

22



**UNIVERSIDAD NACIONAL
AUTONOMA DE MEXICO**

FACULTAD DE INGENIERIA

**INTEGRACION DE INFORMACION
HEMEROGRAFICA Y REFERENCIAL EN REDUNAM-
INTERNET**

T E S I S

**QUE PARA OBTENER EL TITULO DE:
INGENIERO EN COMPUTACION
PRESENTAN:**

**THELMA GUADALUPE DEL CASTILLO MONROY
CARLOS RAFAEL TABOADA GARCIA**

275495

DIRECTOR DE TESIS:

ING. ANTONIO ENRIQUE GONZALEZ VELAZQUEZ

CODIRECTOR DE TESIS:

ING. CARLOS SAUCEDO MACIEL



MEXICO, D. F.

2000



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Doy gracias a Dios por permitirme dar este paso y por darme todo lo que soy y lo que tengo.

Dedico este trabajo a mis padres Lupita y Gabino y les doy las gracias por todo su amor, apoyo y comprensión, por inculcarme los valores para conducirme por la vida con honestidad y dedicación.

Gracias a mis hermanos Katy y Arturo por su ayuda y cariño y por estar conmigo siempre.

A Carlos por todo su apoyo, colaboración y paciencia en todo momento, por su cariño y amor.

A mis verdaderos amigos que me apoyaron en todo momento y me mostraron lo que es la amistad.

Gracias a mis profesores y asesores que compartieron conmigo sus conocimientos y experiencia.

A la UNAM por ser la Institución que me ha dado todo.

Thelma

Dedico el presente trabajo a mis padres, quienes con su ejemplo, dedicación y amor lograron cimentar en mí los valores fundamentales que me han permitido lograr todo lo que soy y tengo.

Agradezco a Thelma por todo su amor, paciencia y apoyo, sin los cuales no hubiera podido sortear todos los momentos difíciles que he enfrentado. Para ti todo mi amor.

Mis más sinceras gracias a mis Asesores de Tesis sobre todo por la paciencia, por su tiempo y por sus consejos. Gracias a ellos, hoy puedo ver este fruto concluido.

Agradezco a mis maestros por haber tenido la paciencia y la dedicación para compartir conmigo algo de su sabiduría y experiencia. A todos ellos gracias.

Pero sobre todo, gracias a Dios por poner en mi camino a todas aquellas personas de las que he aprendido tantas y tantas cosas. Mi agradecimiento a todas ellas.

Carlos

Contenido

Introducción

1. Antecedentes

1.1. Objetivos y funciones de una Biblioteca

- 1.1.1. Historia de la Biblioteca Tradicional
- 1.1.2. Rol de la Biblioteca Tradicional
- 1.1.3. Formas de comunicación en las Bibliotecas Tradicionales

1.2. Impacto de la tecnología en las funciones y actividades de las Bibliotecas

- 1.2.1. Comunicación en la biblioteca actual
- 1.2.2. Redes de computadoras
- 1.2.3. Impacto de las telecomunicaciones

1.3. "Biblioteca Digital"

- 1.3.1. Introducción
- 1.3.2. La infraestructura tecnológica y política
- 1.3.3. Concepto de "lo de afuera hacia adentro"
- 1.3.4. La Biblioteca Digital como un ejemplo de Revolución de Control

1.4. Tendencias mundiales de los servicios que ofrecen las Bibliotecas empleando tecnologías computacionales

- 1.4.1. Biblioteca Digital considerada como Colección Nacional de Información
- 1.4.2. Tendencias
- 1.4.3. Proyectos para las Bibliotecas Digitales
- 1.4.4. La Biblioteca Digital en América Latina. Proyectos Actuales
- 1.4.5. La Biblioteca del Congreso de los Estados Unidos de Norte América
- 1.4.6. Instituto Internacional para la Investigación de la Biblioteca Electrónica
- 1.4.7. Proyectos del Programa de Bibliotecas Digital de Inglaterra
- 1.4.8. Otros proyectos de Biblioteca Electrónica

1.5. Preservación Digital

- 1.5.1. Introducción
- 1.5.2. Proyecto de digitalización de textos de la Biblioteca Nacional de Agricultura de los Estados Unidos de Norte América
- 1.5.3. Proyecto Open Book

1.6. Organización de una Biblioteca Digital Global. Biblioteca del Congreso

1.7. Revistas electrónicas

2. Diseño de un Esquema de Biblioteca Electrónica

2.1. Objetivos de la Biblioteca Electrónica

- 2.1.1. Importancia de la Biblioteca Electrónica
- 2.1.2. Requerimientos de los usuarios y editores para el establecimiento de Biblioteca Digital o Electrónica

2.2. Elementos de una Biblioteca Electrónica

- 2.2.1. Cómo se inicia una Biblioteca Electrónica
- 2.2.2. Principios de implantación
- 2.2.3. Repositorio electrónico
- 2.2.4. Registros electrónicos
- 2.2.5. Colecciones Históricas Digitales: Tipos, Elementos y Construcción
- 2.2.6. Estándares para imágenes digitales
- 2.2.7. Herramientas de acceso y políticas
- 2.2.8. Derechos de Autor

2.3. Técnicas de Automatización y Administración de Bibliotecas Electrónicas

- 2.3.1. Intranet
- 2.3.2. Reingeniería Organizacional
- 2.3.3. Directrices federales
- 2.3.4. Imagen de documentos
- 2.3.5. Sistemas de Administración de Imágenes Electrónicas (EIM)
- 2.3.6. Reconocimiento Óptico de Caracteres
- 2.3.7. Sistemas de Administración de Documentos Electrónicos (EDMS)
- 2.3.8. Flujo de trabajo de documentos
- 2.3.9. Aplicaciones de administración de bibliotecas y la Intranet
- 2.3.10. La gente: la clave para la implantación exitosa de la tecnología
- 2.3.11. Costo y justificación de las tecnologías
- 2.3.12. Conclusión dentro de la automatización de bibliotecas digitales implementadas

2.4. Alcances y limitaciones actuales de la Biblioteca Electrónica en el tratamiento de la información

- 2.4.1. Vías para crear recursos de texto completo
- 2.4.2. Imágenes Digitales
- 2.4.3. OCR
- 2.4.4. Uso del CD-ROM y DVD para almacenar información digital
- 2.4.5. Problemas en el uso de imágenes electrónicas en la preservación de documentos
- 2.4.6. El papel que juega la red de datos
- 2.4.7. Biblioteca Electrónica. Visiones y Realidades

2.5. Publicaciones Seriales Electrónicas en WWW

- 2.5.1. ¿Qué es una publicación serial electrónica?
- 2.5.2. ¿Qué tipo de publicaciones electrónicas seriales están disponibles en el WWW?
- 2.5.3. Similitudes entre publicaciones seriales impresas y electrónicas
- 2.5.4. Diferencias entre las publicaciones seriales impresas y electrónicas
- 2.5.5. La opción de texto completo para publicaciones seriales electrónicas

3. Planteamiento del Problema

3.1. Antecedentes

- 3.1.1. Biblioteca de la Dirección General de Servicios de Cómputo Académico
- 3.1.2. Servicios que ofrece la Biblioteca
- 3.1.3. Tipo de Usuarios de la Biblioteca de DGSCA
- 3.1.4. Horario de servicios de la Biblioteca
- 3.1.5. Políticas de préstamo
- 3.1.6. Personal que labora en la Biblioteca

3.2. Planteamiento del problema

3.3. Objetivos

4. Análisis del sistema

4.1. Identificación de necesidades

4.2. Recursos humanos y de cómputo

4.3. Estudio de viabilidad del servicio

- 4.3.1. Viabilidad económica
- 4.3.2. Viabilidad legal

4.4. Análisis de requerimientos

- 4.4.1. Análisis de la Información
 - 4.4.1.1. Contenido de la información
 - 4.4.1.2. Flujo de la información
- 4.4.2. División del sistema y descripción funcional de cada modulo

4.5. Análisis orientado al flujo de datos

- 4.5.1. Módulo de Incorporación y actualización de información al acervo de revistas
- 4.5.2. Módulo de Publicación en WWW del acervo de revistas
- 4.5.3. Módulo de Búsqueda en WWW de información en el acervo de revistas
- 4.5.4. Modulo de Envío de los artículos a los usuarios finales

4.6. Análisis técnico

- 4.6.1. Elección del servidor WWW
- 4.6.2. Empleo de CGI (Common Gateway Interface)
 - 4.6.2.1. Puentes (Gateways)
 - 4.6.2.2. Programando CGI's
 - 4.6.2.3. Consideraciones para la programación de CGI's
- 4.6.3. Protocolo Z39.50 para información bibliográfica y hemerográfica
 - 4.6.3.1. Breve historia de Z39.50
 - 4.6.3.2. ¿Qué es el protocolo Z39.50?
 - 4.6.3.3. ¿Qué hace Z39.50?
 - 4.6.3.4. Empleando Z39.50 sobre Internet
- 4.6.4. Introducción al SSL y certificación utilizando SSLey
 - 4.6.4.1. Introducción
 - 4.6.4.2. Técnicas criptográficas
 - 4.6.4.3. Certificados
 - 4.6.4.4. SSL
- 4.6.5. EL ISSN (International Standar Serial Number) para identificación de publicaciones seriales
 - 4.6.5.1. Aspectos generales del ISSN
 - 4.6.5.2. Caso de estudio en México
- 4.6.6. SGML y los Sistemas de Biblioteca en Línea
 - 4.6.6.1. Del Marc a las etiquetas, SGML y los Sis temas Biblioteca en Línea

5. Desarrollo del sistema

5.1. Diseño de datos

- 5.1.1. Identificación de los nombres de las publicaciones seriales de la biblioteca de DGSCA.
- 5.1.2. Formato de las tablas de contenido de las revistas

5.2. Diseño Arquitectónico y Procedural

- 5.2.1. Estructura del programa
 - 5.2.1.1. Diagramas de Flujo de datos
- 5.2.2. Estructura de datos
 - 5.2.2.1. Diccionario de datos
 - 5.2.2.2. Representación para el sistema de las revistas y sus tablas de contenido.
 - 5.2.2.3. Datos de salida

5.3. Implantación

- 5.3.1. Configuración de servicios en red
- 5.3.2. Codificación
- 5.3.3. Mantenimiento
- 5.3.4. Costos asociados

5.4. Pruebas

6. Conclusiones

Apéndice A

Apéndice B

Apéndice C

Glosario

Referencias

INTRODUCCIÓN

INTRODUCCIÓN

Como es sabido, en el siglo que está por terminar se aceleró exponencialmente la inventiva humana sobre todo en su última mitad. Los inventos se han sucedido aceleradamente en los campos de la ciencia y particularmente en el de la computación, para que con el transcurso de su desarrollo se han podido conformar, por su trascendencia, dos vertientes fundamentales, una, la de la informática, al integrar a la computación y a las telecomunicaciones y la segunda, la automatización industrial. La primera, desemboca en la ahora llamada Era de la Información, pues reconoce las transformaciones en los procesos internos que reforman profundamente a la organización, administración y servicios de prácticamente todas las organizaciones de todos los sectores de la economía.

Es de esperarse que esta carrera desenfadada en busca de innovación continúe su marcha, pues de esa misma manera se presentan las crecientes necesidades de satisfactores de todo tipo para la aún ascendente magnitud de la población humana.

Es evidente que el buen uso o mal uso de la computadora tendrá influencia definitiva en el futuro de las organizaciones.

La interpretación de las realidades y de los avances tecnológicos será la base para visualizar al futuro y dar el impulso para la selección de los caminos que se deben seguir, para que con ello, se logren organizaciones y servicios más modernos y eficientes, logrando entregar los bienes que la sociedad requiere.

Es indudable que hacia cualquier campo de la tecnología de la información que volteemos la mirada, seguramente nos sorprenderemos al observar sus acelerados cambios y mejoras. La siempre interminable ruta de la miniaturización, la inimaginable velocidad de procesamiento, la multiplicación de las capacidades de almacenamiento, las interfaces hombre máquina con lo llamado multimedia que hasta hace pocos años solamente se soñaba, la automatización y operación de procesos, son campos que día con día se suman a otros antes aparentemente ajenos a los dominios de la Tecnología de la Información.

Actualmente, la tecnología, y particularmente el cómputo y las telecomunicaciones han influido de manera definitiva la mayoría de las actividades humanas y la forma de interrelacionar a los individuos, así como las maneras de intercambiar información, conocimiento e ideas. Muchos de los proyectos de investigación y de educación serían muy complicados de ser llevados a cabo o serían impensables si no existieran las computadoras y las redes informáticas (principalmente Internet).

La aparición y rápida aplicación de tales tecnologías en los ámbitos más diversos del quehacer humano han llevado a tener que replantear de manera importante muchas actividades bien establecidas y tradicionales.

Las bibliotecas, en su papel tradicional como centros de acopio de información y de conocimiento y cuya importancia en el desarrollo de actividades de investigación, educación y difusión del conocimiento ha sido evidente al poner a disposición de los usuarios sus acervos para consulta y préstamo, han tenido que plantear nuevas formas de trabajo, de administración, de servicios internos y externos, así como de atención a los usuarios.

La información debe estar en cualquier lugar y a cualquier hora que se requiera, es una de las funciones más importantes de los sistemas de información hoy en día. Así que las bibliotecas, al

publicar sus acervos a través de redes de comunicaciones (como Internet), plantean la necesidad de que los servicios de la biblioteca estén a disposición del público más allá de los horarios de atención normales, además de necesitar atender a una población geográficamente dispersa y con necesidades distintas.

La sociedad actualmente está consciente de los avances tecnológicos y exige cambios en la forma de adquirir, consultar y recabar información, así como en la calidad de la misma. Además, en algunos sectores ha sido particularmente rápida la adaptación a estos nuevos esquemas y cambios. Precisamente por contar con estas nuevas alternativas en la adquisición y distribución de información y conocimiento, la sociedad exige que los medios y centros de acopio de información se mueva en el mismo sentido. Los medios electrónicos, las redes de computadoras han representado herramientas adecuadas para cumplir con este punto.

Las bibliotecas del futuro, además de otras organizaciones aplicarán la tecnología como un factor básico para lograr los objetivos y metas planteadas en la forma más rápida y eficiente posible. En el caso de la biblioteca, el llevar el conocimiento a los usuarios a través de los medios convencionales y de medios electrónicos. Por sus características, es indudable que Internet es una de las herramientas más importantes a ser utilizada para lograr tales objetivos. El público de las bibliotecas del futuro será mundial y el acceso y atención deberá de ser permanente (24 horas al día siete días a la semana).

El presente trabajo tiene como finalidad por una parte establecer una serie de elementos y esquemas para la incorporación de acervos bibliográficos y hemerográficos a medios electrónicos (principalmente redes TCP/IP, tal como Internet), estableciendo un punto de referencia para la comunidad interesada sobre el análisis, metodología, programas y herramientas necesarias para lograr presencia en la red Internet a través de la publicación de su acervo en el tiempo más corto posible.

Las herramientas aplicadas en este proceso son de dominio público, ya que se parte del hecho de que las bibliotecas (muchas de ellas pequeñas) no cuentan con los recursos económicos necesarios para establecer una infraestructura de cómputo exclusiva y, sobre todo, para adquirir una aplicación comercial para la administración y publicación de su acervo. Además, se trata de un esquema y metodología desarrollado pensando para ser multiplataforma, en el sentido de que no importando el equipo y la arquitectura de la infraestructura de cómputo, si se cuenta con un sistema operativo UNIX y una red basada en TCP/IP, es posible implementar este servicio.

Se presenta un caso práctico que se refiere al establecimiento del servicio de consulta al acervo hemerográfico de la Biblioteca de la Dirección General de Servicios de Cómputo Académico (DGSCA), en donde se han colocado a disposición del público las tablas de contenido de publicaciones seriales recibidas por la biblioteca.

El esquema propuesto, contempla no sólo la publicación de las tablas de contenido, sino con la capacidad de presentar una versión del texto completo de los artículos y hasta la presentación de las imágenes digitalizadas de los mismos. La incorporación de tales elementos puede ser llevado a cabo por etapas bien definidas.

Tratando de sintetizar, podemos decir que la finalidad de todo este nuevo camino ha sido la de aspirar: al mejor producto o servicio, con el menor costo, en el menor tiempo, en el lugar que es demandado, en el tiempo oportuno, con el menor esfuerzo

CAPÍTULO 1

CAPÍTULO 1

ANTECEDENTES

1.1. OBJETIVOS Y FUNCIONES DE UNA BIBLIOTECA

1.1.1. Historia de la biblioteca tradicional¹

La biblioteca tradicional es definida como el local en donde se guardan libros convenientemente ordenados para la lectura y, el conjunto o colecciones de ellos. La historia de éstas colecciones tiene su origen en la más remota antigüedad. Las primeras referencias documentadas que poseemos se refieren a las bibliotecas de los hebreos, egipcios y asirios. Es seguro que los primeros, además de la Biblia "Libro de Libros", poseyeron otras obras referidas por los comentaristas, que debieron conservarse en los templos y palacios custodiados por los sacerdotes que, mas que bibliotecarios, eran guardianes de los textos. También los egipcios tuvieron importantes bibliotecas, y algunos jeroglíficos hacen mención de personas que desempeñaban el cargo de bibliotecarios. Éstos debieron ser los encargados del cuidado y conservación de las enormes colecciones de rollos de papiro, alguna de las cuales, como la llamada de los libros de Toth, se afirma que ascendía a más de treinta mil piezas. Eustacio, el gramático, hace referencia categórica a la gran biblioteca del templo de Menfis, y en la casi totalidad de las tumbas descubiertas en Egipto se han encontrado rollos de papiro que constituían la biblioteca particular del difunto. Entre los Asirios de Nínive aparecen algunas de cuya existencia ha quedado prueba irrefutable. Tal es la biblioteca y archivo del rey Arsubanipal, formada por una extensa colección de ladrillos de arcilla escritos con caracteres cuneiformes, cocidos después para garantizar su durabilidad. Este inapreciable monumento, actualmente en el Museo Británico de Londres, nos permite apreciar cuál era la clase de textos que se coleccionaban en tan remota época, pues junto a documentos comerciales y estadísticos aparecen obras científicas y literarias de gran valor arqueológico. En Sipparce, "la ciudad de los libros", próxima a Bagdad, ha sido encontrada otra biblioteca formada por más de cincuenta mil piezas, y no lejos de allí, Smith descubrió a finales del siglo pasado, las famosas tablillas donde se describe una versión del diluvio universal. Incomparable entre todas las bibliotecas de la antigüedad clásica, es la famosísima Alejandría. Fundada por Tolomeo Soter y continuada por su hijo, Tolomeo Filadelfo, que la llevó a su grado máximo de esplendor, tuvo como asiento el Museion, famoso palacio de las musas, erigido por el primero de los dos sátrapas nombrados. Conforme a la clásica descripción, ocupaba todo el fondo del Museion y se componía de amplias salas y estancias donde los amanuenses, iluminadores y encuadernadores, en número más de cinco mil, se dedicaban a la copia, decoración y guarnecido de los quinientos mil volúmenes que llegaron a albergarse en ella. Calimaco de Cirene la clasificó y catalogó conforme al siguiente procedimiento: en una tira de pergamino que se pegaba en el exterior de cada rollo, se anotaba el nombre del autor o autores, agregándose su biografía y el número de líneas de que se componía el texto. Trabajo este tan paciente como impecable para la historia de la cultura, pues gracias a él sabemos de muchos tesoros que contuvo aquella biblioteca hoy desaparecida por obra de un destino que parecía perseguirla implacablemente.

¹ Referencia de la Enciclopedia Ilustrada Cumbre (Historia de la Biblioteca)

Castigada primero por el saqueo de las tropas de César; rehecha por la donación que del fondo de la biblioteca de Pérgamo hizo Marco Antonio a Cleopatra, desapareció devorada por las llamas como consecuencia de la conquista de la ciudad por las hordas de Omar a las ordenes del fanático Amru. Un cronista de esta espantosa catástrofe relata la respuesta del caudillo a la súplica del gramático que le pidió la entrega de la biblioteca: "Si estos escritos están en el Corán, son inútiles; y si no lo están, deben desaparecer". Como consecuencia del bárbaro dilema, los inapreciables papiros fueron distribuidos entre las termas de la ciudad, donde se utilizaron como combustible para calentar el agua durante más de medio año.

La citada biblioteca de Pérgamo llegó a contar con más de doscientos mil volúmenes escritos en un material especial, el pergamino o pergamena, inventado en esta ciudad que le dio su nombre. La historia de este invento es aleccionadora para quienes tratan de poner obstáculos al afán de cultura, pues el rey de Pérgamo, deseoso de formar una gran biblioteca, se enfrentó con la prohibición de exportar papiro, decretada por Tolomeo, decidido a que no hubiera otra biblioteca que la suya. Mas la necesidad agudizó el ingenio y obtuvo la preparación del pergamino, vehículo de escritura insuperado durante milenios. Entre los griegos, la primera biblioteca pública que se recuerda es la de Atenas, fundada por Pisístrato, que fue saqueada por Jerjes y devuelta luego a la ciudad. De las bibliotecas particulares de los griegos, se tiene referencia escrita de las que poseyeron Euclides, Eurípides, Demóstenes, Aristóteles y otros; la de este último se hizo famosa por el sistema de clasificación ideado por el filósofo, el primero que haya tenido base científica. Las primeras que aparecieron en Roma fueron las que llegaron como botín de guerra acumulado por los conquistadores durante los últimos años de la República; entre ellas era importante la de Cártago, ofrenda de Escipión el Africano. A estas primeras bibliotecas conseguidas manu militari, hay que agregar la colección de textos griegos recogida por Lúculo durante su campaña de Oriente, con los que formó la primera biblioteca en Roma, que estuvo a disposición de los literatos y amigos de su fundador. Muy famosa también es la organizada por Tiranión el Gramático, que acrecentada posteriormente con la donación de la de Aristóteles que Sila le hizo, llegó a contar con más de treinta mil volúmenes. La idea de Julio César de establecer en la capital del imperio una gran biblioteca dividida en dos secciones, griega y romana, dirigida por Terencio Varrón, fue llevada a cabo, después del asesinato del dictador, por Asinio Polión, quien utilizó para ello el recientemente restaurado Atrium Libertatis vecino al foro romano. El fondo de esta biblioteca quedó constituido por los volúmenes robados a los partos, cuatro décadas antes de nuestra era. Después de esta primera biblioteca vienen las dos fundadas por Augusto, la Palatina y la Octaviana, dividida ésta en las dos secciones, griega y latina. A estas dos bibliotecas destruidas por el fuego, sucedieron la Augusta, la del Templo de la Paz, erigida en conmemoración de la conquista de Judea, y la Ulpiana, fundada por Trajano, el emperador romano de origen hispánico, única que perduró hasta pasados quinientos años de nuestra hora. En Roma llegaron a funcionar veintiocho bibliotecas públicas y era raro el ciudadano distinguido que no poseyese su propia colección de libros. Entre las bibliotecas particulares más afamadas de la Roma clásica estaba la de Cicerón. Las bibliotecas romanas se instalaban en salas no demasiado grandes, lujosamente decoradas y confortables, y a las que los huecos abiertos hacia Oriente dotaban de luz suave e igual, matizada por el tono de los mármoles elegidos para su recubrimiento, siempre verdes, por ser el color menos perjudicial para la vista. Los volúmenes se encontraban en estantes adosados a las paredes o colocados en el centro de las estancias, protegidos por cajas o fundas de pergamino. Todas ellas estaban dirigidas por funcionarios especializados. Al trasladarse a Bizancio (Constantinopla) la sede del Imperio Romano, se estableció en la nueva capital una biblioteca formada por las obras cristianas salvadas de la persecución decretada por Diocleciano. Los emperadores Juliano y Teodosio aumentaron sus tesoros hasta hacer de ella la más famosa de su tiempo; gracias a esta colección, los textos clásicos, la mayoría de los conocimientos de la

antigüedad, después de la conquista de Constantinopla por los turcos, pasaron en el bagaje de los humanistas griegos refugiados en Italia, a constituir los pilares de la cultura del Renacimiento.

Durante la propagación del cristianismo hubo también famosas bibliotecas, como la de Pánfilo de Cesárea, citada varias veces por San Jerónimo, víctimas todas del furor de las invasiones bárbaras. En la España visigoda, más tolerante, se salvaron algunas hispanorromanas cuyos volúmenes encontraron asilo en diversos conventos y monasterios: la de Valclara; la del rey Recesvinto; la famosa de Toledo, y la de Sevilla, obra de San Isidoro. Todas ellas, resto de las grandes colecciones romanas y griegas, perduraron, apenas hojeadas, en el fondo de frías estancias donde el polvo, la incuria y la agresión de los insectos y roedores, cuando no la mordedura del fuego, las sometieron a una continua merma de la que no se han salvado más que escasos ejemplares restaurados por obra de la paciencia benedictina de los copistas de la Edad Media. Con el florecimiento de la cultura de los árabes, las bibliotecas alcanzan uno de los puntos culminantes de su desarrollo. Bagdad y Córdoba, sedes de los dos grandes califatos de Oriente y Occidente, tuvieron una importancia trascendental en la historia de la cultura. La Andalucía árabe llegó a contar con más de setenta bibliotecas, con cerca de trescientos mil manuscritos arábigos. En el otro lado del estrecho de Gibraltar era también muy crecido el número de ellas, entre las que destacaremos la de la Academia de Trípoli, que poseía más de tres millones de volúmenes cuando fue incendiada por los cruzados; la de El Cairo, con un millón cien mil manuscritos, y la casi fabulosa de Bagdad, de la que se cuentan que era tal la cantidad de libros conservados en ella, que cuando los mongoles asaltaron la ciudad, los arrojaron al Tigris para formar un dique que les permitió cruzar su ancho y profundo caudal a pie.

Bibliotecas medievales

Durante la Edad Media, fue en los monasterios y abadías donde se conservaban los libros, reliquias del pasado. De estas rarísimas bibliotecas, son famosas la del monasterio del monte Athos, en Grecia; las de los monasterios de Fulda, dividida en cuarenta y ocho secciones; Hildesheim y Ratisbona, en Alemania; la celeberrima de San Gall, en Suiza; las de Cluny, Fleury, Besanzón, San Germán de Auxerre y San Germán des Pres, en Francia; la de la abadía de Montecassino en Italia; las de Glembourg y Lieja, en Bélgica; las de Wearmouth y Canterburry, en Inglaterra; y las de Albeda, Silos Sahagún, Arlanza, Cardeña, San Millán de la Cogulla y Valbanera, en España. Durante la Edad Media, la administración de las bibliotecas se resintió de cierta falta de rigor. Los libros se depositaban sin ningún orden, tendidos sobre mesas o colgados sobre atriles a los que estaban unidos con cadenas, como pueden verse todavía en diversos monasterios europeos. Los libros así dispuestos se denominaban libri catenati o barbetos, y lo estaban así para evitar las sustracciones.

El Renacimiento

Con la llegada del Renacimiento, las bibliotecas adquieren un renovado impulso. Son famosas las de San Marcos en Venecia, primera que tuvo un edificio exclusivo para ella, formado con el legado de la biblioteca particular del príncipe de los humanistas, Francisco Petrarca, quien la dejó con la condición de que estuviera al público; la de los ermitaños de Florencia, legada con igual fin que la anterior por Juan Boccaccio; la famosa Medicea, fundada por Lorenzo de Médicis y dirigida por el gramático griego Juan Lascaris, que pasó posteriormente a poder del cardenal de la Rovere, y la de

Urbino, base de la hoy inapreciable biblioteca Vaticana de Roma. En la Francia anterior a los días de la fundación de la Soborna, cuya biblioteca contaba en tiempos de San Luis con mil volúmenes, 4 fueron famosas las bibliotecas de San Mauricio, en Valais, y la de Tours y la de San Dionisio. En Inglaterra, la de Oxford fundada por Ricardo de Burry, y la de la biblioteca Bodleiana, conocida hasta 1550 como Biblioteca del Duque de Gloucester, reorganizada y enriquecida por el diplomático inglés sir Tomás Bodley, a quien debe su nombre actual. Contiene un millón y medio de volúmenes y cuarenta mil manuscritos, en su mayor parte originales de los más grandes escritores ingleses. Alejado de la corte isabelina por las antipatías del conde de Essex, sir Tomás Bodley se refugió en Oxford, dedicándose a restaurar la biblioteca quemada por la saña anticatólica de los ministros de Eduardo VI, y la dotó de ochenta y cuatro mil volúmenes y de una renta de doscientas libras para sueldos de bibliotecarios. Hoy las leyes inglesas protegen su incremento.

En Europa cabe destacar, además, las bibliotecas de Praga, Heidelberg, Erfurt, Danzing y Budapest; en España, las de Salamanca, fundada por Alfonso el Sabio, la Rabinica de Zaragoza y la de Almanzor, en Córdoba, cuyos restos se conservan hoy en El Escorial; la de la Universidad Complutense, fundada por Cisneros, y la de Poblet. Con la invención de la imprenta de tipos móviles se inicia un desarrollo definitivo de las bibliotecas, cuyo crecimiento ha tropezado siempre con la intolerancia, enemiga eterna de la cultura. Ejemplos irreparables de estas destrucciones han sido la de la biblioteca de Alejandría por romanos y árabes; la de Constantinopla, aventada a la caída de la ciudad en poder de los turcos; la de Córdoba, destruida por los católicos, al igual que la de Trípoli, incendiada por los cruzados; la de Bagdad, arrojada al Tigris por los mongoles; la de Cluny, víctima del fanatismo de los protestantes; la de Heilderberg, requisada por el Vaticano y robada por Napoleón; la de Poblet, quemada durante la guerra carlista española; los millones de volúmenes desaparecidos en la Primera y en la Segunda Guerra Mundial. En las bibliotecas del Renacimiento y épocas siguientes a la invención de la imprenta, la colocación y la clasificación de los libros, si bien progresó algo, no avanzó demasiado, pues aún aparecían colocados en grandes atriles o en estantes con los títulos sobre el lomo o el canto, según cual fuera de los dos el que miraba hacia el exterior. Ejemplos de libros con el título sobre el canto dorado, escrito con letras negras, se hallan en la famosa biblioteca del Escorial, en España. Las estanterías y muebles destinados a albergar los libros poseían más bien un carácter decorativo que práctico, y su ornamentación solía estar de acuerdo con el resto del estilo que imperaba en el edificio. Ya muy avanzado el siglo XVIII empiezan a aparecer, en las estancias destinadas a biblioteca, muebles contruidos con este exclusivo fin y sobre los que solían colocarse los libros a gusto del dueño o simplemente por orden alfabético.

Con la llegada de la época contemporánea se inicia el intento de instalar las bibliotecas en edificios capaces de contener mayor o menor cantidad de volúmenes. Esta innovación dio lugar a la creación de bibliotecas de tipo central y de otras destinadas a depósito o almacén. La primera de tipo central, construida en Francia, es la de Santa Genoveva de París, obra de Labrouste, cuya descripción bastará para darse una idea suficiente del propósito que animó a sus creadores. Erigida de acuerdo con las ventajas que proporcionaba el empleo del hierro como material de construcción, su planta se reduce a la creación de una fábrica de dos pisos, destinado el inferior a la administración y a la sección de duplicados, estampas y manuscritos; y el superior, formado por una sala de gran altura rodeada de estanterías, con pupitres para los lectores, y una tarima central desde donde atienden los bibliotecarios y mozos de servicio. Este tipo de biblioteca, con salón central de lectura y depósito de libros alrededor, ha sido ya desechado por los siguientes inconvenientes fáciles de advertir: imposibilidad de ampliación, dificultades de limpieza, incomodidad para el lector producida por las actividades de los empleados y necesidad casi continua de luz artificial, etc. En las bibliotecas - almacén o depósitos se trató de eliminar los defectos antes señalados, procurando aislar completamente al lector, creando al efecto secciones independientes dentro del ámbito del edificio

(raros, incunables, manuscritos, estampas, mapas), conforme a diversas ramas de una misma actividad (poesía, teatro, novela, ensayo) y de acuerdo con una clasificación más general (arquitectura, escultura, pintura, grabado), estableciendo grandes grupos (literatura, ciencias, artes) o simplemente para destacar una especialidad de la institución ("Cervantes", "Cuentistas", "Americanos", "Deportes") con objeto de situar al lector dentro de un medio acorde con la lectura de su predilección. La primera biblioteca - almacén construida fue la del Museo Británico de Londres, considerada también como la más perfecta hasta la erección de la del Congreso, en Washington, obra de Lincoln Casey, y quizás la más rica y lujosa del mundo. En ésta última el lector tiene a su disposición toda clase de servicios (cafetería, restaurante, baños) y asesoramiento por parte de su numeroso personal (ficheros de otras bibliotecas, repertorios bibliográficos, servicios de copias en microfilm) así como un fondo de consulta que pasa los diez millones de volúmenes. Todo el movimiento administrativo funciona con independencia del público, que así puede dedicarse tranquilamente a la lectura, sin la molestia que supone el continuo ir y venir del personal o la entrada y salida de un público siempre renovado. Con los ficheros a su disposición, el lector no tiene más que hacer su pedido para que el libro requerido, por los procedimientos más veloces (tubo neumático, teléfono, ascensor o plano inclinado) esté lo más pronto posible a su disposición. En estas bibliotecas todo es silencio, sosiego y recogimiento, y el lector ignora lo que pasa más allá de la pared que le separa del almacén, que da nombre a las mismas. Los almacenes unidos con los departamentos de entrega y devolución por los veloces medios antes indicados, no son otra cosa que naves de diversas amplitudes, dentro de las que albergan los pisos de estanterías donde están depositados los libros. Las estanterías, siempre de hierro para evitar en lo posible el peligro del incendio, tienen una altura fija de dos metros y medio, de forma que el encargado de tomar un libro pueda hacerlo siempre sin ayuda de escalera o banquillo alguno. La limpieza, desinfección y sistemas de alarmas de incendio son siempre de los más modernos, así como el alumbrado y calefacción, cuyo aislamiento es casi siempre perfecto. Con esto se evita cada vez más el peligro del fuego, terrible enemigo que, pese a las precauciones adoptadas, suele producir desastres tan graves como el ocurrido en Lima (Perú), cuya biblioteca fue completamente devorada por las llamas. Dada la altura de los pisos de las estanterías y separación existente entre ellas, es fácil comprender la cantidad de volúmenes que pueden almacenarse, convenientemente clasificados y prontos para su rápido manejo, en la menor cantidad de espacio posible. Si se une a esto el servicio de reproducción en microfilm, existente en la mayoría de las grandes bibliotecas actuales, cuyo fin principal es el de combatir la invasión del espacio disponible por el fabuloso crecimiento de las colecciones de diarios y revistas, se advertirá que la unidad de estos edificios está garantizada para un largo número de años, pese al continuo y casi mítico producir de las prensas del mundo. Especialidades existentes en estas y otras bibliotecas son las salas de lectura para ciegos; salas de textos musicales, con instrumentos y discoteca a disposición del estudioso; salas de lectura infantil, etc. Variante muy curiosa adoptada últimamente por diversas bibliotecas de todo el mundo, es la de la Biblioteca de Nueva York. En esta hay dos grandes secciones: la biblioteca propiamente dicha, con más de seis millones de libros de fondo, y la biblioteca de préstamo, con dos millones de libros y más de cincuenta sucursales. De entre los servicios de biblioteca de préstamo a establecimientos de enseñanza alejados de los centros de población, es de señalarse aquí que suele llevarse a cabo en algunas naciones por intermedio de misiones pedagógicas. Uno de los medios más eficaces con que cuentan esas misiones culturales, es la de dejar en cada sitio de actuación una biblioteca modelo de doscientos a cuatrocientos volúmenes, una discoteca también seleccionada con su correspondiente aparato reproductor y un proyector cinematográfico, dotado de un cierto número de películas cómicas y educativas. Estando además esas bibliotecas compuestas de colecciones de volúmenes no coincidentes entre sí, basta una simple permuta entre dos vecinas, para renovar por un cierto tiempo el fondo de lectura local. Semejantes en cierto modo a esas bibliotecas que hoy pueden verse en buques, ferrocarriles y hasta en camiones y aviones.

Catalogación de bibliotecas

Una biblioteca sin clasificación es como un cuerpo sin sentidos. En ella puede haber de todo, pero ¿cómo buscarlo?. El conocimiento de lo que los libros contienen es ya en sí un gran saber. De la importancia concedida a la catalogación por el hombre, a lo largo de la historia de la cultura, es muestra suficiente el que los sistemas de catalogación inventados desde Aristóteles hasta nuestros días pasen de varios centenares. Conrado Gessner, humanista, naturalista y médico suizo, a quien con razón se califica de padre de la bibliografía moderna, fue el primero en crear un sistema de catalogación bibliográfica, científico y racional en el que se procura conciliar la tradición escolástica con las exigencias del progreso suscitado por el Renacimiento. Para ello divide las artes y las ciencias en preparatorias y substanciales. Y las primeras, a su vez, en de adorno y necesarias, y éstas en parlantes.

Un gran polígrafo español, Nicolás Antonio, en el Índice de Materias de su Biblioteca Nova ordenó los libros en los doce grupos siguientes:

I	Teología	VII	Traducciones
II	Filosofía	VIII	Humanidades, Gramática, Filología y Lógica
III	Medicina	IX	Historia
IV	Derecho Canónico y Civil	X	Poesía
V	Política y Economía	XI	Miscelánea
VI	Matemáticas	XII	Literatura

La clasificación decimal

Los dos sistemas bibliográficos más completos son los de Brunet y Dewey. De ellos, aunque el primero esté en uso en algunas bibliotecas, daremos únicamente el segundo, que es el ya adoptado por todas las grandes bibliotecas del mundo y el aceptado por el Instituto Internacional de Bibliografía, con sede en Bruselas (Bélgica). Con este ingenioso método, llamado de clasificación decimal, se puede llegar al ideal en cuanto a organización de bibliotecas se refiere, que es la unificación de la signatura de los ficheros, pudiendo saberse, con sólo una mirada al número de la signatura, si en esa biblioteca se encuentran obras referentes al tema deseado, por muy especializado que sea. Mediante la utilización del sistema decimal han quedado corregidos los inconvenientes de la clasificación alfabética, que presupone el conocimiento de todos los idiomas por el lector, y de la cronológica, que supone, asimismo, el conocimiento de la disciplina por el que acude en busca de los textos. Otra ventaja del sistema decimal, tan importante como fácil de comprender, es que, al ser uniforme la signatura de las fichas de todas las bibliotecas del mundo, llegará el día en que, mediante un intercambio de fichas fácilmente realizable por duplicado, podrá saberse en qué lugar del mundo está la biblioteca que posee el libro deseado, ahorrándose así la inmensa pérdida de tiempo que costaba antes la sola localización de un título. Base de esta aspiración es el gran fichero, poseedor de más de treinta millones de fichas, del citado Instituto Nacional de Bruselas, cuyo aporte al progreso de la cultura universal ha dado ya frutos evidentes.

El método de clasificación decimal, creado por el bibliotecario inglés Melvil Dewey, reparte las ramas del conocimiento en las 10 siguientes secciones:

0	Generalidades	5	Ciencias Puras
1	Filosofía	6	Ciencias Aplicadas
2	Religión	7	Bellas Artes
3	Ciencias Sociales y Derecho	8	Literatura
4	Filología	9	Historia y Geografía

Cada una de estas grandes divisiones iniciales se encuentra subdividida en otros diez grupos, que pueden subdividirse hasta conseguir un número de apartados que permitan la exacta clasificación de toda obra sin error posible. Así tenemos que, por ejemplo, la Sección 3 (Ciencias Sociales y Derecho) se divide en las siguientes diez secciones generales:

30	Sociología General	35	Administración Pública
31	Estadística	36	Beneficencia Social, Seguros
32	Política	37	Educación y Enseñanza
33	Economía Política	38	Comercio, Transporte
34	Derecho	39	Etnología, Costumbres, Usos

A estos grandes grupos, se puede aplicar otra división decimal, que bastará para dar una idea completa de la infinidad de matices logrados por este procedimiento. Si tomamos el punto de la anterior clasificación, número 33 (Economía Política), veremos que se subdivide en:

330	Nociones generales	335	Socialismo
331	Trabajo y trabajadores	336	Finanzas Públicas
332	Economía financiera, Bancos, Moneda, Crédito	337	Aduanas. Proteccionismo. Libre cambio.
333	Propiedad y sus diferentes formas	338	Producción de las riquezas
334	Cooperación, Mutualidad	339	Reparto y consumo

La clasificación general de Dewey constituye un sistema muy aproximado a la perfección; todas las subdivisiones se realizan agregando números dígitos del uno al nueve, separados de tres por tres por un punto que carece de significación especial. Así, pues, un número clasificador abarca toda una materia, y los temas analíticos que de ella se derivan están representados por un número clasificador, y todos los números clasificadores constituyen lo que se llama "notación del sistema".

Para redactar correctamente una ficha bibliográfica corriente, los manuscritos, incunables, textos orientales, etc., son catalogados por especialistas que tienen en cuenta estos dos aspectos principales:

1. Consignar el nombre del autor con todos sus apellidos y empezar por ellos (por ejemplo: Cervantes Saavedra, Miguel de)

2. El título de la obra, al que hay que agregar siempre la edición con su fecha correspondiente. Por ejemplo: Quijote de la Mancha (El Ingenioso Hidalgo): y en la ficha duplicada El Ingenioso Hidalgo don Quijote de la Mancha.

Ambas con el pie de edición. Así, pues, tenemos:

Ficha por autor:

*CERVANTES SAAVEDRA, Miguel de
El Ingenioso Hidalgo don Quijote de la Mancha
Buenos Aires, Ediciones Jackson
1950, 740 páginas*

La del mismo libro por título sería:

*El Ingenioso Hidalgo don Quijote de la Mancha
Buenos Aires, 1950, 740 páginas
Ediciones Jackson
V. Cervantes Saavedra, Miguel de*

O bien:

*Quijote de la Mancha (El ingenioso Hidalgo don)
Buenos Aires, 1950, 740 páginas
Ediciones Jackson
V. Cervantes Saavedra, Miguel de*

Se tienen, pues, tres tipos de referencia, por autor, por título completo y por título más común en la memoria del lector medio. Estos datos son más que suficientes para que el lector encuentre el título deseado en el fichero. Cada ficha debe llevar en una de las esquinas superiores, el número de clasificación que corresponda al libro. Ese mismo número, bien visible, deberá aparecer en el lomo del libro. Así se establece la relación entre el libro, colocado en su estante respectivo, entre otros miles de volúmenes, y la ficha clasificada en las gavetas del fichero. Por cada libro puede haber, además de las fichas ya mencionadas, la ficha de clasificación por materia y la ficha analítica.

Las principales colecciones

Las bibliotecas más importantes del mundo, en la actualidad, son difíciles de indicar. Hay bibliotecas famosas por un solo libro, como la de la Rávena, que posee el código de Aristófanes, en que se encuentran las once comedias que han llegado hasta nosotros; en tanto que otras bibliotecas, que almacenan centenares de miles de libros, nada de irremplazables pueden ofrecer. Por el número de los volúmenes en ella almacenados, el valor de las mismas y su perfecta organización, aparece al frente de las bibliotecas públicas la ya varias veces citada del Congreso de Washington, cuyo valor en cifras es fácil de apreciar: diez millones de volúmenes; ciento cuarenta y siete mil

tomos de periódicos y revistas encuadernadas; catorce millones de manuscritos; dos millones trescientos mil mapas y vistas; cien mil rollos de micro películas; dos millones de partituras musicales; dos millones doscientas mil fotografías y negativos. Siguen en importancia en cuanto a número, la Pública de Nueva York, la del Museo Británico de Londres, la Nacional de París, la de Viena, la de Leningrado y la de Roma. Por el valor de los libros conservados en ellas, son famosas las bibliotecas, ya citadas algunas, Laurentiana, de Florencia; Vaticana de Roma, Ambrosiana de Milán, de San Marcos, de Rávena, Borbónica de Nápoles, y la de la Catedral de Verona, en Italia; la Nacional Mazariana, la del Arsenal, de Santa Genoveva, de Montpellier y de Besanzón, en Francia; las de Berna, Ginebra, Zurich, Saint-Gall y Basilea, en Suiza; la Británica, de Oxford y Cambridge, en Inglaterra; la de Leyden en Holanda; la de Bruselas, en Bélgica; las de Heidelberg, Berlín, Leipzig, y Munich, en Alemania, aunque sufrieron los estragos de la guerra; las de Copenhague, Viena, Leningrado, Moscú, Madrid, Toledo, Sevilla, y el Escorial en el resto de Europa; y las de Harvard, Yale, Hispanic Society, Washington y Sociedad Geográfica Americana, en los Estados Unidos de América, poseedoras todas de textos únicos y libros de gran valor bibliográfico.

Bibliotecas Especializadas

La multiplicidad de ediciones y números siempre en aumento de obras de especialización, han convertido muchas de las primitivas secciones de las grandes bibliotecas en instituciones aparte, en las que el especialista o el simple lector puede encontrar a mano el tipo de lectura o información que necesita. Entre esta nueva clase de bibliotecas hay que distinguir, aparte de las pertenecientes a centros de especialización, y cuyo sólo nombre basta para caracterizarlas, las bibliotecas de ingeniería, arquitectura, bellas artes, medicina, botánica, etc., otras que cumplen fines educativos, sociológicos y hasta complementarios del complejo vivir del hombre moderno. Son éstas las bibliotecas infantiles, bibliotecas circulantes, bibliotecas al aire libre y las bibliotecas de barrio, instituciones cuyo establecimiento tiende más que a proporcionar material cultural e informativo, a satisfacer una necesidad vital.

¿Cómo se forma una biblioteca?

En la actualidad, la formación de una biblioteca de conocimientos generales se realiza aplicando el criterio de clasificación Dewey, antes estudiado. Basada cada sección en libros fundamentales correspondientes a cada una de las diez cifras iniciales, se procede a su ampliación adquiriendo volúmenes correspondientes a los grupos de dos cifras, suficientes para extender cien veces el fondo anterior. Realizado esto, la especialización se iniciará insistiendo en una o varias de las secciones existentes o se procederá a continuar prestando idéntica atención a todas las secciones, según se desee establecer una biblioteca general o especializada. Es de notar que toda la biblioteca cuyas fichas del sistema Dewey acusen más de tres cifras en algunas de sus secciones, puede considerarse como especializada en ese punto. Este criterio científico e imparcial de formación es el más seguro y eficaz, porque se encuentra del todo alejado de cualquier particularismo o tendencia.

Las bibliotecas en América Latina

Las colecciones de libros tienen un doble origen en el Nuevo Mundo: algunas fueron establecidas en la época colonial y otras surgieron mientras el nuevo impulso de la libertad alentaba a los constructores de los diversos países. No habían transcurrido setenta años desde el descubrimiento de Colón, y ya surgía en Lima la primera biblioteca del Nuevo Mundo. Corría el año 1551 y faltaban ocho décadas para que los colonos norteamericanos fundaran en Harvard su primera biblioteca. Algunos años antes, en 1538, se había formado en Santo Domingo una colección de menor importancia, consultada por los estudiantes de la Pontificia Universidad de Santo Tomás de Aquino, la más antigua del Nuevo Mundo. El proceso iniciado en Lima y Santo Domingo se extendió a través de todas las regiones: en México ya había, a mediados del siglo XVI, colecciones especializadas en teología, artes, leyes, gramática y medicina. En ellas estudiaron Juan Ruiz de Alarcón, sor Juana Inés de la Cruz y otros preclaros genios. En el Cuzco, en Guadalajara, en la ciudad peruana de Huamanga, en Santiago de Chile, Caracas, Quito, Guatemala y otros centros de la cultura colonial, abundaban en el siglo XVII las bibliotecas de primera calidad. Aunque las nuevas ideas europeas penetraban en ellas con lentitud, éstas instituciones no tardaron en convertirse en focos de agitación revolucionaria. La Biblioteca de la Universidad de Caracas, fundada en 1624, fue un ejemplo típico; en sus salones estudiaron catorce de los veintinueve firmantes del acta de la independencia argentina. Después de la emancipación continuó este proceso, aunque sujeto a los naturales altibajos de la evolución política. El ejemplo más notable fue proporcionado por la Biblioteca Pública de Buenos Aires. Fundada poco después de la revolución libertadora gracias al genio profético de Mario Moreno, fue equipada al mismo tiempo que las expediciones militares al Paraguay y al Alto Perú. Convencido de que en la cultura popular residía el cimiento más firme del orden revolucionario, Moreno no se detuvo ante ningún obstáculo para completar su dotación de libros. Fue ésta la primera biblioteca americana que existió con independencia de los colegios y universidades, y que estuvo destinada a todos los sectores del pueblo. Creada el 07 de septiembre de 1810 con los volúmenes legados por el obispo Azamor, las colecciones de Moreno y Belgrano y las donaciones de muchos patriotas, ésta biblioteca ha llegado a ser una de las más importantes de la América Latina.

Las bibliotecas surgidas durante el siglo XIX han adquirido un nuevo impulso gracias al intercambio con sus similares de otros continentes, y de modo especial a la ayuda técnica prestada por la Fundación Hispánica de la Biblioteca del Congreso de los Estados Unidos. Cuando en 1943 un pavoroso incendio destruyó las valiosas colecciones de la biblioteca de Lima, las instituciones de todos los países hermanos donaron grandes cantidades de volúmenes para reparar la pérdida sufrida por la cultura hispanoamericana. En 1947 quedó constituida la Asamblea de Bibliotecarios de las Américas, cuya secretaría funciona en la ciudad de Washington.

1.1.2. Rol de la Biblioteca Tradicional

El rol tradicional de las bibliotecas ha sido concentrar, almacenar y compartir el conocimiento, la historia y la cultura. Ellas ofrecen el acceso al conocimiento e información que representan los más diversos puntos de vista y las más diversas fuentes. Las bibliotecas son junto con la educación la base para generar la innovación del pensamiento, un estímulo para la cultura, y una ayuda para el desarrollo del individuo. También son los repositorios de la memoria intelectual, cultural e histórica de la comunidad. Las bibliotecas acopian, catalogan, ponen a disposición, y preservan colecciones en todos los formatos.

El compromiso de la biblioteca es el de encontrar y resolver las necesidades de sus usuarios y mantener una posición favorable para sus proveedores de información.

El rol de las bibliotecas es el defender y ayudar a la disponibilidad de la información de manera igualitaria para el público. Las bibliotecas continuarán coordinando y facilitando la preservación de los registros y expresiones de la vida cultural en formato tradicional.

Las bibliotecas tradicionales sirven como lugar de encuentro para la comunicación y colaboración. Por ejemplo, investigadores o estudiantes se conocen en las bibliotecas para trabajar conjuntamente o para intercambiar información. Además, el uso de las bibliotecas tradicionales es frecuentemente interactivo y colaborativo, ya que el papel de los bibliotecarios como patrones de la localización de información así lo permiten. Actualmente la biblioteca de DGSCA cumple con su papel como biblioteca tradicional, promoviendo la difusión del conocimiento y presta todos los servicios que cualquier otra biblioteca.

El papel de los bibliotecarios en la sociedad es maximizar el empleo de los registros gráficos para el beneficio de la propia sociedad. En otras palabras, su función es fungir como mediador entre el hombre y los registros gráficos; no sólo libros sino también sonidos, fotografías, carteles, y cualquier cosa que contribuya con el avance del conocimiento humano. El objetivo de la biblioteca es unir al ser humano y el conocimiento registrado en una relación fructífera tan profunda como sea posible.

1.1.3. Formas de comunicación en las bibliotecas tradicionales

En la comunicación siempre están involucrados tres elementos: una fuente, uno o varios destinatarios, y un medio. La fuente proporciona la información que es transmitida por el medio (es decir, texto o sonido) que finalmente es recibido por los destinatarios.

Dependiendo de las condiciones y situaciones, la comunicación puede tomar diversas formas, y diversos propósitos que se persiguen con ella. A continuación se emplean dos dimensiones de "tiempo" (mismo, diferente) y "lugar" (mismo, diferente) para caracterizar la comunicación:

- “ La comunicación es *síncrona y no distribuida*, si la fuente y el destinatario se están comunicando mutuamente al mismo tiempo y en el mismo lugar, es decir, *hablan*.
- La comunicación es *síncrona y distribuida*, si la fuente y el destinatario se están comunicando al mismo tiempo pero en diferentes lugares, es decir, realizan una llamada telefónica.
- La comunicación es *asíncrona y no distribuida*, si la fuente y el destinatario se están comunicando en el mismo lugar en diferentes tiempos, es decir, un periódico mural.
- La comunicación es *asíncrona y distribuida*, si la fuente y el destinatario se están comunicando en diferentes tiempos y diferentes lugares, por ejemplo a través de correo electrónico.
- Adicionalmente, también se toma en consideración cómo se inicia la comunicación sincrónica (*formal o informalmente*). En la comunicación formal siempre se tiene como

base en una cita concertada, mientras que la comunicación informal es espontánea y toma lugar de forma accidental. Por ejemplo, una videoconferencia en tiempo real es una comunicación formal sincrónica y distribuida.”²

Las bibliotecas tradicionales son lugares de encuentro en donde los estudiantes e investigadores se conocen unos a otros ya sea accidentalmente o por cita concertada. Generalmente las bibliotecas tradicionales cuentan con áreas de colaboración en donde la gente puede trabajar conjuntamente. Y más aún, la mayoría de los servicios que la biblioteca presta son altamente interactivos y están basados en la comunicación entre las personas que involucra.

Debido a esto, todas las formas de comunicación existentes en las bibliotecas tradicionales son:

- La comunicación es formal, sincrónica y no distribuida cuando los bibliotecarios presentan los patrones de uso de la biblioteca.
- La comunicación formal, sincrónica y distribuida es requerida si la colección documental o la biblioteca está distribuida en diversos edificios. En este caso, un sistema de conferencia telefónica puede ser considerado como este tipo de comunicación.
- La comunicación informal, sincrónica y no distribuida es encontrada en las bibliotecas tradicionales debido a que los estudiantes e investigadores frecuentemente se conocen unos a otros accidentalmente en las bibliotecas. Está es probablemente la forma más común de comunicación en las bibliotecas tradicionales.

² Referencia tomada del Centro para el estudio para Biblioteca Digitales de la Universidad de Texas A&M

1.2. IMPACTO DE LA TECNOLOGÍA EN LAS FUNCIONES Y ACTIVIDADES DE LAS BIBLIOTECAS

Las bibliotecas acopian, catalogan, ponen a disposición, y preservan colecciones de información en todos los formatos. Estas colecciones consisten tradicionalmente de materiales almacenados en lugares específicos que limitan el acceso a solo aquellas personas que puedan trasladarse físicamente a la biblioteca o que puedan recibir el ejemplar a través de préstamo interbibliotecario. Cuando un ejemplar está en uso, no está disponible para la demás gente.

Sin embargo las bibliotecas tienen la oportunidad de cambiar este concepto con ayuda de la tecnología. Con este apoyo, existen tres tareas que las bibliotecas tienen que cubrir en los próximos años y que son:

1. La actualización de los servicios existentes haciendo uso de la tecnología.
2. Planear y desarrollar nuevos servicios y productos que empleen la tecnología.
3. Estudiar el comportamiento, perfil y necesidades de información de los usuarios de las bibliotecas.

La llave para la evolución de las bibliotecas en la era de la información, no es introducir tecnología en forma indiscriminada a la biblioteca y su funcionamiento; es necesario considerar las metas que cualquier biblioteca persigue, de la misma forma no se debe perder de vista que uno de los principales objetivos de la biblioteca es ver las necesidades de los usuarios, además de evaluar lo más adecuado que nos brinda la tecnología para alcanzar estas metas.

La información en el futuro será producida, transmitida y recuperada en forma electrónica. Los libros impresos serán remplazados por nuevas formas electrónicas; hay estadísticas que dicen que las bibliotecas basadas en papel con sus esquemas de indización darán pauta a las bibliotecas digitales dinámicas con mecanismos flexibles y eficientes para la localización, organización y personalización de gran cantidad de información multimedia.

En la década pasada proliferó un gran número de fuentes de información digital. Los avances en los sistemas de computación y redes, además de la revolución de las comunicaciones, han creado la habilidad para generar, procesar y diseminar información digital. De la misma forma, estos avances han generado nuevos mecanismos factibles en la forma de almacenar, liberar, representar, presentar y recuperar la información; Esquemas como interacción computadora-humano, sistemas de hiperbase e hypermedia, de multimedia distribuida y redes de banda amplia están considerados dentro de los nuevos proyectos que se pueden aplicar en una biblioteca.

La respuesta a la pregunta de cómo tomar ventaja de este tipo de tecnología (llamada "tecnología de información") requiere resultados empíricos y teóricos al poner en práctica un conjunto de prototipos experimentales que se aplicarían a los problemas reales de las bibliotecas.

1.2.1. Redes de computadoras

Las ventajas de las redes de computadoras incluyen el encontrar, intercambiar y compartir información desde y hacia cualquier parte del mundo. La Internet (la red de redes de computadoras más grande del mundo) ofrece la ventaja potencial de que si se publica algún tipo de información, está disponible para todos los usuarios conectados a la red. Este tipo de ventajas es actualmente tomado en cuenta por las bibliotecas para el desarrollo de nuevos servicios ofrecidos a sus usuarios.

1.2.2. Impacto de las telecomunicaciones

Desde los 80's los sistemas de recuperación de información basados en computadoras han sido muy populares en las bibliotecas y usados por los proveedores de servicios de información. En muchas bibliotecas los catálogos basados en computadoras han reemplazado sus carpetas de tarjetas tradicionales, ofreciendo ventajas tales como, acceso simultáneo para múltiples usuarios, búsquedas de información por palabra y acceso remoto. Con la aparición del CD-ROM, mucha información original puede ser accedida a través de las redes de computadoras.

Dentro de las tecnologías de telecomunicaciones existen aplicaciones que son empleadas en las bibliotecas electrónicas.

Transferencia de archivos

Tomemos como premisa que tenemos un sistema e información de una biblioteca y que con el acceso remoto se puede ver a través de una terminal virtual en nuestra computadora personal información de la biblioteca en forma digital, sin embargo, sólo se muestra en pantalla y logramos obtener pequeñas notas. Si se desea adquirir documentos más extensos, se puede ofrecer la transferencia de archivos a través de una red de computadoras. Tradicionalmente se puede adquirir esta información (libros, diarios, artículos) acudiendo personalmente a la biblioteca o con un sistema postal, mientras que la transferencia de archivos electrónica permite ventajas tales como el obtener documentos que pueden no existir en forma impresa u otra forma tradicional con una velocidad en la entrega muy superior.

Búsquedas en bases de datos

La búsqueda en bases de datos, en su forma más general, es una aplicación basada en una red de computadoras. Las bases de datos sobre las que se aplica la búsqueda fueron organizadas históricamente como una única base en una única máquina. Esto está gradualmente cambiando con el surgimiento de las bases de datos distribuidas, en donde lógicamente se ve como una única base pero está físicamente distribuida en varias máquinas. Muchas de las bases de datos basadas en CD-ROM siguen esta característica, de tal forma que son una colección de bases independientes a las que se les realiza un acceso en forma separada.

En años anteriores las búsquedas en bases de datos se han efectuado conectándose a la máquina que almacena la base a través de un protocolo de acceso remoto (tal como Telnet). En años recientes las búsquedas basadas en los estándares Z39.50 y Z39.59 de la Organización Nacional de Estándares de Información y la Institución Nacional de América de Estándares han comenzado a eliminar la necesidad de que los usuarios se conecten a las máquinas y deban, de esta forma, tener claves de acceso a las mismas. Estos estándares permiten ofrecer los resultados de las búsquedas a un usuario final empleando un protocolo de acceso remoto estándar. Esta forma de búsqueda en bases de datos es más eficiente y flexible para la red de computadoras y para la máquina donde reside la base, ya que el equipo sólo atiende las peticiones de búsqueda y no ofrece sus recursos a otros usuarios que se conectan directamente a ella por medio del telnet.

En el World Wide Web, los motores de búsqueda (tales como Lycos y Yahoo) han surgido para facilitar la localización de información en ambientes descentralizados. Estos sistemas crean un índice de páginas de WWW que se presentan como resultado de una búsqueda.

Esta forma es diferente de los sistemas tradicionales de búsqueda, ya que estos motores deben compilar y actualizar la información constantemente y el World Wide Web es tal vez el ejemplo más representativo de bases de datos distribuidas y descentralizadas.

1.3. BIBLIOTECA DIGITAL

1.3.1. Introducción

La potencialidad de las aplicaciones de la Tecnología de Información es una de las áreas más importantes y que han sido puestas a consideración por el área de Bibliotecología. La Tecnología de Información hace posible que muchos tipos de información sean almacenados, accedidos y transmitidos electrónicamente y sin hacer referencia a la impresión tradicional en papel. Como una ilustración de estas posibilidades el concepto más adecuado es el de "Biblioteca Electrónica" o también llamada "Biblioteca Digital".

En una "Biblioteca Digital"³, toda la información puede ser almacenada electrónicamente. El usuario que trabaja en una computadora en su trabajo u hogar podrá tener acceso a los datos, sin importar su localización física, ya que ésta es irrelevante para él, como también lo es el lugar donde se encuentran los datos. Los bibliotecarios y usuarios esencialmente estarán en posibilidades de acceder la información, sin importar su forma y su localización.

La posibilidad de las bibliotecas electrónicas de almacenar y compartir conocimiento, historia y cultura será primordial en los futuros desarrollos y acuerdos a los que se estén llegando. Una biblioteca electrónica es realmente una biblioteca con colecciones electrónicas presentes en una gran variedad de formatos y almacenados en diversos lugares. Progresivamente los materiales están siendo adquiridos en forma electrónica; las bibliotecas están convirtiendo sus colecciones de papel a formatos electrónicos tanto por razones de preservación como por razones de espacio.

El término de biblioteca electrónica es empleado como un agregado, lo que implica acceso electrónico a diversos recursos de información digital. Esto incluye a las bibliotecas pero no excluye a corporaciones, gobierno y entidades de investigación.

Como lo hacen actualmente, el rol de las bibliotecas en el futuro será el defender y ayudar a la disponibilidad de la información de manera igualitaria para el público. Las bibliotecas continuarán coordinando y facilitando la preservación de los registros y expresiones de la vida cultural en formato tradicional y digital. Las bibliotecas electrónicas serán fuentes de información digital gratuita o a un costo muy bajo, proveerán acceso al flujo de información electrónica de cualquier tipo y recursos digitalizados alrededor del mundo; atenderá solicitudes de distribución de documentos almacenados remotamente, realizará reproducciones digitalizadas de material único que es de dominio público, respetando los derechos de autor.

La biblioteca digital no está limitada a los libros y los periódicos. Los servicios ahora ofrecidos por empresas e industrias de información comercial y de entretenimiento son indicativos de la información con la cual las bibliotecas trabajan y las necesidades de cambio además de los deseos y necesidades de sus usuarios. La biblioteca del futuro se extenderá tanto, que, proveerá acceso inmediato a una amplia variedad de fuentes de información impresa y electrónica en el momento que los usuarios lo necesiten. Muchas bibliotecas ya están haciendo uso de la tecnología para ofrecer más facilidades de uso, incrementar su productividad y proveer más y mejores servicios a los usuarios. Ellos ofrecen por ejemplo:

³ Referencia tomada del Centro para el estudio para Bibliotecas Digitales de la Universidad de Texas A&M

- Acceso al catálogo en línea en una computadora en red o vía módem.
- Accesos a índices de periódicos, bases de datos de texto completo y otro conjunto de datos a través de catálogos o de sistemas de CD-ROM y DVD-ROM.
- Coordinación de otros registros guardados en bibliotecas y bibliografía.
- MIS (Sistemas Administrativos de Información), para la planeación y administración de las bibliotecas.

Con el surgimiento de las bibliotecas electrónicas, se incrementan las necesidades de información de sus usuarios; los recursos electrónicos vendrán a suplir los recursos tradicionales de impresión, y la biblioteca será un centro de acceso a nuevos servicios y productos. Algunas áreas de estudio y desarrollo para el futuro incluyen:

- Desarrollo de tecnologías para la administración y comunicación (Correo electrónico interno, acceso a correo de Internet, establecimiento de intranets, etc.).
- Desarrollo de tecnologías para la recuperación de información y para las respuestas a preguntas, auxiliándose de sistemas expertos, acceso a direcciones electrónicas de especialistas, búsqueda de fuentes para proveer información comercial, sistemas para ayudar a los usuarios a la elección del mejor recurso de información para sus necesidades.
- Investigación del impacto de la tecnología en los usuarios, analizando también la competencia entre los servicios de información comercial y la industria del entretenimiento para algunos servicios de información, tales como acceso al público a grandes bases de datos comerciales o las compras electrónicas. También implica instruir a los usuarios en el uso de las nuevas tecnologías, así como el aceptar que se puede presentar el rechazo de cierta tecnología pero abriéndose a los cambios que ofrezcan resolver las necesidades de ellos.

La actualización de los servicios existentes de las bibliotecas y la implementación de otros trae consigo ciertas dificultades. Se debe asegurar que la tecnología realmente resuelve las necesidades de los usuarios manteniendo un costo razonable al invertir en esta solución.

La biblioteca tiene la oportunidad de ser un líder de innovación de recursos de información. Las bibliotecas electrónicas del futuro no son sólo lugares donde se pueda obtener un libro o preguntar por una referencia; es una parte vital de la sociedad de la información.

La evolución de la infraestructura para el manejo de la información está cambiando dramáticamente, así como las operaciones tradicionales entre la gente encargada de las bibliotecas, los proveedores y los usuarios. Esto ofrece nuevas características. Nuevas formas de materiales digitalizados sin publicación (versión impresa) correspondiente están surgiendo a partir de los millones de usuarios que se encuentran conectados a la red mundial (Internet). El volumen de material digital nuevo, si estuviera en papel, empujaría a las colecciones existentes. Esta situación es particularmente compleja ya que la información digitalizada puede ser manipulada, actualizada y combinada con otros materiales con gran facilidad y mostrada en muy diversas formas. Así, los datos digitalizados crean un nuevo cúmulo de conocimientos que puede ser

accedido y manipulado a través de computadoras, con una posible existencia temporal y sin ser almacenado permanentemente. Las diferentes instituciones, incluyendo las bibliotecas, pueden proporcionar acceso a dichos materiales simultáneamente.

Un concepto que introducimos es el de *biblioteca virtual* que ha sido definida con el concepto de acceso remoto a los contenidos y servicios de las bibliotecas y otras fuentes de información, combinando una colección de material (actualmente muy consultado) en forma impresa y electrónica, con una red electrónica que proporciona el medio de acceso, los servicios de bibliotecas diseminadas en el mundo, información comercial y fuentes de conocimiento. En esencia, al usuario se le proporciona el efecto de emplear una biblioteca que es una sinergia creada al brindar tecnológicamente en conjunto los recursos de muchas bibliotecas y muchos servicios de información.

1.3.2. La infraestructura tecnológica y política

La infraestructura es convergente, y es necesario comprender la convergencia y potencialidad. Se han desarrollado extensiones por décadas para la biblioteca local, pensando en la cooperación creativa de una gran variedad de herramientas, tecnologías, consensos, políticas y técnicas de soporte. Actualmente se tiene un número crítico de estructuras para dar soporte a accesos en tiempo real a colecciones bibliográficas más allá de los muros de la biblioteca local:

- Redes de comunicaciones nacionales e internacionales con la velocidad y ancho de banda esenciales para transmitir los más grandes y complejos archivos de texto completo e imágenes digitalizadas.
- Estándares y protocolos que facilitan la comunicación computadora a computadora y base de datos a otra base de datos.
- Dispositivos de digitalización automática como equipos digitalizadores y telefacsimiles que permiten la transmisión de contenidos por volumen y en tiempo real.
- Disponibilidad de sistemas en línea disponibles a través de una gran variedad y tamaño de redes.
- Los derechos de autor y experiencia en arrendamiento serán esenciales al trabajar con una estructura legal y financiera como parte del soporte de la biblioteca.
- Una propuesta de acceso abierto y libre de pago a colecciones bibliográficas y programas son puntos importantes de tomar en cuenta al momento de crear y emplear enfoques alternativos en respuesta a los nuevos patrones de publicación y uso.

1.3.3. Concepto de "lo de afuera hacia adentro"⁴

La mayor parte de las discusiones de biblioteca digital se centra en la comprensión de la tecnología. Por lo tanto veremos la manera en que los estudiantes, maestros, investigadores y ciudadanos en general seremos afectados por tal tecnología y cómo tratamos de convivir juntos. Este enfoque es conocido como "lo de afuera hacia adentro". El corazón de la biblioteca, el compromiso de la biblioteca ha sido algo diferente a pesar de todo.

El propósito de las bibliotecas es "generar entre los seres humanos y el registro del conocimiento una relación fructífera lo más humanamente posible". Existen profundas complejidades intelectuales involucradas con el hecho de que estamos en el inicio de entender y apreciar los problemas en cuanto a información se refiere que puedan tener los usuarios y a sus necesidades individuales. la realización que da relevancia a un documento en particular es el juicio de un sólo individuo, no una constante universal.

Se cree que la biblioteca digital llegará a ser una realidad no sólo por los fundamentos tecnológicos, sino más importante aún por el revestimiento que toma el hecho de relacionar al ser humano y el registro del conocimiento. Se desea plantear un modelo de entendimiento. Además se piensa que se ganara una perspectiva robusta para el análisis posterior si se establece una perspectiva completamente contextual.

1.3.4. La Biblioteca Digital como un ejemplo de la Revolución de Control⁵

Se puede empezar por establecer una perspectiva desde un punto de vista de la sociedad de información y la biblioteca digital. La proposición establecida es que la biblioteca digital es una metáfora de la revolución de control social. El origen de dicha revolución se remonta a la mayor crisis económica y financiera en el siglo pasado. En los E.U.A., por ejemplo, aplicaciones de fuerza de vapor en los finales de 1800 presentaron un crecimiento dramático en velocidad, volumen y complejidad de procesos industriales, haciéndolos muy difícil de controlar. Los resultados del problema no se hicieron esperar: colisiones fatales de trenes, retraso de los carros frigoríficos por meses, pérdida de embarques, incapacidad para mantener altas tasas de falta de inventario. Inevitablemente la Revolución Industrial, con su enorme uso de energía para enviar los materiales a proceso, requirió de un crecimiento correspondiente en la explotación de información: La Revolución del Control.

Entre 1840 y 1920 se dieron los más importantes avances tecnológicos en comunicación y procesos de información que aún se emplean hoy en día: telegrafía, burocracia moderna, impresión por rotativos, estampilla postal, papel moneda, máquina de escribir, teléfono, proceso de tarjetas perforadas, imágenes en movimiento, radio y televisión. Se puede mostrar claramente que los desarrollos más recientes en microprocesadores, computadoras y telecomunicaciones sólo son una pequeña continuación de la Revolución del Control. Así se puede sostener:

La Sociedad de la Información no ha sido resultado de cambios recientes, pero sí del incremento en la velocidad de procesamiento de los materiales y del flujo de la economía que empezó hace más

⁴ Referencia tomada del Centro para el estudio para Bibliotecas Digitales de la Universidad de Texas A&M

⁵ Referencia tomada del Centro para el estudio para Bibliotecas Digitales de la Universidad de Texas A&M

de un siglo. De forma similar, el microprocesador y la tecnología computacional, contrariamente a las opiniones de moda, no representan una nueva fuerza sólo desencadenada recientemente en una sociedad no preparada, además de que son meramente los desarrollos más recientemente instalados en la continua evolución de la Revolución del Control. Esto explica por qué tantos componentes de control por computadora han sido anticipados, tanto por visionarios como Charles Babbage o por innovadores prácticos como Daniel McCallum, desde los primeros signos de una crisis de control en los finales del siglo XIX.

Se puede comenzar a concluir que hay fuertes implicaciones de esta evolución para el entendimiento cabal de las bibliotecas, tecnología y la evolución de la Revolución del Control apuntando hacia la biblioteca digital. Podemos expandir nuestra perspectiva de observar el control de la información y sus interacciones cíclicas con la innovación tecnológica como un mecanismo de análisis social. Esto es, existe un contexto social poderoso considerando la manera en que se crea y se usa la biblioteca digital.

La importancia de la Revolución del Control para el entendimiento de la sociedad contemporánea, especialmente el impacto continuo de computadoras y microprocesadores, radica en la lección más útil relativa a nuestro entendimiento de la vida social de forma más general. La cumbre de la Sociedad de la Información por sí misma, mas aún que el desarrollo paralelo de la teoría formal de la información, ha expuesto la centralización del procesamiento de información, comunicación, y control de todos los aspectos de la sociedad humana y el ambiente social. Son con estos conceptos fundamentales de información, que los científicos sociales pueden esperar reducir el proliferado pero aún ampliamente no sistematizado conocimiento de la estructura social y proceso.

1.4. TENDENCIAS MUNDIALES DE LOS SERVICIOS QUE OFRECEN LAS BIBLIOTECAS EMPLEANDO TECNOLOGÍAS COMPUTACIONALES

En la década pasada el número y tipo de fuentes de información digital proliferó ampliamente. Los avances en los Sistemas de Computación y la constante revolución en las redes y las comunicaciones han provocado una expansión en la habilidad para generar, procesar y diseminar información digital. Además con estos desarrollos se han creado mecanismos de emancipación y almacenamiento de información que resultan más factibles.

Antes de que estos avances puedan ser combinados de forma efectiva dentro de bibliotecas digitales, la investigación en este aspecto debe estar enfocada en áreas tales como la representación, presentación y recuperación de información; así como en esquemas de interacción humano-computadora, sistemas hyperbase, hypermedia, de multimedia distribuida y ancho de banda para red. Para la implementación de las tecnologías computacionales en las bibliotecas digitales, se requieren resultados empíricos y prácticos de estudios bien diseñados además de prototipos.

1.4.1. Biblioteca Digital considerada como Colección Nacional de Información

Las Bibliotecas Digitales han sido identificadas como "Colección Nacional"⁶ en la IITA (Aplicaciones Tecnológicas en la Infraestructura de la Información) que forma parte del HPCC (Programa de Comunicación y Computación de Alto Rendimientos en E.U.A.). La Colecciones Nacionales son fundamentalmente aplicaciones que tienen un amplio y directo impacto en la ciudadanía. Un objetivo primario de esta iniciativa nacional es el establecer mejores conexiones entre los desarrollos tecnológicos y científicos que directamente soporta la NII (Infraestructura de Información Nacional).

En respuesta a esta Colección Nacional, muchas agencias federales de Estados Unidos (NFS, NASA y ARPA) han iniciado programas de investigación de Bibliotecas Digitales, además las bibliotecas nacionales y museos han comenzado a poner sus colecciones electrónicas disponibles a través de Internet. Además de esto, muchas universidades han implementado programas de Bibliotecas Digitales.

1.4.2. Tendencias

Las Bibliotecas Digitales se enfocarán en el futuro a proporcionar las bases para un amplio conjunto de actividades humanas, tal es el caso del trabajo corporativo soportado por computadora, educación a distancia, comercio electrónico y entretenimiento. Esta transición a un sitio de trabajo

⁶ Referencia tomada de los Reportes del Grupo de Trabajo de las Bibliotecas Digitales de la IITA (Aplicaciones Tecnológicas en la Infraestructura de la Información)

con información electrónica ya ha comenzado con gran esfuerzo. Las Bibliotecas Digitales impactarán significativamente a la calidad de la educación y por lo tanto a la calidad de vida durante la siguiente década.

Una Biblioteca Digital o Electrónica puede ser el foco para muchas aplicaciones productivas (que van desde el aspecto comercial, académico y público).

La biblioteca digital no está limitada a los libros y los periódicos. Los servicios ahora ofrecidos por la información de tipo comercial y por la industria del entretenimiento son indicativos de la información con la cual las bibliotecas trabajan y las necesidades que se tienen de cambio, además de las necesidades de sus usuarios.

Una biblioteca digital que desee el éxito se extenderá tanto que, proveerá acceso inmediato a una amplia variedad de fuentes de información impresa y electrónica en el momento en que los usuarios lo necesiten.

Muchas bibliotecas ya están haciendo uso de la tecnología para actualizar la información, incrementar su productividad y proveer un mejor servicio a los usuarios. Ellos ofrecen por ejemplo:

- Acceso al catálogo en línea via módem o en una computadora en red.
- Accesos a índices de periódicos, bases de datos de texto completo y otro conjunto de datos a través de catálogos o de sistemas de basados en CD-ROM y DVD.

Netscape - [LJ Digital]

http://www.bookwire.com/ljdigital/

AN EDUCATION RENAISSANCE IS HERE!

Celebrating Israel
50
 16 Books for a Nation's Founding

LIBRARY JOURNAL
 119111

HOME	NEWS 4/20	EDITORIAL 4/20
NEWS	<u>Public libraries called top Internet access point outside home, work, or school</u>	<u>I'm Proud To Be a Librarian: Thanks to Arthur Curley, a truly learned person...</u>
VIEWS	<u>Oklahoma ups libraries'</u>	
INFOTECH		

- Análisis de otros registros guardados en bibliotecas y bibliografía.
- Con sus siglas en inglés MIS (Sistemas Administrativos de Información), para la planeación y administración de las bibliotecas.

Con el surgimiento de las bibliotecas electrónicas, se incrementan las necesidades de información por parte de los usuarios; los recursos electrónicos vendrán a suplir los recursos tradicionales de impresión, y la biblioteca será un centro de acceso a nuevos servicios y productos. Algunas áreas de estudio y desarrollo actuales y para el futuro incluyen:

- Desarrollo de tecnologías para la administración y comunicación (Correo electrónico interno, acceso a correo de Internet, implementación de intranets, etc.).
- Desarrollo de tecnologías para la recuperación de información y para dar respuestas a las preguntas más frecuentes de los usuarios, auxiliándose de sistemas expertos, acceso a correo electrónico de especialistas, además se deben crear fuentes para proveer información comercial y sistemas para ayudar a los usuarios a la elección del mejor recurso de información para sus necesidades.
- La investigación en cuanto al impacto que tiene la tecnología en los usuarios, analizando también la competencia entre los servicios de información comercial y la industria del entretenimiento para algunos servicios de información, tales

como el acceso al público a grandes bases de datos comerciales y a las compras electrónicas. También implica instruir a los usuarios para el empleo de las nuevas tecnologías, así como el análisis para descartar el uso de cierta tecnología, pero abriéndose a los cambios que encuentren cómo satisfacer las necesidades de los usuarios.

La actualización de los servicios existentes de la biblioteca y la implementación de otros trae consigo ciertas dificultades. Se debe asegurar que la tecnología realmente resuelva las necesidades de los usuarios manteniendo un costo razonable.

1.4.3. Proyectos para las Bibliotecas Digitales

La biblioteca tiene la oportunidad de ser un líder de innovación de recursos de información. Las bibliotecas electrónicas del futuro no son sólo lugares donde se pueda obtener un libro o preguntar por una referencia; es una parte vital de la sociedad de la información. Los proyectos que se tienen contemplados para la siguiente década son:

- 1) La Hiper-Biblioteca. Navegar a través de la biblioteca. Desde ofrecer sus acervos bibliográficos hasta poder tener el acceso en línea a algunos servicios.
- 2) Sistemas expertos para referencias de bibliotecas. Sistemas que puedes almacenar el conocimiento de los expertos en el área y solucionar problemas mediante la deducción lógica.
- 3) La realidad virtual para el aprendizaje y el entretenimiento.
- 4) La Literatura en Multimedia.
- 5) La literatura cooperativa. El empleo de varios medios.
- 6) Las bibliotecas electrónicas las 24 horas.
- 7) Amplia gama de recursos para acceder información.
- 8) La biblioteca como un Depósito de Información Comunitaria.
- 9) La biblioteca como un lugar de encuentro.
- 10) La biblioteca del mundo.

Hoy en día la Joint Academic Network (JANET) conjunta referencias a 150 lugares en Reino Unido y provee conexiones a otras partes del mundo, ofrece correo electrónico, transferencia de archivos, uso directo y remoto de máquinas, encaminadores para redes internacionales, acceso a los servicios de la Biblioteca Nacional y alrededor de 100 catálogos de biblioteca en línea.

1.4.4. La Biblioteca Digital en América Latina. Proyectos Actuales

Bibliotecas Nacionales

Desde 1988 las bibliotecas de 18 países de la región han trabajado en conjunto en la Asociación de Bibliotecas Nacionales de Ibero América (ABINIA) entre ellos México.

El trabajo cooperativo entre los miembros de esta asociación ha dado resultados satisfactorios tales como el CD-ROM "Catálogo colectivo de fondo antiguo siglos XV-XIX", el cual contiene los registros bibliográficos de monografías impresas desde 1900, *basadas en la información proporcionada por las bibliotecas nacionales de los diferentes países.*


Algunos países comenzaron a emplear en su bibliografía nacional, tomando como consideración de cómputo, el formato CALCO, diseñado por bibliotecas brasileñas y basado en MARC II.

También se comenzaron a emplear programas comerciales disponibles internacionalmente, tales como NOTIS en Venezuela, Puerto Rico y Chile, Dynix en México, y CDS-ISIS en pequeñas bibliotecas y proyectos individuales. El sistema NOTIS consiste de diferentes tipos de registros los cuales coexisten en una serie de relaciones; algunas de estas relaciones son permanente; los diferentes tipos de registros son enlazados de forma tal que las relaciones entre ellos no pueden ser rotas. El sistema Dynix está basado en un esquema cliente servidor que es proporcionado por las bibliotecas a través del servicio de Telnet, las bases de datos son tablas de correspondencia entre *diversos campos dentro de los registros.*

U.S. National Library of Medicine (NLM) - Netscape

File Edit View Go Communicator Help

Bookmarks Location: http://www.nlm.nih.gov/



NATIONAL LIBRARY OF MEDICINE

Free MEDLINE Search NLM Site Health Info Contact NLM Site Map

Hot Topics

- ▶ [Search MEDLINE free](#) (Internet Grateful Med or PubMed)
- ▶ [Telemedicine](#)
- ▶ [Unified Medical Language System](#)
- ▶ [Visible Human Project](#)

General Information

- ▶ [Welcome](#)

News

- ▶ [NSF-NLM Digital Library-2 Grant Program West Coast Public Briefing Announced \(4/2/98\)](#)
- ▶ [Nominations invited for Frank B. Rodgers Award \(4/1/98\)](#)
- ▶ [NLM/MLA Teleconference: Colloquium to Honor Medical Librarians/Joseph Leiter Lecture, May 12 \(4/1/98\)](#)
- ▶ ["HI98": Conference on the Health Information Infrastructure, April 27-29, Sponsored by Friends of the NLM \(2/26/98\)](#)

Bibliotecas Universitarias

Este tipo de bibliotecas cuentan con un grupo de profesionales expertos en la materia, así como de una infraestructura adecuada para proporcionar mejores servicios de información a estudiantes y profesores. En muchas ocasiones las bibliotecas de las facultades y "campus" universitarios son coordinadas por una unidad central. Algunos de sus procedimientos son centralizados y usan programas de computadoras, algunas veces desarrollados individualmente. En otros casos algunas grandes universidades emplean programas disponibles internacionalmente tal como Dynix, TINLIB y Aleph. Pequeñas universidades han adaptado Micro-ISIS para su operación. Esto hace posible la recopilación de los catálogos de monografías y publicaciones periódicas en bases de datos en línea y en CD-ROM. Muchos trabajos han sido desarrollados para coordinar estas bibliotecas a través de agrupaciones, sistemas nacionales, redes, etc., tal como Red Nacional de Bibliotecas Nacionales (RENDU) o la Asociación Nacional de Universidades e Institutos de Educación Superior (ANUIES) de México.

Notable por el significado de su educación y programas de investigación, además por el número de estudiantes y por la influencia que tienen en su país y en la región, son la Universidad Nacional Autónoma de México (143 bibliotecas) y la Universidad de Sao Pablo Brasil (38 bibliotecas); aquí los sistemas de bibliotecas son muy importantes en términos de logros bibliográficos y desarrollos, además del beneficio que se aporta a los usuarios y al estudiantado de América Latina fuera de la región. Estas bibliotecas ofrecen bases de datos y catálogos en papel, CD-ROM, en línea o como

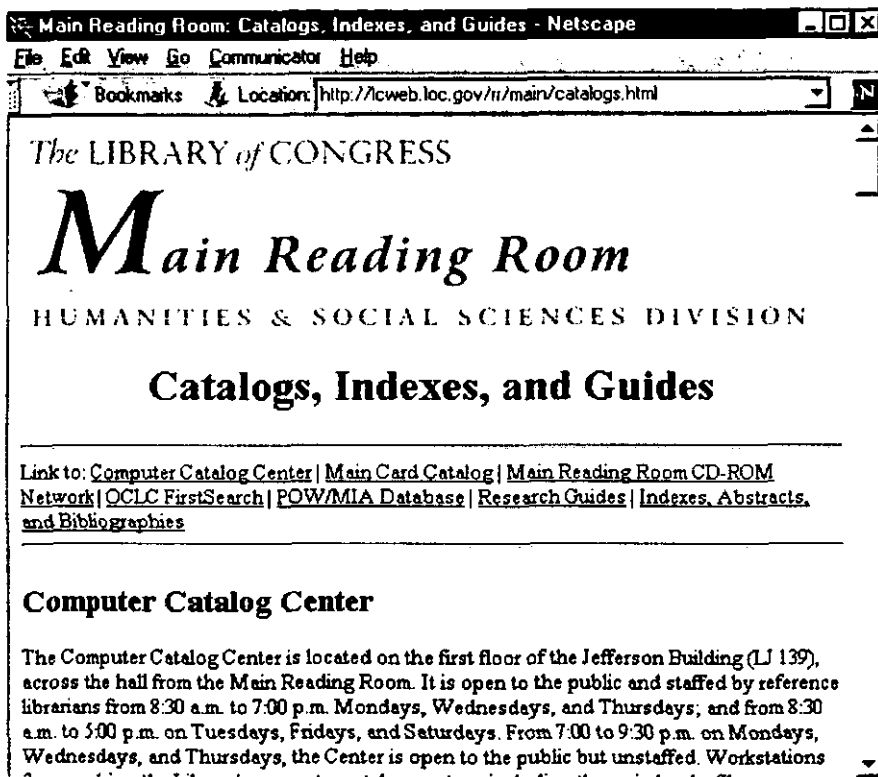
Catálogos de Acceso Público en Línea (Online Public Access Catalogues, OPACs), así como literatura representativa de la región en audiovisuales y material electrónico de buena calidad como la información en forma impresa.

Estas bibliotecas han sido pioneras en la integración y desarrollos individuales de tecnologías que brindan a los usuarios contacto con la información en su lugar de trabajo o alrededor del mundo, empleando recursos de telecomunicaciones, particularmente Internet, servicio gopher y servicio de World Wide Web. Este tipo de servicio es ofrecido por las bibliotecas de las universidades antes mencionadas además de la Universidad de Chile y las Universidades Católicas de Perú y Chile.

The image shows a Netscape browser window with the title "Búsqueda en todos los campos - Netscape". The address bar contains the URL "http://132.248.139.210:4500/ALEPH/spa/L87/L87". The main content area displays the logo of the Faculty of Engineering (Facultad de Ingeniería) of the National Autonomous University of Mexico (UNAM). Below the logo, the text reads "UNAM Búsqueda en todos los campos, usando operadores booleanos." There is a search form with a text input field and two buttons labeled "Encontrar" and "Borrar". At the bottom of the page, there are three small search icons, each with the text "Busca x" and a small graphic.

1.4.5. La Biblioteca del Congreso de los Estados Unidos de Norte América

La Biblioteca del Congreso cuenta con un acervo de información tal que, se han construido una serie de sitios donde se puede acceder a los catálogos, índices y guías de información; dentro de estos lugares se encuentran el Centro de Catálogos por computadora, la Sala Principal de Tarjetas de Catalogación y la Red Principal de Salas de Lectura de CD-ROM. Sin embargo, en estos sitios no se cuenta con un acceso remoto a través de redes tales como Internet.



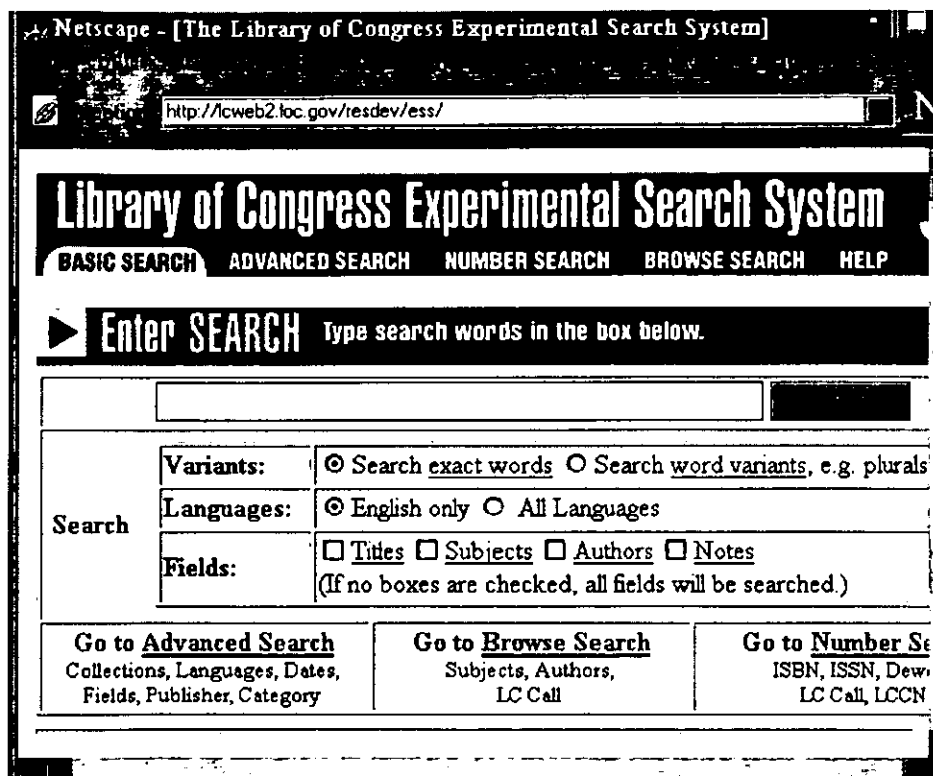
El Sistema Experimental de Búsqueda

El sistema de Experimental de Búsqueda (ESS) es uno de los primeros esfuerzos de la Biblioteca del Congreso para poner a disposición del público catálogos selectos y los recursos de biblioteca digital sobre el World Wide Web empleando una única interfaz de "apunta y selecciona", tal como un visualizador de WWW.

Esta interfaz consiste de páginas en formato HTML que contienen solicitudes de búsqueda (básica, avanzada, numérica, y una pantalla de navegación) y otras con los resultados arrojados por dichas búsquedas (una lista de elementos con información en forma resumida y un elemento de despliegue completo), junto con ayudas breves las cuales se ligan directamente con las palabras significativas y resaltadas en estas páginas. Empleando las ventajas de los vínculos de hipertexto y de un motor

de búsqueda de gran relevancia, se espera que el ESS proporcione una nueva y más intuitiva forma de buscar el tradicional Catálogo de Acceso Público en Línea (OPAC).

Al mismo tiempo, la Biblioteca posee colecciones que no son libros, tales como microfichas, publicaciones seriales, manuscritos, etc., y que aun no están representadas en sus catálogos en línea, excepto por un registro de colección. También la mayoría de los libros en las colecciones de la Biblioteca no están en computadora, a menos que algunos incluyan PREMARC. Un conjunto relativamente no editado de registros de catalogación para los libros registrados antes de 1968 están disponibles a través de LOCIS (Sistema de Información de la Biblioteca del Congreso basado en el empleo del protocolo Telnet).



Netscape - [The Library of Congress Experimental Search System]

http://lcweb2.loc.gov/tesdev/ess/

Library of Congress Experimental Search System

BASIC SEARCH **ADVANCED SEARCH** **NUMBER SEARCH** **BROWSE SEARCH** **HELP**

▶ **Enter SEARCH** Type search words in the box below.

Search

Variants: Search exact words Search word variants, e.g. plurals

Languages: English only All Languages

Fields: Titles Subjects Authors Notes
(If no boxes are checked, all fields will be searched.)

Go to <u>Advanced Search</u> Collections, Languages, Dates, Fields, Publisher, Category	Go to <u>Browse Search</u> Subjects, Authors, LC Call	Go to <u>Number Search</u> ISBN, ISSN, Dewey, LC Call, LCCN
--	--	--

Servicios del Sistema Experimental de Búsqueda

Una lista actual de las bases de datos y colecciones disponibles a través del ESS es mostrada en la página en formato HTML; muchos detalles de información pueden ser obtenidos dando el nombre de la colección en particular. Para aclarar qué colecciones contienen sólo catálogos y cuales se ligan a objetos digitales, por ejemplo, texto completo e imágenes, las colecciones se han dividido dentro de Registros de Catalogación y Multimedia.

El registro de catalogación es únicamente la referencia al elemento (libro, revista, etc.), tal como su ficha bibliográfica.

Existen registros de catalogación de cerca de 4 millones de libros; 263, 000 millones de películas cinematográficas, video film, y otros trabajos visuales; 200.000 registros de sonido y marcas digitales; más de 150.000 mapas; y 4.300 archivos de computadora (por ejemplo registros de catalogación desde 1968). ESS también contiene los catálogos de 140.000 fotografías y manuscritos en la Memoria Americana del Programa de Biblioteca Digital Nacional, vinculando a más de 70.000 fotografías digitales e imágenes disponibles en línea. Indizando el trabajo seleccionado y organizado por la página en formato HTML "Libros en Línea" de la Universidad de Carnegie Mellon, los vínculos apuntan a texto completo de 2500 libros en línea desde diferentes lugares a través de Internet.

El catálogo de multimedia de la colección de la División de Impresión y fotografía también está disponible, respetando los derechos de patente. Estos objetos digitalizados están disponibles sólo a través de un programa instalado en las computadoras que se encuentran físicamente en la Biblioteca del Congreso.

La importancia del ESS radica en la habilidad de mostrar la relación entre los términos de texto disponible, los vocabularios controlados y la catalogación.

Cada vez más documentos del gobierno y de dominio público están disponibles en texto completo, y pueden ser ligados a los registros de catálogos para proporcionar un único punto de entrada a la infraestructura de la Biblioteca Digital Nacional. Esto permitirá a los usuarios tener una experiencia enriquecida de información, vocabulario controlado, búsquedas con gran relevancia, y lograr que ellos puedan rápidamente identificar, localizar y acceder materiales que han sido convertidos a formatos digitales (que son representados como texto completo, imágenes, vídeo y audio), todo esto en múltiples instituciones, no solo en la Biblioteca del Congreso.

Mientras el ESS no reemplace los sistemas de esta biblioteca, el principio que se tiene es proporcionar un acceso inmediato a materiales de texto completo y objetos digitales con el objetivo de propiciar la migración de los sistemas existentes.

Los resultados de las búsquedas son ordenados y enviados por correo electrónico.

Netscape - [The Library of Congress Experimental Search System]

http://lcweb2.loc.gov/resdev/ess/numera.html

Library of Congress Experimental Search System

BASIC SEARCH ADVANCED SEARCH **NUMBER SEARCH** BROWSE SEARCH HELP

▶ Number Search

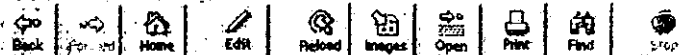
ISBN
 ISSN
 Dewey
 LC Class/Call
 LCCN

Note: Read help before searching Dewey.

- This is an experimental online public retrieval system. It is the result of a rapid prototyping effort here at The Library of Congress.
- This database is built with *Inquery*, developed by Sovereign Hill Software.

Servicios alternativos al Sistema de Búsqueda Experimental

El ESS puede ser empleado en conjunto con otros catálogos de la Biblioteca del Congreso. Estos incluyen LOCIS, el cual es un sistema que está disponible a través de Telnet; búsqueda por visualizador de WWW, una navegación y un proceso seleccionado, incluyendo referencias a través de títulos; y búsqueda por palabra, a través de una interfaz Z39.50.



Location: <http://lcweb.loc.gov/z3950/>

[What's New?](#) [What's Cool?](#) [Deliverables](#) [Net Search](#) [People](#) [Software](#)

Z39.50 GATEWAY

Welcome to the Library of Congress Page for gateway access to LC's catalog and those at many other institutions.

Contents: [Search Library of Congress Catalog](#) | [Search Other Catalogs](#) | [About the Z39.50 Gateway](#)

Search Library of Congress Catalog

- [Simple Search \(title or personal name\)](#)
- [Advanced Search \(multiple terms using Boolean operators\)](#)
 - [Tips for Using the Advanced Search](#)

Search Other Catalogs

Where applicable, the software vendor for the target Z39.50 implementation is given in parentheses after the name of the institution.

[Acadia University -- Nova Scotia, Canada \(SIRS\)](#)

[Albion College -- Albion, Mich. \(INNOPAC\)](#)

[Amarillo College \(DRA\)](#)

[Amarillo National Resource Center for Plutonium, Electronic Resource Library \(DRA\)](#)


[Anne Arundel County, Md. Public Library \(DRA\)](#)

Catálogos de la Biblioteca del Congreso

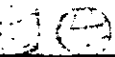
Búsqueda por palabra

Se emplea una forma dentro una página en formato HTML para buscar por palabras o números específicos en los catálogos de la biblioteca. Esta interfaz de Z39.50 proporciona una búsqueda simple para solicitudes de autores y títulos y una búsqueda avanzada que permite emplear operadores booleanos (and, or, y not), con búsquedas por nombres, títulos, series, notas, y varios números. Algunos de estos registros tienen vínculos directos a materiales digitalizados.

Location: http://lcweb.loc.gov/cgi-bin/zgate?ACTION=INFORMATION_HOST_PORT=/prod/www/data/z3950/loc.html.lib2.loc.gov



LIBRARY OF CONGRESS CATALOGS
Advanced Search



Select Database to be Searched:

Additional Information About LC Database:

Select Preferred Record Display: Brief Full Tagged

Enter Term 1:

AND OR AND NOT


Enter Term 2:

AND OR AND NOT

Enter Term 3:

Database Availability: (4:35 a.m. - 9:30 p.m. Mon-Fri) (4:35 a.m. - 5:00 p.m. Sat) (11:30 a.m. - 5:00 p.m. Sun) *All times USA Eastern*

For title and personal name searching only, use [Simple Search \(title or personal name\)](#)
For derived-key searching, use [Derived-Key Search \(author or title are known\)](#)
To search a different database, use [LC WWW/Z39.50 Gateway](#)



Library of Congress
Comments: cweb@loc.gov (03/26/98)

Búsqueda por visualizador

Se emplea una forma dentro una página en formato HTML para navegar y seleccionar el material a consultar desde "índices alfabéticos para los catálogos de la biblioteca", incluyendo referencias a través de "Tema". Se puede navegar por "Tema", autor (personal, corporativo), conferencias, títulos, series, clasificación de la Biblioteca del Congreso (número parcial), Número Decimal Dewey. También busca por números estándares exactos como el Número Internacional Estándar de Publicaciones Seriales (ISBN), Número Serial Estándar Internacional (ISSN) y el Número de Control de la Biblioteca del Congreso.

Búsqueda por comando

La forma en que se trabaja es a través de una sesión remota por "Telnet" (telnet o tn3270) y se emplean comandos que el usuario proporciona para buscar por palabras y para navegar por índices en los catálogos de la biblioteca, y para archivos adicionales no catalogados. Este método proporciona acceso a LOCIS (Sistema de Información de la Biblioteca del Congreso, el sistema de recuperación de información original basado en mainframe), con índices en los que se puede navegar, buscar palabras, tener combinación booleanas, tener opciones de despliegue, creación de conjuntos y opciones avanzadas para limitar y refinar los resultados de las búsquedas.

```
Trying 140.147.254.3...
Connected to locis.loc.gov.
Escape character is '^]'.
  L O C I S :  LIBRARY OF CONGRESS INFORMATION SYSTEM

      To make a choice: type a number, then press ENTER

1  Library of Congress Catalog           4  Braille and Audio
2  Federal Legislation                   5  Foreign Law
3  Copyright Information
*  *  *  *  *  *  *  *  *  *  *  *
7  Searching Hours and Basic Search Commands
8  Documentation and Classes
9  Library of Congress General Information
10 Library of Congress Fast Facts
11 * * Announcements * *

      The Organizations (NRCM) file is no longer created or supported by LC.
      It has been removed from LOCIS.

12 Comments and Logoff
    Choice:
1
CHOICE                                LC CATALOG                                FILE
1  BOOKS cataloged from 1898 to 1949    LOC1
    (most older records are in PREM, option 4 below)
2  BOOKS cataloged from 1950 to 1974    LOC2
```

Adquisiciones electrónicas

La Biblioteca del Congreso adquiere materiales en todos los formatos, videos, música, sonidos grabados, de todas las áreas, y de todas partes del mundo.

Catalogación

Los materiales de los catálogos en 470 idiomas y en todos los formatos, incluyendo archivos de computadora, música, sonidos grabados, hacen una colección accesible para el Congreso, investigadores y el público en general además de compartirla con las bibliotecas de todo el mundo.

Estándares

La Biblioteca del Congreso es el lugar de referencia de los estándares clave empleados en la comunidad de información, incluyendo los formatos MARC, protocolo de recuperación de información Z39.50, la Definición de Tipo de Documentos EAD (Descripción de archivos codificados) para el SGML (lenguaje de etiquetas generalizado estándar), y el Número Serial Estándar Internacional (ISSN).

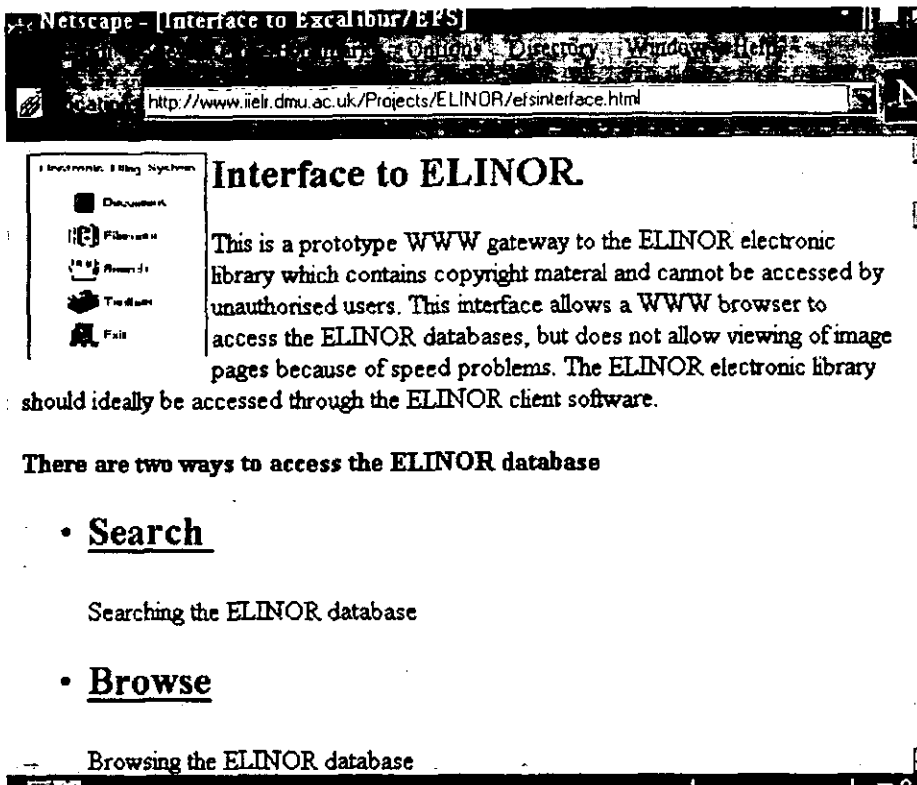
Catálogos de otras bibliotecas

Empleando el protocolo de recuperación de información Z39.50, la biblioteca ahora proporciona un acceso a búsquedas a través de formas dentro páginas en formato HTML sobre 220 catálogos en línea, de bibliotecas, dentro y fuera de los Estados Unidos de Norte América.

1.4.6. Instituto Internacional para la Investigación de la Biblioteca Digital

Elinor fue el primer proyecto de la Universidad De Montfort en el área de bibliotecas electrónicas, y fue el primer proyecto en cualquiera de las universidad Inglesas para construir un sistema de biblioteca electrónica que trabaja con texto completo para ser usado por los estudiantes. El proyecto comenzó en 1992 y fue fundado por la Universidad De Montfort, El Departamento de Desarrollo y Investigación en la Biblioteca Inglesa y por el Centro Científico de IBM de Reino Unido. Este proyecto apunta al desarrollo de la biblioteca electrónica en un ambiente de aprendizaje y enseñanza, en el cual una gran colección electrónica de información y materiales de enseñanza en texto e imágenes son accesibles para estudiantes y académicos a través de estaciones de trabajo.

A principio de 1996, el sistema contenía 120 libros de texto completos, y muchos otros materiales, tal como análisis de publicaciones y periódicos. Los resultados de este proyecto implicaron muchas horas de negociación con los editores sobre los derechos de patente y el proyecto en conjunto trajo consigo gran experiencia en el área. Para simplificar el proceso de negociación, un *modelo de aceptación de licencias* fue desarrollado. El grupo que generó este modelo recientemente ha recibido una propuesta para efectuar futuros desarrollos de software de administración de derechos de patente.



El proyecto técnicamente comenzó por las tecnologías de procesamiento de imágenes y documentos para convertir materiales impresos a imágenes (TIFF del Grupo IV) para navegar y leer páginas en pantalla. El sistema seleccionado fue el programa de administración de documentos EFS de Excalibur. Para crear textos en ASCII en los que se puedan buscar, se empleó la conversión OCR (reconocimiento óptico de caracteres). El programa EFS incluye un motor de recuperación de texto para buscar en el contenido de documentos siguiendo ciertos patrones (lo cual evita problemas con errores del OCR o algunas formas de deletrear por los usuarios).

Un pequeño proyecto de ELINOR es ELVIS, fundado por el Consejo de Educación Superior de Inglaterra, el cual desarrolló un prototipo de una interfaz de usuario para el programa ELINOR para ser usado por estudiantes con problemas de visión.

1.4.7. Proyectos del Programa de Bibliotecas Electrónicas de Inglaterra

El programa de Biblioteca Electrónica eLib⁷, es fundado por los Cuerpos Fundados de Educación Superior de Inglaterra; alrededor de 60 proyectos han surgido comenzando en 1995. El programa no es abierto y su principal enfoque es proveer recursos electrónicos y servicios para la educación superior de Inglaterra e impulsar un cambio cultural que propicie la aceptación del empleo de estos

⁷ La eLib es el Programa de Bibliotecas Electrónicas de Inglaterra.

recursos y servicios en lugar del uso de los tradicionales métodos de acceso y acopios de información.

Biblioteca Híbrida

El reto de la eLib es tener juntas las tecnologías de los nuevos desarrollos de los proyectos de bibliotecas digitales, más los productos electrónicos y servicios ya disponibles en las bibliotecas, incluyendo las funciones históricas de las bibliotecas físicas y locales, dentro de una bien organizada y accesible biblioteca híbrida.

Se están presentando proyectos pilotos, integrando recursos de biblioteca nueva y tradicional. Mientras las bibliotecas físicas son limitadas por el espacio, localizadas dentro de edificios actuales y usualmente basadas en una organización (posiblemente distribuida), la biblioteca electrónica está potencialmente libre de estas circunstancias. Así las bibliotecas electrónicas o digitales pueden ser organizadas en diferentes formas, incluyendo bases institucionales, regionales o basadas en áreas. Se busca que las organizaciones de esta naturaleza puedan tener su biblioteca híbrida, integrando acceso remoto a colecciones físicas y electrónicas y a servicios.

Biblioteca Híbrida, Proyecto Cabecera

El objetivo de este proyecto es integrar recursos relacionados a las Ciencias Sociales y de Negocios empleando las fuentes comerciales y académicas, vía una interfaz común de usuario basada en WWW. Este proyecto planea tener una combinación de elementos de investigación, aprendizaje y enseñanza de la educación superior de Inglaterra. El proyecto es llevado por la Escuela de Economía de Londres con socios en la Escuela de Negocios de Londres y la Universidad de Herfordshire.

Los recursos son objetos del mundo real con los cuales la biblioteca trabaja: los catálogos, bases de datos, aplicaciones, servicios electrónicos, recursos de WWW, etc., los cuales construyen la biblioteca híbrida. El modelo está pensado para ser aplicable y escalable para grupos de bibliotecas.

La información de estas áreas (negocios, financieras, económicas) está disponible electrónicamente empleando una variedad de medios de acceso y métodos de entrega de información, los cuales van desde interfaces propietarias de Windows hasta productos basados en WWW.

El modelo propuesto para la biblioteca híbrida está centrado en el usuario, y el concepto fundamental de diseño es que éste opere con un ambiente WWW definido precisamente para ese usuario. El sistema de cómputo consiste de una base de datos que contiene información de los usuarios que en este caso son estudiantes y académicos de educación superior de Inglaterra. El sistema tiene acceso (vía el proceso "login" de un equipo multiusuario) a los detalles administrativos del usuario (estado del área, cursos registrados, etc.) y usa esta información para dar un ambiente a la medida del usuario. El sistema tiene retroalimentación de las acciones de los usuarios, proporcionando el ambiente más adecuado al propio usuario. De esta forma el sistema es capaz de responder a las elecciones efectuadas, y ofrecer nuevos recursos (basados en meta datos vinculados a áreas del usuario) dependiendo de las elecciones realizadas.

Biblioteca Híbrida Proyecto Malibú

El proyecto se enfoca a las disciplinas Humanísticas y busca proveer formas innovadoras para conocer los requerimientos de información por parte de estudiantes y maestros a través de cooperativas de recursos compartidos.

El término de “biblioteca híbrida” se refiere a las necesidades de los usuarios para obtener acceso en forma efectiva a recursos digitales y no digitales en un ámbito de información común. El proyecto no busca crear nuevas herramientas o recursos, pero sí integrar éstos a los ya desarrollados por la comunidad de Inglaterra en este ámbito. Se crea un servicio operacional que permita a los usuarios tomar ventajas de las nuevas herramientas, recursos e infraestructura. El proyecto propone modelos para administrar y organizar la biblioteca híbrida, de soportar la mezcla de los servicios anteriores y nuevos, así como la forma de migrar desde las estructuras actuales a las nuevas.

Biblioteca Híbrida Proyecto Hylife

Este proyecto busca establecer, probar y evaluar el conocimiento de las prácticas de operación de la “biblioteca híbrida”, para que entonces pueda ser diseminada a toda la comunidad. El proyecto es administrado por la Universidad de Northumbria y es coordinado también por la Universidad del Centro de Lancashire.

Se diseña y desarrolla un programa de computadora genérico el cual es ajustado a las necesidades de grupos particulares de usuarios. Estos usuarios son estudiantes, investigadores de medio tiempo y tiempo completo, usuarios en ambientes distribuidos y de muchas áreas. Seis implementaciones de interfaces son desarrolladas y serán abiertas e independientes del sistema donde se ejecuten y constarán de dos etapas; inicialmente con una interfaz WWW empleando las bondades del programa Z39.50 dando acceso a una amplia gama de servicios, incluyendo los catálogos de la biblioteca.

El proyecto está enfocado a integrar los servicios de entrega de información electrónica e impresa en un acceso electrónico.

Biblioteca Híbrida Proyecto Builder

Este proyecto es dirigido por la Universidad de Birmingham y como los anteriores proyectos permitirá el acceso a fuentes de información electrónica e impresa, local y remota empleando una interfaz basada en WWW. Consta de seis modelos:

- registro e inducción de usuarios
- organización y entrega del material
- índices de meta datos para fuentes electrónicas e impresas (incluyendo catálogos de biblioteca, bases de datos en línea, paquetes CAL, recursos Internet y de CD-ROM)
- enseñanza y aprendizaje (integrando materiales existentes, el proyecto de eLib, y para el soporte a la Intranet).
- publicaciones y digitalización
- acceso a puentes de tipo Internet para recursos fuera de esta red.

Aplicaciones SGML de la biblioteca digital en Inglaterra (ELSA)

El proyecto ELSA se encamina al empleo de documentos en las bibliotecas que han sido elaborados en formato SGML (Lenguaje Estándar Generalizado de Etiquetas). El proyecto construye un acopio electrónico de documentos los cuales tomarán están en un servidor en una red. Las computadoras clientes estarán aptas para acceder el material del servidor, obtenerlo y ponerlo disponible a bibliotecas y usuarios finales para su uso y manipulación. El proyecto establecerá una especificación detallada del ambiente de trabajo del usuario final basada en un modelo de cliente servidor que permita la consulta a estos usuario y a los bibliotecarios. La especificación y análisis de estándares relevantes dirigirán la construcción de una interfaz de usuario la cual será producida para proporcionar la habilidad de buscar, recuperar y visualizar material.

Se proporcionará un acopio local de documentos para los usuarios finales en las máquinas de los mismos para la recuperación local de esta información. El sistema resultante es probado en un ambiente apropiado con retroalimentación de los bibliotecarios y los usuarios finales. El enfoque de este proyecto será la habilidad de los sistemas (integrados a este proyecto) para mejorar los ya existentes que proveen información y permitir que nuevos sean desarrollados. Un componente esencial del proyecto será la diseminación de información como indicativo de progreso del mismo. El proyecto es fundado a través de la Comisión Europea DG XIII y tiene tres socios; Jouve System D'Information (de Francia), el socio principal, el cual provee el motor de búsqueda, los programas de interfaz de usuario y los programas de cliente servidor, Ciencia Elservier el cual provee los documentos y la Universidad De Montfort que desarrolla la interfaz de usuario y realiza las pruebas.

1.4.8. Otros proyectos de Biblioteca Electrónica

Empleo de los Textos Electrónicos por los Estudiantes de Ciencias Sociales y Humanidades⁶

Es necesario explorar las formas en que los textos electrónicos están siendo utilizados por la comunidad. Se asegura que actualmente los eruditos emplean la tecnología de la información y las tendencias claves de la tecnología de la información más relevantes en el proceso de investigación, para predecir cómo los estudiosos de ciencias y humanidades están preparados para emplear textos electrónicos. Como punto introductorio a este tópico, el empleo actual de textos electrónicos es explorado ampliamente en el contexto de la comunicación humanista. Desde esta perspectiva, el trabajo de los eruditos de las ciencias sociales y humanidades involucra cinco procesos: 1.- identificación de las fuentes, 2.- comunicación con los colegas, 3.- interpretación y análisis de los datos, 4.- diseminación de los resultados de las búsquedas y 5.- desarrollo e instrucción curricular. La extensión del impacto o influencia de la computación actualmente en los aspectos de la comunicación humanista representa una variable indicativa de la prospección de los textos electrónicos.

⁶ Referencia tomada de los Grupos de trabajo de la IITA (Aplicaciones Tecnológicas en la Infraestructura de la Información)

Se están analizando las tendencias clave en el uso humanista de la tecnología de información. Esto incluye las tendencias de cómputo del usuario final y la conectividad, lo cual proporciona el marco de dispersión del uso de textos electrónicos en este milenio. Además es necesario considerar el impacto que tendrá el uso de textos electrónicos en los proveedores de información.

Uso del proyecto Memoria Americana en bibliotecas públicas y escolares

Se llevó a cabo un estudio que se centró en las aplicaciones no especializadas de materiales bibliográficos electrónicos, especialmente en el proyecto Memoria Americana de la Biblioteca del Congreso, en un reducido número de bibliotecas públicas y escolares a través de los Estados Unidos. El proyecto Memoria Americana consiste de materiales del acervo primario de la Biblioteca del Congreso almacenados en medios ópticos (CD-ROM y videodiscos), y presentados con poca o nula edición. Diversas colecciones están acompañadas de introducciones electrónicas y guías de usuarios ofreciendo información adicional y contexto histórico. Las colecciones se presentan en una gran variedad de formatos incluyendo fotografías, artes gráficas, películas, grabaciones sonoras, música, manuscritos, libros y panfletos.

La Biblioteca del Congreso comenzó una evaluación a nivel nacional de la Memoria de América en diversas instituciones. Los sitios de prueba incluyeron bibliotecas públicas, bibliotecas de escuelas primarias y secundarias, estatales, de bachillerato, universitarias y especiales. Esta evaluación dio como resultado que el uso no especializado de la Memoria de América se centra en las bibliotecas públicas como institución cultural comunitaria, frecuentemente vinculo entre la educación formal e informal. El uso por parte de estudiantes y maestros ha revelado cuestiones generales sobre el uso de recursos electrónicos; también ha permitido definir beneficios para los usuarios no especializados. Algunos temas incluyen el problema de entendimiento de contenido y contexto en un ambiente electrónico, y las habilidades e intereses surgidos a partir del uso de recursos electrónicos.

El proyecto Perseo: Fuentes interactivas y estudios en la Grecia Clásica

El Proyecto Perseo liberó su primera versión pública con base de datos de hipertexto y materiales multimedia de la Grecia Clásica. Perseo fue diseñado para ser empleado por un gran público, conformado principalmente por lectores en niveles escolares. Por lo tanto, debe ser capaz de localizar información empleando distintas estrategias, conteniendo suficiente detalle para satisfacer las distintas necesidades de los usuarios. Además, debe ser entregado a los usuarios en forma tal que resulte muy sencillo de entender.

Para lograr tales objetivos, el grupo de desarrollo decidió realizar una separación meticulosa de la base de datos textual y de imágenes, seleccionada y convertida (en un sentido), y de la colocada en el sistema de entrega (por otro). De esta manera es posible generar los datos electrónicos sin tener que pensar en las restricciones en los sistemas de entrega. Se ha hecho un gran esfuerzo para elegir formatos para los datos independientes de la plataforma, y se ha puesto trabajo e imaginación en estructurarlos para que la traducción de papel a formato electrónico tenga de valor agregado.

Mucho del trabajo en Perseo se ha enfocado a la recolección y conversión de datos en los que el proyecto está basado. Al mismo tiempo, es necesario proporcionar medios de acceso a la información, de tal forma que sea útil, y luego investigar cómo es empleada. Conforme se aprende

qué hacen los estudiantes y expertos con el sistema, se puede ir haciendo el ajuste a la colección de datos, y modificar el sistema para mejorarlo. Al crear un sistema de entrega para uso general, se ha tratado de evitar favorecer un tipo de uso, permitiendo diversas maneras de acceso y navegación a través del sistema.

La manera en que el texto es manejado, ejemplifica algunos de estos principios. Todos los textos en Perseo presentan formato SGML, siguiendo los lineamientos de la Text Encoding Initiative (TEI). Estas etiquetas son empleadas para indizar el texto y procesarlo de tal forma que pueda ser importado en HyperCard. Ninguna etiqueta SGML permanece en el texto que le aparece al usuario, debido a que resulta muy costoso crear un sistema que trabaje con SGML en tiempo real. Sin embargo, la regularidad proporcionada por SGML es esencial para verificar el contenido de los textos, y hace más rápidos los procesos realizados sobre ellos. El hecho de que el texto exista en formato SGML asegura que será relativamente sencillo portar a distinta plataforma y ambiente operativo, restándole importancia a la plataforma actual. Finalmente, las etiquetas SGML incorporan los sistemas de referencias canónicas existentes (capítulo, estrofa, línea, etc.); la indización y navegación están basadas en estas características. Esto permite asegurar que la misma referencia canónica conducirá al mismo punto dentro del texto, y que todas las versiones de nuestros textos, independientemente de la plataforma de entrega, funcionarán de la misma manera.

Para proporcionar herramientas a los usuarios, el texto es procesado por un analizador morfológico, y el resultado es almacenado en una base de datos. Junto con los índices, el Léxico Griego - Inglés, y el índice de todas las palabras en inglés dentro de la definición del léxico, el análisis morfológico comprende un conjunto de herramientas lingüísticas que permiten al usuario de todos los niveles trabajar con información textual y llevar a cabo diversas tareas. Por ejemplo, estudiantes que no leen griego, pueden explorar un concepto como si apareciera en textos griegos empleando índices Inglés - Griego, y entonces buscar palabras en el texto y se hace la traducción, o los estudiantes pueden realizar estudios morfológicos detallados de palabras empleando análisis morfológico del texto. Debido a que estas herramientas no son diseñadas para un uso específico, pueden ser empleadas junto con los mismos datos por estudiantes y eruditos.

El papel de las etiquetas SGML en el proyecto CORE

La importancia de redes de telecomunicaciones de alta velocidad como característica básica para el lugar de trabajo, está dirigiendo la demanda de entrega de documentos electrónicos. Tres distintas categorías de publicación / republicación electrónica son necesarias para apoyar la demanda de acceso en este ambiente inicial:

- Conversión de archivos de papel o microfilmación a formato electrónico
- Conversión de archivos electrónicos a formatos apropiados para la recuperación y despliegue
- Edición electrónica primaria (materiales para los que las versiones electrónicas sean el formato primario)

OCLC (Online Computer Library Center) tiene actividades experimentales y desarrollo de productos en cada una de estas áreas. Entre los retos en que se está trabajando está la integración del almacenamiento de los tres tipos de información en sistemas distribuidos coherentes.

El Proyecto CORE (Chemistry Online Retrieval Experiment) es un modelo de conversión de colecciones de grandes textos y gráficas para los que archivos electrónicamente generados están disponibles (categoría 2). La Sociedad Química Americana ha puesto a disposición archivos tipográficos que datan de 1980 para sus 20 publicaciones. Esta colección de 250 publicaciones / año está siendo convertida a un formato electrónico que estará disponible a través de diversas aplicaciones para usuario final.

El uso de Standard Generalized Markup Language (SGML) ofrece el medio de capturar la riqueza estructural de los artículos originales en una forma que soportarán una variedad de opciones de recuperación, navegación y despliegue necesario para consultar de manera efectiva grandes bases de datos de textos

Un documento SGML consiste de un texto con etiquetas descriptivas que especifican la función de un elemento dado dentro del documento. Como una construcción formal de lenguaje, un documento SGML puede ser comparado con una definición de tipo de documento (Document-Type Definition, DTD) que determina sin ambigüedades qué elementos son permitidos y en qué parte del documento pueden o deben ocurrir. Este mapa formalizado de la estructura de un artículo permite que el diseño de la interfaz con el usuario sea totalmente independiente del sistema de bases de datos, dando un importante paso hacia la interoperabilidad. Una demostración de esta separación es parte del proyecto CORE, donde el diseño de la interfaz del usuario nació con filosofía muy distinta pero acceden a la misma base de datos.

El proyecto CORE es una colaboración entre la Biblioteca de la Universidad de Cornell, Bell Communications Research (Bellcore), la American Chemical Society (ACS), la Chemical Abstracts Service (CAS), y OCLC.

1.5. PRESERVACIÓN DIGITAL

1.5.1. La Biblioteca de la Universidad de Cornell

La Biblioteca de la Universidad de Cornell, Cornell Information Technologies y Xerox Corporation, con el apoyo de la Comisión de Preservación y Acceso, y Sun Microsystems, Inc., han colaborado en un proyecto para someter a prueba un prototipo de grabación de libros maltratados en forma de imágenes digitales, produciendo en demanda, archivos de reemplazos de alta calidad. El proyecto es más ambicioso, explorando además algunos aspectos de la digitalización, almacenamiento, recuperación, y acceso a las imágenes en un ambiente de red.

El estudio conjunto comenzó en enero de 1990. Xerox proporcionó la aplicación College Library Access and Storage System (CLASS), un prototipo de digitalizador de 600 puntos por pulgada, y el equipo necesario para proporcionar impresión en red en una impresora DocuTech ubicada en el Cornell's Computing and Communications Center (CCC).

El grupo de personas de Cornell que emplean la aplicación se constituyó para el proceso de desarrollo y pruebas para las mejoras al sistema CLASS. La naturaleza de colaboración de la relación resultó en un sistema específicamente diseñado para la preservación de acervos de este tipo.

Una biblioteca digital de 1,000 volúmenes (o aproximadamente 300,000 imágenes) fue creada y almacenada en un robot de discos opto magnéticos ubicado en el CCC. La biblioteca incluye una colección de monografías selectas de matemáticas proporcionando a los maestros en esta rama una posibilidad para emplear la biblioteca electrónica. Los volúmenes restantes fueron elegidos de la biblioteca para probar diversas capacidades del sistema de digitalización.

Un objetivo del proyecto es proporcionar a los usuarios de la biblioteca de Cornell y a sus administradores la facilidad de solicitar facsímiles de imágenes digitalizadas o recuperar la imagen actual a través de la navegación. Un prototipo de "equipo de escritorio para visualización" creado por Xerox permite experimentar con patrones de acceso a las imágenes que constituyen la biblioteca digital; proporciona búsqueda, recuperación y funciones de impresión con mejoras para facilitar la navegación a través de múltiples documentos.

Cornell comenzó a trabajar en extender el acceso a la biblioteca digital a lectores empleando estaciones de trabajo desde sus oficinas. Hubo un arduo trabajo en el desarrollo de un servidor de conversión y entrega de imágenes en ambiente de red, y aplicaciones cliente que soportaran visualizadores en plataforma Apple Macintosh, IBM y Sun. El equipo para este desarrollo fue proporcionado por Sun Microsystems con apoyo de la Comisión de Preservación y Acceso.

La versión final del "equipo de escritorio para visualización" proporciona al usuario otro mecanismo de acceso a la biblioteca digital, dando la oportunidad de observar las imágenes directamente. No se requieren de aplicaciones especiales, sin embargo es muy recomendable tener una máquina poderosa con buen manejo de gráficos.

El Estudio Conjunto en Preservación Digital ha generado un gran interés en la comunidad bibliotecaria. Afortunadamente, este proyecto ha servido para solucionar un gran número de problemas alrededor del uso de tecnología digital para la preservación y uso de materiales bibliográficos deteriorados, que proyectos subsecuentes necesitarán examinar.

Este proyecto resultó en la captura de 1,000 libros deteriorados como imágenes digitales de 600 puntos por pulgada y, la producción en demanda de reemplazos del papel de alta calidad. El proyecto, apoyado por la Comisión de Preservación y Acceso, también investigó algunos de los problemas derivados de la digitalización, almacenamiento, recuperación y acceso a las imágenes digitales en ambiente de red.

Al enfocarse únicamente en los problemas relacionados con la digitalización, tenemos identificado al Proyecto Cornell Xerox. Dentro de los problemas estudiados están: imagen contra texto capturado; indización y acceso; capacidades de captura de imágenes; comparación entre fotocopia y microfilmación; análisis de costos de producción; formatos de almacenamiento; protocolos y estándares; uso de la tecnología de digitalización para propósitos de preservación.

Las imágenes de 600 ppp producidas en el proyecto proporcionan una calidad altamente aceptable para crear reemplazos de los originales en papel ya deteriorados. Los 1,000 volúmenes digitalizados proporcionan un conjunto de retos de captura de imágenes que son comunes en las técnicas de impresión del siglo pasado y del material deteriorado, y que retan el uso de los procesos de conversión a texto. Estos retos incluyen contraste disminuido entre el texto y el fondo, páginas frágiles y deterioradas, impresión no clara, tipos de letra muy elaborados, adyacencia de texto desvanecido y oscura, anotaciones y textos escritos a mano, lenguajes no romanos y la proliferación de material ilustrado incrustado en el texto. Las últimas categorías incluyen la frecuencia de los medios tonos, fotografías de tono continuo, esquemas matemáticos complicados, mapas, dibujos en colores invertidos, grabados, relieves, etc.

El sistema prototipo de digitalización de Xerox proporciona un número importante de características para la captura de material diverso. Los técnicos emplean múltiples definiciones de inicio, filtros, arte lineal y definiciones de medio tono, segmentación automática, y programas de edición para optimizar la captura de la imagen. Al mismo tiempo, el proyecto se enfoca a la producción. El objetivo fue lograr hacer la digitalización tan sencilla y aceptable como el fotocopiado y la microfilmación para la preservación. El estudio de tiempo - costo de este proyecto reveló la viabilidad económica de la digitalización.

El Proyecto Cornell Xerox fue pensado para emplear estándares no propietarios y para el uso de protocolos comunes cuando no existiera un estándar. Los archivos digitales fueron creados como imágenes TIFF que fueron comprimidas empleando el algoritmo de compresión CCITT nivel 4. La aplicación de Xerox es en ambiente MS-DOS y emplea programas como Microsoft Windows y Wang Image Wizard. La biblioteca digital está diseñada para ser independiente de la plataforma de cómputo y proporcionar servicios de intercambio con otras instituciones a través de conexiones de red. El acceso a los archivos digitales es proporcionado en dos formas: registros bibliográficos para computadoras son creados en sistemas locales en Cornell accediendo las imágenes digitales actuales y, a través de la estructura de control y los servidores de archivos en red.

1.5.2. Proyecto de digitalización de textos de la Biblioteca Nacional de Agricultura de los Estados Unidos de Norte América

La Biblioteca Nacional de Agricultura (NAL) ha acumulado una amplia experiencia con la digitalización de materiales impresos. Desde finales de la década de los 80's, la biblioteca ha participado en el Proyecto Nacional de Digitalización de Textos de Agricultura, un esfuerzo de cooperación entre la NAL y 45 bibliotecas universitarias especializadas en ciencias de la tierra.

Uno de los aspectos más relevantes de este proyecto se refiere a la conversión de microfilmaciones a imágenes electrónicas con el fin de proporcionar acceso a ellas. Con la cooperación de la Universidad Tuskegee, la NAL seleccionó tres carretes de microfilm de una colección de 67, conteniendo papeles, cartas y dibujos de George Washington Carver. Los tres carretes fueron convertidos en 3,500 imágenes electrónicas empleando un digitalizador especial.

1.5.3. Proyecto Open Book

El Proyecto Open Book, es el esfuerzo de la Biblioteca de la Universidad de Yale para convertir 10,000 libros de microfilmación a imágenes digitales. Después de la etapa de planeación y organización, se han seleccionado a las más grandes empresas para involucrarlas como patrocinadoras del proyecto y como integradoras del proyecto. En la propuesta se establece que el patrocinador ayudará a aislar las áreas de riesgo e incertidumbre así como para resolver los problemas clave que se presenten durante el tiempo de vida del proyecto. La Biblioteca de Yale decide sobre el orden en el que el material sea convertido a imágenes digitales, y se encarga del proceso de obtener fondos para el proyecto total.

La propuesta de Yale aceptó que para la implementación del Proyecto se proporcione al final de tres fases un subsistema de conversión, estaciones de navegación distribuidas en la red del campus y en la Biblioteca de Yale, un subsistema de almacenamiento de las imágenes de los 10,000 libros a una resolución de 200 y 600 ppp, y acceso a través de la red a impresoras.

Yale seleccionó a su patrocinador en un proceso formal, parcialmente apoyado por la Comisión de Preservación y Acceso. Después de una solicitud de propuesta, la Universidad de Yale seleccionó a dos empresas como finalistas para trabajar con el personal de la Biblioteca para generar un análisis detallado de requerimientos para el Proyecto Open Book. Cada empresa empleó los resultados de análisis de requerimientos para generar y enviar una propuesta formal del proyecto completo. Este proceso competitivo no sólo dio la posibilidad a la Biblioteca de Yale a seleccionar a su patrocinador principal, sino que le reveló un panorama muy claro del estado de la industria del manejo de imágenes, las variantes, los esfuerzos corporativos y las dinámicas de las organizaciones para el desarrollo de los mercados.

El Proyecto Open Book está enfocado específicamente en la conversión de imágenes de microfilmación a formato digital. La tecnología de digitalización de microfilmaciones está disponible pero cambia rápidamente. En los requerimientos del proyecto, la Biblioteca de Yale enfatizó las características de la tecnología que afectan la calidad técnica de la producción de imágenes digitales y los costos de crear y almacenar el acervo de imágenes: ¿qué niveles de resolución digitales deben ser empleados para digitalizar las microfilmaciones?, ¿cómo las variaciones en la calidad de la microfilmación, particularmente en película producida por

estándares de preservación, afecta la calidad de las imágenes digitales?. ¿qué tecnología puede operar efectiva y económicamente cuando se digitaliza una película para separar dos imágenes y para controlar y corregir imperfecciones en la imagen?. ¿cómo el control de calidad puede ser integrado en el flujo de trabajo de la digitalización de las imágenes que incluye la indización de los documentos y su almacenamiento?.

Los usos actuales y esperados de las imágenes digitales - almacenamiento, navegación, impresión y OCR - ayudan a determinar los estándares para determinar su calidad. La navegación es especialmente importante, pero las facilidades disponibles para los lectores al navegar en documentos es quizá el aspecto más débil de la tecnología de imágenes y la más necesitada de desarrollo. Al definir sus requerimientos, la Biblioteca de Yale se concentró en aspectos fundamentales de utilidad de los documentos: ¿proporciona el sistema suficiente flexibilidad para manejar un amplio rango de tipos de documentos, incluyendo monografías, conjuntos multipartes y multivolumen, seriales, y colecciones de manuscritos?. ¿qué convenciones son necesarias para identificar de manera inequívoca un documento para su almacenamiento y recuperación?. ¿dónde colocar la base de datos de registros con información bibliográfica de los documentos?. ¿cómo se presentan físicamente los documentos en la pantalla del lector?.

La Biblioteca de Yale diseñó el Proyecto Open Book basado en el hecho de que la microfilmación es el medio más adecuado para la preservación del contenido de materiales bibliográficos deteriorados. Es claro que el reto de la tecnología de imágenes digitales y la clave del éxito de esfuerzos como éste, es proporcionar un medio para la preservación y acceso a los materiales deteriorados.

1.6. ORGANIZACIÓN DE UNA BIBLIOTECA DIGITAL GLOBAL. BIBLIOTECA DEL CONGRESO

Convenciones de Nombres

En sus inicios en 1984, Z39.50, el protocolo de recuperación de información fue contemplado como una herramienta de búsqueda y recuperación de información de computadora a computadora que permitía a una interfaz única de usuario acceder diversos sistemas. En su última versión, Z39.50 se muestra ya como un protocolo maduro, muy rico y complejo, tan complejo que a diferencia de ver al estándar como un todo que debe ser implementado en cualquier sitio, se tienen perfiles selectos del estándar que pueden implementarse conforme se requieren. Los perfiles son conjuntos de acuerdos de lo implementado especificando el uso de un estándar en particular para apoyar una aplicación específica, función, clase de información o ambiente. Los desarrolladores, clientes, y usuarios acuerdan de conformidad un subconjunto de un estándar más amplio, correspondiente a los requerimientos relevantes específicos. Los vendedores desarrollan productos que cumplan con las especificaciones para que los productos resultantes puedan operar entre ellos. GILS, el Servicio de Localización de Información Gubernamental de los Estados Unidos de Norteamérica ha adoptado un perfil de Z39.50, y algunos otros para búsquedas autor / título / tema, búsqueda de información geoespacial y de objetos de museos.

Actualmente está en desarrollo el Perfil de Acceso a Colecciones Digitales. Bibliotecas e instituciones crean colecciones, organizadas temáticamente (es decir, por creador, tema período histórico, etc.) con diversos y numerosos objetos, tanto físicos como digitales. Estas colecciones están organizadas frecuentemente en forma jerárquica y distribuidas entre diversas instituciones. A pesar de que numerosos recursos son invertidos en la digitalización, incremento, organización y descripción de la información en una colección, para un usuario (o cliente) la colección puede parecer una simple acumulación de objetos de datos sin diferencia alguna. Este problema se deriva del hecho que no hay una semántica predeterminada para revisar dicha colección, para localizar y recuperar un objeto de interés. Los objetivos primordiales del Perfil de Acceso a Colecciones Digitales son permitir a los servidores designar con claridad qué es contenido y qué es descripción; para permitir a un cliente navegar entre información descriptiva asociada (como son ayudas en la localización o registros catalográficos); para modelar las relaciones entre colecciones y objetos; para proporcionar una semántica a las solicitudes para ser interpretadas consistentemente entre distintos servidores; para proporcionar estructuras de datos que proporcionan la semántica para la navegación, y finalmente para permitir a los usuarios la localización y recuperación de los objetos de interés.

El Perfil para el Acceso a Colecciones Digitales, aunque está limitado en su alcance, explícitamente anticipa el desarrollo de diversos perfiles adicionales. Los perfiles adicionales son extensiones compatibles o subconjuntos del perfil de Colecciones Digitales, para aplicaciones específicas o clases de información en disciplinas especializadas, por ejemplo, colecciones de museos o datos geoespaciales.

Repositorio Digital de la Biblioteca del Congreso. Evolución de los Aspectos Técnicos

Los esfuerzos relativos al Repositorio Digital comenzaron a principios de los 80's como un esfuerzo para compartir las únicas colecciones visuales de la Biblioteca, preservando las imágenes empleando tecnología de videodiscos como parte del proyecto Memoria Americana, con el intento de poner a disposición los discos a otras bibliotecas e instituciones. Como el volumen de material a ser digitalizado fue enorme, los esfuerzos fueron enfocados a poner a disposición los materiales digitales en línea, fortuitamente al tiempo que el surgimiento del World Wide Web se convirtió en una realidad viable.

Como cantidades masivas de material digital fueron creadas, la Biblioteca se hizo más consiente de la naturaleza y estructura del repositorio digital; y además se creó el Grupo de Nombres del Repositorio para considerar la extensión de las configuraciones establecidas en los diversos proyectos y esfuerzos iniciales que pudieran ser empleados para escalar el ambiente a un repositorio de biblioteca. Diversos factores jugaron un papel importante en el proceso de toma de decisiones, principalmente relacionados con el ambiente, incluyendo el hecho de que las colecciones digitalizadas fueron primordialmente de las colecciones retrospectivas de materiales distintos a libros de la Biblioteca, la mayor parte de los cuales no se encuentran bajo la protección de las restricciones de los derechos de propiedad. Esto significa que la Biblioteca no ha señalado aquellos materiales que comienzan su vida en forma digital, y el gran énfasis de "colecciones" significa que no han sido señalados ejemplares sin relación con un grupo significativo.

El grupo además ha heredado algunas convenciones de los proyectos de videodiscos, especialmente aquellos relacionados a la catalogación de imágenes individuales en la colección y la conexión entre la descripción bibliográfica y la imagen relativa. Los esfuerzos de digitalización son usualmente hechos en lotes enormes por contratistas que se especializan en este tipo de trabajos, los contratos para este trabajo son preparados mucho antes de que los materiales digitales sean producidos, y las personas que hacen esta labor deben tener mecanismos para identificar las imágenes digitales y relacionarlas con los objetos fuente de donde fueron digitalizados. Sin embargo, la identificación de archivos necesita ser determinada proactivamente, no después de la entrega a la Biblioteca. Aun cuando estos archivos son eventualmente transferidos a un ambiente UNIX, las convenciones de nombres más flexibles en UNIX no están disponibles en el proceso de producción, como la entrega de los archivos a los contratistas y el proceso de control de calidad los cuales son realizados en ambiente DOS. Finalmente, los proyectos exploratorios en la Biblioteca han resultado en diversas reglas empleadas para organizar los diferentes grupos de recursos digitalizados. Consecuentemente, aplicaciones locales han sido adaptadas para cada colección para apoyar la entrega de imágenes cuando éstas son solicitadas.

Tomando estos factores en cuenta, el Grupo de Nombres de Repositorios ha adoptado el siguiente conjunto de hipótesis:

- El grupo sólo puede interesarse con activos físicamente almacenados en los repositorios de la Biblioteca
- Los activos digitalizados pueden no ser directamente buscados, pero sí accedidos vía registros catalográficos o ayudas de localización
- La Biblioteca puede continuar con las restricciones de DOS para los identificadores de archivos

- En la medida de lo posible, los identificadores de archivos pueden no contener significados inteligentes respecto a sus contenidos, tales como los mnemónicos
- No es práctico simplemente asignar números consecutivos a ejemplares digitales en el espectro total y por el volumen del repositorio
- Los activos pueden ser digitalizados y agregados al repositorio en grandes grupos, o colecciones, no individualmente
- Una aplicación de administración de almacenamiento jerárquico generalizado puede ser adoptada para localizar cualquier grupo identificado o activos específicos - esto se conoce como sistema de localización de documentos -
- Los identificadores necesitan ser asignados en forma eficiente conforme los activos van siendo digitalizados, y permanecen apropiadamente cuando son instalados en los repositorios, o reubicados dentro del repositorio

Trabajando dentro de este marco, el grupo ha producido un conjunto de lineamientos para establecer ejemplares digitales con identificadores únicos y persistentes que pueden ser proporcionados a los contratistas durante el inicio del proceso de digitalización. El resultado es un nombre de dos partes a partir del cual un URN (Uniform Resource Name) puede ser derivado. La primera parte del nombre lógico es un "agregado" - que representa el nivel más bajo de agrupación reconocido por la división de conservación de la Biblioteca al digitalizar los ejemplares -. Estas agrupaciones, o lotes, no siempre son equivalentes a las colecciones "tangibles" existentes, y en algunos casos, puede reflejar sólo el nombre de agrupación para almacenamiento. El requisito más importante para los agregados del nombre es que el nombre sea único en todo el repositorio, lo que puede permitir que el agregado y todos sus archivos puedan ser movidos de una máquina a otra sin cambiar la información en la descripción bibliográfica.

La segunda parte del nombre lógico es un "identificador de ejemplar", que representa la unidad de más bajo nivel de contenido descrito por la división de conservación de la Biblioteca, independientemente de si es un libro, una fotografía, una carpeta manuscrita o una grabación sonora. Los nombres asignados deben ser únicos dentro del agregado, y al igual que los agregados del nombre, deben ser compatibles con DOS. Debido a que las políticas tratan de no incluir en los nombres el significado basado en el contenido o en la semántica, los identificadores de ejemplares están cerrados a números consecutivos. Sin embargo, la mayoría de los ejemplares colocados en los repositorios son conformados por diversos archivos. Las fotografías pueden ser representadas por una imagen de referencia, una imagen mínima y un archivo de imagen de alta resolución sin comprimir; un libro puede tener sus páginas digitalizadas, transcripciones de texto en formato HTML, y versiones SGML; y una carpeta manuscrita puede tener una imagen digitalizada del frente y parte posterior de cada documento en la carpeta. En este caso, los identificadores de ejemplares pasan a ser directorios que contienen los diversos archivos que conforman cada ejemplar. Los archivos pueden ser distinguidos unos de otros por la adición de las extensiones de archivo (es decir, gif o tif), o empleando uno o más de los 8 caracteres aceptados en los nombres de archivos en DOS para representar cosas tales como el uso del archivo (es decir, "m" para mínima), para mostrar las relaciones estructurales (cuando los ejemplares son demasiado largos para ser digitalizados como un todo y tiene que ser segmentado para su digitalización), o para indicar información sobre la constitución física del ejemplar (empleado para reconocer dónde un nuevo documento comienza cuando el identificador de ejemplar es para una carpeta de documentos).

Reconociendo que estos lineamientos son adecuados para el corto plazo y sólo son el comienzo para las necesidades a largo plazo para los repositorios de la Biblioteca, se contrató a CNRI (Corporation for National Research Initiatives) e IBM para explorar caminos para construir un sistema de repositorio con nivel de producción que incluyera las capacidades para cumplir con el

control de calidad en ambiente UNIX, y para crear y mantener suficiente información para recuperar y administrar los archivos. Aunque CNRI ha propuesto un enfoque de programación orientada a objetos e IBM está trabajando en un enfoque de base de datos relacional, se espera que los sistemas puedan operar entre sí dado que ambos esfuerzos emplean el mismo conjunto base de meta datos de elementos físicos, o aquellos elementos necesarios para identificar la información estructural y administrativa requerida para emplear los archivos digitales.

Aunque este proyecto para crear un repositorio digital de la Biblioteca ha dado como resultado el entendimiento de muchos temas complejos, nuevas preguntas sobre la implementación y construcción del repositorio, nombre de archivos, indización, y meta datos descriptivos (es decir, registros MARC y apoyos de localización SGML) surgen continuamente.

PURL's y Nombres de los elementos en Internet

Los URN's (Uniform Resource Name) son identificadores independientes de la localización de recursos en Internet. Los URN's han mejorado a los URL's (Uniform Resource Locator) debido a que estos son contemplados como identificadores únicos globalmente e independientes de la localización, que pueden ser empleados para la identificación de los recursos, y de esta manera facilitar el acceso tanto al recurso por sí mismo como a los meta datos sobre él. Aunque los URL's han sido ampliamente usados y aceptados como método de localización de recursos en Internet, se tienen diversos problemas asociados con el movimiento de recursos de una localización a otra, o cuando las localizaciones mismas cambian de nombre o se vuelven obsoletas. Sin embargo, el concepto de nombre universal ha sido cordialmente aceptado por la comunidad bibliotecaria y por la de red, aun cuando la convergencia en los detalles ha sido elusiva hasta los meses recientes.

Puesto en forma sencilla, los URN's son nombres que identifican recursos o unidades de información independientemente de su localización. El recurso puede incluir ejemplares bibliográficos, ejemplares de colecciones bibliográficas, registros catalográficos, computadoras, buzones de correo electrónico, etc. Los URN's tienen tres componentes: el nombre del esquema, un sistema acordado de nombres únicos reservados (seis esquemas URN han surgido hasta el momento); la autoridad de nombre, una entidad que está autorizada para crear URN's dentro de un esquema específico; una cadena local única, o cadena de caracteres asignada por la autoridad de nombre sólo una vez. Por ejemplo, en el URN "hdl:cnri.dlib/arms-09" el nombre del esquema es "hdl" (referente al esquema Handle, una implementación completa de URN desarrollada por el CNRI), "cnri.dlib" es la autoridad de nombre (un nombre único global dentro del esquema), y "arms-09" es la cadena local única (la cadena debe ser única para la autoridad de nombre dada). Para emplear un URN, debe existir un servicio de resolución que pueda relacionar el nombre al recurso correspondiente y regresar una o más localizaciones. Así, el URN "hdl:cnri.dlib/arms-09" puede relacionarse con el URL "http://www.dlib.org".

En Octubre de 1995, representantes de los seis esquemas de URN's reunidos en la Universidad de Tennessee acordaron trabajar juntos para: adoptar una sintaxis para URN's, desarrollando registros para relacionar nombres de esquemas a los servicios de resolución (debido a que los nombres de esquemas no están asociados a sistemas de resolución específicos); aprobar a otros URN's; y realizar un diseño conjunto de una interfaz para el explorador.

Hasta que la tecnología de URN's se desarrolle y sea ampliamente aceptada, lo impredecible en la localización de los recursos en Internet continuará siendo un problema para los bibliotecarios que intentan proporcionar información URL en los registros catalográficos.

En un esfuerzo para minimizar el mantenimiento de la información de URL's en los catálogos OCLC desarrolló una nueva forma de localización, el PURL (Persistent Uniform Resource Locator). Aunque un PURL se parece mucho a un URN en varios aspectos, un PURL es actualmente un URL, que en lugar de apuntar directamente a la localización de un recurso en Internet, apunta a un servicio de resolución intermedio. Este servicio de resolución asocia el PURL a un único URL actual, que es regresado al cliente (el explorador de WWW) que puede entonces completar la transacción del URL en forma acostumbrada, regresando el recurso de Internet al cliente. Como un "redirect" estándar de http, el PURL es resuelto en forma sencilla empleando los exploradores de WWW en uso actualmente. OCLC mantiene un servidor de PURL, para su proyecto Intecat; OCLC está apoyando la adopción del modelo de PURL poniendo a disposición de otras organizaciones la aplicación de servicio con el compromiso de mantener la persistencia en los esquemas de nombres.

Un programa de colaboración entre el CNRI y OCLC está actualmente en desarrollo para explotar las ventajas de los sistemas Handle y PURL. CNRI agregará el protocolo PURL al sistema Handle, y OCLC agregará las bibliotecas cliente de Handle a los servidores PURL. El resultado final es que las bibliotecas no terminarán con datos en sus catálogos que necesiten ser cambiados cuando el recurso cambie. CNRI y OCLC también planean trabajar con la comunidad bibliotecaria para desarrollar convenciones para los nombres de elementos en Internet, empleando tanto nombres de dominio como PURL's o el universo de nombres Handle. CNRI y OCLC consideran las implicaciones políticas y técnicas de adaptar el servidor local Handle para poder ser empleado como servidor de PURL's

A continuación se mencionan algunas convenciones que es necesario observar para el empleo de URN's en objetos bibliográficos:

- No emplear nombres semánticos (los nombres pueden ser mnemónicos, pero la información implícita en un nombre es codificada como meta dato)
- Dar nombres distintos al contenido que pueda ser usado por separado
- Combinar elementos de contenido que son siempre empleados conjuntamente en un único objeto digital, con un nombre individual
- Emplear meta objetos para describir estructuras y relaciones entre objetos digitales (un meta objeto es un objeto cuyo propósito primario es proporcionar vínculos que describen la estructura de un conjunto de objetos).
- Esperar que los meta datos sean diseminados entre diferentes localizaciones. Los meta datos para un objeto digital individual son almacenados en diversas localizaciones, así como también son almacenados inmersos en el contenido (es decir, un encabezado TIFF), en el repositorio como parte del objeto digital, como parte del meta objeto que se vincula al objeto digital, como meta dato sobre la colección o repositorio (es decir, ayuda para la localización), en un catálogo o índice local, en un catálogo o índice independiente, etc.

1.7. REVISTAS ELECTRÓNICAS

Dentro de las colecciones que están presentes en material electrónico y que pueden ser difundidas por este medio, se encuentran las revistas electrónicas; servicios tales que pueden ofrecer un vasto acervo de publicaciones seriales están presentes en las bibliotecas del mundo.

Se presentan servicios desde donde puedes consultar los nombres, índices, hasta artículo por artículo.

La modalidad de revista en texto completo electrónico ya se está ofreciendo en algunas denominadas bibliotecas electrónicas; los mecanismos varían desde los más novedosos donde la tecnología ha resuelto la edición original en forma electrónica, y entregada directamente al suscriptor por vía electrónica, transmitida directamente del editor al usuario; hasta la otra modalidad con base en la conversión de la revista cuya versión original fue en papel. Lo que no se ha resuelto por completo, dado los diferentes enfoques y la protección a diferentes intereses, es lo que se refiere a derechos de autor, pago de regalías y facturación de servicios, sobre todo en países como E.U.A.

CAPÍTULO 2

CAPÍTULO 2

DISEÑO DE UN ESQUEMA DE BIBLIOTECA ELECTRÓNICA

2.1. OBJETIVOS DE LA BIBLIOTECA ELECTRÓNICA

Una biblioteca es un sistema en el cual un gran volumen de información que proviene de muchas fuentes es ensamblada, organizada, presentada y puesta a disposición sin especificar o considerar el conocimiento que pueda tener el que usa la información.

Una **Biblioteca Electrónica**¹ es una reunión de cómputo digital, un centro de almacenamiento, y equipo de comunicaciones además del contenido y programas de computadora necesarios para reproducir, emular y extender los servicios proporcionados por las bibliotecas convencionales basadas en papel y otros materiales que manejan la catalogación, las colecciones, recuperación y diseminación de información. Un servicio completo de Biblioteca Electrónica debe abarcar todos los servicios esenciales de las bibliotecas tradicionales y también explotar las ventajas del almacenamiento, las búsquedas de información y comunicación de tipo digital.

Los elementos que conforman la definición de la Biblioteca Electrónica son los siguientes:

- La Biblioteca Electrónica no es una entidad única
- La Biblioteca Electrónica requiere tecnología para vincular muchos recursos
- Los enlaces entre varias Bibliotecas Electrónicas y los servicios de información son transparentes para los usuarios finales
- El acceso universal a las Bibliotecas Electrónicas y a los servicios de información es un objetivo
- Las colecciones de las Bibliotecas Electrónicas no están limitadas a los substitutos de documentos impresos, se extienden a elementos digitales que no pueden ser representados o distribuidos en formatos impresos.

Una Biblioteca Electrónica es una biblioteca donde la información es almacenada y procesada en formato digital (el World Wide Web es un ejemplo). El sistema de Biblioteca Electrónica contiene muchos componentes, diferentes estructuras de soporte, y son administradas por diversas organizaciones. La Biblioteca Electrónica es también llamada Biblioteca Digital, el primer término es usado en la Europa y el segundo en América del Norte.

¹ Referencia tomada de los talleres de la IEEE (Instituto de Ingenieros Eléctricos y Electrónicos)

2.1.1.Importancia de la Biblioteca Electrónica

Las Bibliotecas Electrónicas son importantes ya que retienen registros sociales, científicos, legales, entre otros propios de nuestra cultura; ellas proveen un acceso económico y amplio; y también proporcionan acceso a estos registros soportando el desarrollo cultural y económico de la sociedad.

La importancia también radica en que tienen el potencial para ofrecer servicios de biblioteca más efectivos, pueden almacenar información que existe sólo en forma digital y ofrecen nuevas oportunidades para organizar y disseminar información.

La investigación en las Bibliotecas Electrónicas debe enfocarse a atacar las dificultades técnicas que se presentan cuando se desea continuar dentro de la era digital y para realizar sus nuevas funciones, por ejemplo, para interoperar entre bibliotecas digitales o electrónicas.

Para soportar el desarrollo de Bibliotecas Electrónicas, la investigación está dirigida necesariamente a problemas centrales en redes de computadoras; ésta es la clave para el desarrollo de muchas otras áreas de gran interés, tal como comercio electrónico.

2.1.2. Requerimientos de los usuarios y editores para el establecimiento de una Biblioteca Digital o Electrónica

Una Biblioteca Digital es una colección de información electrónica organizada para su uso extensivo. La Biblioteca Electrónica tiene que advertir y satisfacer los requerimientos que tienen los usuarios. Para solucionar las necesidades de los usuarios, los encargados de una Biblioteca Electrónica deben de tomar en cuenta dos tareas: establecer el repositorio de material electrónico, e implementar las herramientas para usarlo. Más importante aún, los comités conformados por los responsables de las bibliotecas son necesarios para que la información esté disponible sobre periodos más grandes de tiempo.

Requerimientos de usuarios para las bibliotecas digitales

¿Qué es lo que los usuarios necesitan de una Biblioteca Digital?. La respuesta une el aspecto histórico, las capacidades y misiones de las bibliotecas tradicionales, con la ciencia de la computación para producir un nuevo servicio, encontrando y satisfaciendo las necesidades definidas de los usuarios. La misión de las bibliotecas digitales es el adquirir información, organizarla, ponerla disponible y preservarla; es importante aclarar que éste ha sido el papel más importante y exitoso de las bibliotecas tradicionales trabajando con material impreso desde hace muchos años. Una misión implícita de la ciencia de la computación es tomar los beneficios de la tecnología para ser usada por la sociedad (incluyendo el campo de las bibliotecas). Pero todo esto necesitará consciencia y esfuerzos planeados por las bibliotecas y el área de la computación, para

lograr que la información proporcionada por las bibliotecas hasta ahora sea mostrada en forma electrónica.

Muchas bibliotecas en la actualidad están intentando proveer un gran volumen de información electrónica a sus clientes.

Cuidados que se deben de tener

Actualmente la información necesita ser proporcionada en forma electrónica en colegios y bibliotecas de investigación alrededor del mundo. Esta situación da por el hecho de que la información electrónica puede ser distribuida de forma rápida y eficiente a diferencia de la tradicional, en mayor cantidad, además de reducir los costos de envío; por todas estas razones se toma como una tendencia mundial. Sin embargo, en el caso particular de las bibliotecas de investigación aun no han tomado la organización y preservación de la información con la misma importancia como se ha hecho para el material impreso; esto es por el hecho de que es una tarea cara, poco explorada y difícil.

Hasta que los administradores de las bibliotecas no tomen a su cargo estas tareas, las soluciones propuestas actualmente sólo tomarán efectos temporales. Por ejemplo, el hecho de catalogar recursos en red debe de tomarse con reserva, hasta que los recursos que comiencen a ser catalogados tengan una presencia permanente en la red y que se garantice que esta catalogación sea de calidad. De la misma forma, el empleo de los productos (de gran costo) para la digitalización desde material de microfilmación a forma digital y viceversa será un riesgo después de algunos años, si los administradores de las bibliotecas no están en capacidad de garantizar su preservación.

Contribución de los usuarios

En la creación de Bibliotecas Electrónicas es necesario tomar en cuenta el papel que juega el usuario como la población que va emplear la información electrónica y contribuir a su creación.

Respecto a este punto, para que la comunidad escolar pueda contribuir a la creación de información electrónica, se debe de asegurar la validez de la misma, así como de su mantenimiento íntegro y auténtico. Es de suma importancia la contribución de la información escolar, ya que ayuda a aliviar algunas crisis de tipo social o económicas, es por ello que las universidades y bibliotecas académicas tienen el interés en ayudar a que las publicaciones electrónicas escolares tengan éxito, y por lo tanto de establecer bibliotecas digitales de investigación seguras.

Necesidades de los usuarios

Para los usuarios de las bibliotecas, la transición de bibliotecas tradicionales a bibliotecas digitales les debe traer más beneficios que dificultades (por ejemplo, técnicas, uso, etc.).

Las necesidades de los usuarios de las bibliotecas continúan siendo las mismas: el encontrar información de interés. Los usuarios requieren información que sea localizada en forma segura, así que cuando vayan al lugar donde se encuentre ésta (asistiendo personalmente a un lugar o a través de la red), ellos encuentren la información por la que buscaron. Desean que la información sea de fácil acceso, que la catalogación sea clara y correcta, además de que se tengan métodos de recuperación apropiados. El usuario busca un servicio global que integre todas las respuestas a sus requerimientos de información, este beneficio para el usuario es el que se debe tomar en cuenta en el diseño de la biblioteca electrónica. En el ambiente electrónico la necesidad de acceder

herramientas es más evidente, y el usuario espera un programa de computadora apropiado y estándar que realmente esté disponible. Los usuarios esperan encontrar en las Bibliotecas Electrónicas la misma información que ha estado disponible por mucho tiempo en la biblioteca tradicional, y que ofrezcan integridad de la misma.

Necesidades de los editores y creadores de información en las Bibliotecas Digitales

Actualmente, la información que está disponible en las Bibliotecas Electrónicas (libros, publicaciones electrónicas, etc.) sufrió el proceso de pasar de un formato en papel a un formato digital, a través de la infraestructura de cómputo adquirida por las bibliotecas. Las bibliotecas también adquieren la información electrónica a través de empresas que se especializan en digitalización. Sin embargo, existe la propuesta de que la información sea originada directamente en formato digital, y sea distribuida en las Bibliotecas Electrónicas.

Independientemente del método empleado para contar con la información digital, el valor y potencial inicial de las Bibliotecas Electrónicas depende de las necesidades de los que originan las obras en ellas. Los que originan las obras son los creadores y editores (personas u organizaciones), los cuales generan, organizan, o crean materiales que son distribuidos en forma digital. Dichas necesidades son:

1. **Diseminación.** La necesidad básica de los que dan origen a la información es una infraestructura que soporte la distribución ampliamente difundida de los objetos de la Biblioteca Digital con poco esfuerzo.
2. **Acceso.** La segunda necesidad es un sistema propio de la biblioteca que provea acceso a objetos de la misma, tales como catálogos e índices, y sistemas para administrar los accesos, tal como autenticación y pagos.
3. **Archivar.** Las personas que originan información esperan que su material sea preservado por un periodo amplio de tiempo. Ellos requieren sistemas que aseguren el acceso no importando cambios futuros que se puedan dar en las organizaciones y en tecnologías.
4. **Control.** Cuando se distribuye el material de los autores, se requiere algún tipo de control sobre como éste es usado. Dicho control incluye el cómo está colocado el material en el dominio público para aplicar restricciones en el acceso. Además se deben incluir toma de decisiones acerca de cómo se puede alterar el material y otras consideraciones de integridad.
5. **Aspecto legal y social.** La diseminación ordenada del material es crucial para una sociedad. Las áreas legales incluyen los derechos de autor y otras propiedades intelectuales, como privacidad. Las prácticas de negocios incluyen políticas de uso aceptable, códigos y contratos estándares.
6. **Herramientas.** Los autores necesitan computadoras, redes, herramientas y programas de computación para la creación en forma ordenada, distribución y acceso a todos los tipos de información.

2.2. ELEMENTOS DE UNA BIBLIOTECA ELECTRÓNICA

2.2.1. Cómo se inicia una Biblioteca Electrónica

Es importante definir qué se debe hacer para que una Biblioteca Electrónica tenga éxito una vez conocidos los requerimientos de los usuarios. El requerimiento primario para una Biblioteca Electrónica, es que desde el inicio se definan procedimientos para organizar, almacenar y proveer información electrónica por periodos de tiempo muy largos. La implantación de una Biblioteca Electrónica requiere seis tipos de tareas:

- Tomar en cuenta los **principios de implantación**
- Establecer el **repositorio electrónico de la información**
- Administrar los **registros electrónicos**
- Definición de las **colecciones históricas digitales**
- Tomar en cuenta los **estándares para imágenes digitales**
- Implementar las **herramientas para el uso de la Biblioteca Electrónica**
- Tomar en cuenta los **Derechos de Autor**.

además de nuevos compromisos. Siguiendo estas tareas los problemas técnicos son más fáciles de resolver.

2.2.2. Principios de implantación

Cuando se construye una Biblioteca Electrónica, es importante considerar principios clave para que dichas bibliotecas sean fácilmente empleadas, y tengan valor archivístico de largo plazo.

- Deben emplearse representaciones aseverativas de los documentos (libros, revistas, enciclopedias, etc.).
- Los componentes de los documentos deben ser representados empleando formas naturales, nombrando a los objetos de tal forma que puedan ser manipulados por los usuarios.
- Los vínculos (elementos electrónicos que representan los documentos de la biblioteca) deben ser almacenados, preservados, organizados y generalizados.
- Debe existir una separación clara entre la Biblioteca Digital y la interfaz con el usuario.
- La búsqueda de información debe hacer uso de métodos de recuperación avanzados.
- Los sistemas de cómputo abiertos que incluyen al usuario, y donde muchas de las funciones de los bibliotecarios son llevadas a cabo por la computadora, también deben desarrollarse.

- El acceso a los archivos electrónicos orientado a las actividades inherentes y propias (organización y flujo de información entre otras) de la biblioteca, también debe ser soportado.
- El enfoque de desarrollo orientado al usuario debe ser adoptado.
- Los usuarios deben trabajar con objetos al nivel adecuado de generalidad.

2.2.3. Repositorio electrónico

La Biblioteca Electrónica será para los usuarios como una colección de información existente en varios lugares (no siempre evidentes), accesible mediante el uso de herramientas ampliamente disponibles. El lugar de la información es llamado el repositorio electrónico.

Existirán muchos repositorios electrónicos que responderán a los requerimientos de redundancia (tener la misma información almacenada en distintos lugares) y las necesidades individuales de las instituciones, situación que no se presenta en la colección impresa de información. De esta forma, si por ejemplo estamos hablando de aplicaciones en red (Internet como caso concreto) dicha información podrá ser accedida desde cualquier parte del mundo.

Los repositorios que son parte de las Bibliotecas Electrónicas tendrán muchas características en común.

- **Contenido de megadocumentos:** Un repositorio inicial puede incluir muchos gigabytes de información, lo que da un desarrollo rápido de millones de documentos electrónicos. El medio para almacenar la información (disco) es económico y los posibles recursos son abundantes.
- **Fuentes participantes y potenciales:** Es fácil citar el número de instituciones escolares que participan en la construcción del repositorio electrónico, algunos ejemplos son:

Escuelas de nivel medio superior

Universidades

- **Mecanismos de respaldo:** Los procedimientos de respaldos y restauración de información deben efectuarse en algún lugar, ser automatizados y ser económicos. Los respaldos deben abarcar a todas las dependencias donde está distribuida la misma información, es decir, ligas de esta información, en ellos se debe usar almacenamiento remoto, con simulaciones y pruebas de pérdida y recuperación de información.
- **Acceso representado:** En el ámbito computacional, "representar" se refiere al empleo de métodos mecánicos para almacenar datos que esperan ser grabados para posteriormente ponerlos en línea cuando son solicitados. No todos los datos necesitan estar disponibles en el medio más caro y rápido de almacenamiento (lo que implica estar en línea). Deben existir alternativas para proveer acceso en línea inmediato de enormes cantidades de información. ¿Qué es lo que puede estar fuera de línea y cómo pueden ser recuperables? Presentes alternativas incluyen discos magnéticos, disco ópticos, discos ópticos en gavetas, cintas magnéticas locales y remotas y métodos de almacenar automáticamente en cintas magnéticas.

- Estándares en la estructura de datos: En un repositorio, ¿la información existe simplemente como es (es decir, como fue creada en un principio) o existe información complementaria (metadatos) asociada con ésta? Muchos ejemplos incluyen encabezados SGML (Standard Generalized Markup Language), códigos de libros, encabezados de descripción, ligas de hipertexto y versiones de software para usar con los archivos de datos. Existe la necesidad de ligar imágenes de páginas que fueron transformadas a ASCII con las versiones de texto del contenido de las mismas. Si existe una asociación es a través del uso de encabezados de un archivo o a través de archivos anexos.
- Actualizando mecanismos: El refrescar o actualizar mecanismos es aceptado por el hecho de que se debe dar la preservación de estos a través de los avances en técnicas de cómputo y programas. Algunos ejemplos de esto pueden ser: el hecho de cuando la información es copiada desde un disco magnético u óptico a otro medio los derechos de autor deben ser reconocidos; la automatización deberá necesariamente reducir costos. Otros factores para que se de la actualización de mecanismos incluyen el flujo de trabajo y el hecho de guardar registros, técnicas de migración y estándares y técnicas que aplicarán independientemente de la tecnología.
- Autenticación e integridad. La preservación intelectual va mas allá de la preservación del medio y de la tecnología, se tiene que asegurar la protección de la estructura intelectual de la información tal como fue creada por su autor. Las Bibliotecas Electrónicas deben implementar técnicas que combinen seguridad matemática (código de redundancia cíclica) de fácil uso, integridad pública y protección privada. Por ejemplo, patrones de bits de textos en los documentos, los sonidos y las imágenes deben ser preservadas a través de criptografía y métodos de codificación tales como las técnicas de marcas de tiempo digital.
- Redundancia: Es muy importante establecer estándares para determinar el número de sitios repositorios necesarios para asegurar la existencia y el acceso de cierta información electrónica. ¿Cuántos lugares serán necesarios para ello? Las instituciones separadamente o en conjunto deberán definir los repositorios; aún no está claro cuanta redundancia es la deseada y necesaria. Se ha pensado en tener múltiples copias permanentes de la información electrónica. La localización geográfica, nacional, regional y hasta continental, jugará un importante papel para este tipo de situaciones. Se tendrán que tomar en cuenta para esto ciertos puntos, tales como la relación entre el costo de la respuesta de la red y el costo por almacenar la información. Muchas bibliotecas tendrán que platicar de la redundancia con los proveedores de servicios de red.

2.2.4. Registros electrónicos

El reto electrónico

La revolución electrónica está cambiando de forma definitiva la manera en la que los registros se crean, mantienen y administran. Este proceso tiene un efecto dramático en la posibilidad de proporcionar una rápida respuesta a cualquier tipo de reporte solicitado.

Un aspecto fundamental de nuestra sociedad es la historia en forma de registros, que son creados, mantenidos, preservados y que se espera estén a disposición en un futuro cercano. Con la creciente cantidad y diversidad de registros electrónicos, enfrentamos el reto de desarrollar estrategias, estándares y procesos para asegurar que los registros electrónicos estén a disposición tanto tiempo como sea necesario.

Las bibliotecas tienen una responsabilidad muy grande al jugar un papel de líder en el régimen de administración de los registros, independientemente de su formato, asegurando de esta forma que los registros archivísticos sean debidamente identificados, preservados y permanezcan a disposición del público en cualquier momento y que no sean destruidos en forma inadvertida.

¿Qué es un registro? La adivinanza electrónica

Los registros han sido tradicionalmente definidos como objetos físicos, como documentos de papel, cintas, discos, etc. Dentro de una biblioteca tradicional un registro es considerado desde la ficha técnica en papel de los ejemplares en la biblioteca (libros, revistas, colecciones particulares etc.), hasta el mismo ejemplar. De la misma forma un registro puede ser cada uno de los elementos de una colección, o una representación resumida de un libro o colección, de esta forma el término registro es muy versátil.

Tales definiciones son problemáticas cuando se trata de registros electrónicos. En una primera instancia, un disco puede contener registros. Sin embargo, si el disco no puede ser leído, el registro efectivamente ya no existe. Con los registros electrónicos, bajo esta perspectiva, el objeto físico o disco no es considerado un registro.

El registro digital

Ya no se define registro en términos de los objetos físicos que los contienen. Un registro puede ser contenido o creado empleando diferentes medios y formatos. Contrario a la visión de Marshall McLuhan, el medio no es el mensaje. Los registros pueden ser creados y mantenidos en archivos tradicionales basados en papel o digitalmente en un ambiente computacional, existiendo en un solo medio o, con la actual tecnología, como multimedia (texto + sonido + imágenes). Los registros pueden continuar existiendo aun cuando supuestamente han sido eliminados del medio que los almacenaba.

Dentro del ámbito electrónico y como una tendencia muy empleada un registro para ser considerado como evidencia, debe poseer contenido, estructura y contexto.

Además, los registros pueden ser transferidos de un medio a otro y de un contexto a otro a través de la copia de imágenes o la transferencia digital. Los registros electrónicos, en particular, están basados en las transferencias que ocurren dentro y entre computadoras. En muchos de los casos, dichos registros pueden ser actualizados, borrados, alterados o manipulados sin intervención humana. En este proceso, las características esenciales del registro - contenido, estructura y contexto - pueden ser alteradas o perdidas. Las computadoras no crean o mantienen registros; una

planeación e intervención específica son requeridas para asegurar que las características esenciales del registro sean colocadas en los sistemas de información electrónicos y sean mantenidas.

Cambiando la tradición

Tradicionalmente, los archivos han mantenido las características esenciales de los registros - contenido, estructura y contexto - preservando los portadores físicos de los registros en el orden original en el que fueron creados y acumulados. En el caso de los registros electrónicos, el concentrarse en la preservación de los portadores físicos no es suficiente. Primero, como se menciona, los sistemas de cómputo deben ser específicamente diseñados con la funcionalidad de mantenimiento de registros para asegurar que las características esenciales de ellos serán conservadas. En segundo lugar, mientras los portadores (discos, cintas, CD-ROM, etc.) pueden existir por periodos de tiempo relativamente largos, los registros que almacenan pueden dejar de existir debido a que la tecnología necesaria para su recuperación no está disponible o soportada por el constructor. Considérese el caso de tratar de recuperar un registro de una aplicación actual en treinta años.

En resumen, medidas específicas deben ser tomadas en los sistemas de cómputo para garantizar que los registros - en términos de sus características esenciales - sean mantenidos. Sin tales medidas, toda la comunidad sufriría pérdidas irreversibles. Registros sin contexto, estructura o contenido son simplemente "basura"; los discos que no pueden ser leídos son sólo plásticos.

La Estrategia

Los mejores prospectos para la administración de registros electrónicos y asegurar su disponibilidad todo el tiempo son aquellos registros que permanecen en las bibliotecas, que los crean o manejan. Esta estrategia asegura que las características esenciales de los registros son mantenidas. Además, las bibliotecas en forma individual entienden mejor sus sistemas electrónicos y las aplicaciones específicas requeridas para mantener sus registros. Más aún, esta estrategia asegura además que los registros permanezcan disponibles. Como la tecnología va cambiando, las bibliotecas tienen mejor oportunidad de asegurar que los registros de valor incalculable sean transferidos satisfactoriamente o migrados a un sistema más evolucionado.

El éxito de este enfoque requiere de la identificación y disposición permanente de registros con valor administrativo limitado y la preservación del acceso a aquellos registros que poseen valor administrativo constante. Los sistemas de almacenamiento y manejo de registros eficientes redundan en una planificación estratégica y administración del control. Los sistemas de almacenamiento y manejo de registros sin planeación y sin control constituyen un peligro para la administración, sus costos y por último, para la comunidad.

El papel del Archivo

Si bien las bibliotecas están en posición de manejar registros electrónicos, no cuentan con los recursos tecnológicos para manejar los registros electrónicos de todos los sistemas y aplicaciones existentes. Aun cuando los recursos estuvieran disponibles, al tomar este papel, las bibliotecas se convertirían en museos de tecnología obsoleta. Como consecuencia de ello, la biblioteca sólo tendrá bajo su custodia aquellos registros electrónicos que puedan ser manejados por sus recursos tecnológicos.

Estos casos son típicamente cuando:

- La biblioteca que ha creado los registros electrónicos está por desaparecer o ya desapareció y no hay aun una definición sobre quién asumirá las funciones y el control del acervo
- La biblioteca suscribe un convenio para compartir el acervo y la custodia de los registros electrónicos

Cuando una biblioteca acepta la custodia de registros electrónicos, es necesario asegurar que los registros son transferidos apropiadamente y que están acompañados de la información y metadatos necesarios para mantener el acceso a ellos.

El éxito de esta estrategia depende de la estrecha cooperación y apoyo entre bibliotecas. Se trabaja activamente en el establecimiento de políticas, guías y estándares para apoyar la administración y manejo de registros electrónicos. Además se trabaja en la identificación de registros electrónicos actualmente en custodia y que necesitan ser administrados y puestos a disposición para cumplir con los requerimientos de los usuarios, identificar y adaptar los registros que no tienen considerable valor, reconocer los metadatos o la información que se necesita capturar y mantener para que los registros de gran valor permanezcan identificables y disponibles, dar consejo en cuanto al acceso a los registros electrónicos archivados para cumplir con los estándares y mantenerlos seguros, dar asistencia en el desarrollo de sistemas de almacenamiento y manejo de registros determinando cómo y cuándo se deben crear.

En resumen, es necesario trabajar para asegurar que los sistemas de administración de registros electrónicos sean manejables y permitan consolidar los registros electrónicos, además de identificar los procedimientos de administración adecuados. Además es necesario determinar el periodo de tiempo durante el cual los distintos tipos de registros electrónicos deben ser mantenidos. El objetivo es asegurar que no se utilicen recursos para mantener registros efímeros o no necesarios.

2.2.5. Colecciones Históricas Digitales: Tipos, Elementos y Construcción

Tipo de Colecciones

Colecciones históricas. En la Biblioteca del Congreso, la piedra angular para los materiales históricos en línea es la colección, un cuerpo coherente de materiales relacionados. Los tipos de colecciones incluyen:

Grupos archivísticos

Algunos ejemplos incluyen papeles personales de algún personaje, o los negativos fotográficos de alguna compañía.

Conjuntos acumulados

Algunos ejemplos incluyen conjuntos de impresiones y documentos dentro de un género o categoría y acumulados por la división de colecciones especiales de la Biblioteca durante años.

Compilaciones especiales

Grupos de materiales relacionados especialmente ensamblados para un proyecto digital.

Elementos de las Colecciones Digitales de la Biblioteca

Cada colección en línea de una Biblioteca consiste de los siguientes elementos:

Marco. Es un conjunto de archivos de texto con formato electrónico, como por ejemplo HTML, con imágenes los cuales sirven como la página inicial de la colección. Este marco es así nombrado, debido a que proporciona un entorno intelectual para la colección y acoge a los otros elementos de la misma. La familia de textos incluye resumen y notas de contenido, cronologías, notas técnicas y bibliografías. Este marco puede ser comparado con la página del título principal de un libro, tabla de contenidos, portada y contraportada.

Ayuda al acceso. Es el conjunto de datos que describe a los ejemplares en la colección y facilita la búsqueda y revisión del usuario. La ayuda para el acceso puede ser considerada como metadatos descriptivos. En las colecciones de la Biblioteca del Congreso en línea, la ayuda al acceso más frecuente y familiar es el catálogo específico de la colección que consta de registros bibliográficos. Para algunas colecciones en las Bibliotecas, el acceso es proporcionado por un conjunto de menús que sirve como ayuda a la localización de información. En el futuro, el acceso a algunas colecciones será proporcionado por ayudas a la localización con base en el modelo Encoded Archival Description (EAD).

Ejemplares. En el contexto de las colecciones digitales, la Biblioteca emplea el término ejemplar en una forma flexible. Dependiendo del juicio de los archivistas y diseño de la colección, un ejemplar puede ser:

- una fotografía
- un grupo de fotografías íntimamente relacionadas
- una carpeta de manuscritos conteniendo un número de documentos
- un contenedor de carpetas conteniendo un número de carpetas
- un libro reproducido como texto de búsqueda completa y un conjunto de imágenes de páginas
- un video o película
- una grabación de audio
- una revista

La definición archivística de un ejemplar reflejará la determinación del nivel de detalle deseado (o producido) en el sistema de ayuda de acceso a mano. Cuando altos niveles de detalle son ofrecidos, los ejemplares de la colección será verdaderamente una unidad discreta, por ejemplo un ejemplar puede ser un fascículo, inclusive un artículo. Cuando descripciones a nivel de grupo son proporcionadas, el término ejemplar se tomará en realidad como un grupo de ejemplares. Los lectores deben notar, de todas formas, que en este contexto el ejemplar es definido como la unidad de contenido referenciada por un registro bibliográfico individual o por una línea en el sistema de ayuda para la localización. En los registros bibliográficos de la Biblioteca, los ejemplares son referenciados por el subcampo f del campo 856 MARC². La Biblioteca aun no ha establecido una convención paralela para los sistemas de ayuda de localización EAD.

Reproducciones. Son las imágenes digitales, textos para búsquedas, archivos de grabaciones sonoras y archivos de video que reproducen ejemplares originales. Algunos ejemplares son reproducidos por una combinación de los anteriores. Cuando son almacenados en un sistema de cómputo, las reproducciones digitales contienen y/o son asociados con metadatos administrativos y estructurales, los cuales pueden ser empleados para obtener la localización de las reproducciones en el sistema, reporte del tipo o formato digital específico, control de acceso, indicaciones de relaciones para las reproducciones multipartes, además de proporcionar más información técnica.

Programas suplementarios. Son los elementos que dan una introducción a los usuarios de las colecciones, explicando el contenido y cómo pueden ser empleadas. Llamadas Presentaciones Especiales en los sitios WWW de las Bibliotecas, en su forma más simple, son textos breves o “presentaciones”, mientras que en las más elaboradas, toma la forma de programas interactivos multimedia.

Diversas Ayudas de Acceso

Las colecciones históricas digitales incluyen una variedad de elementos que pueden ser empleados para el acceso. Esto no sólo incluye los registros catalográficos y ayudas para la localización mencionados con anterioridad, sino también textos completos (cuando, por ejemplo se dice que un libro impreso ha sido convertido en un texto electrónico con búsqueda), y material explicativo introductorio. Aunque en la Biblioteca no han instituido un sistema de búsqueda que contemple todos estos elementos (desde luego, ayudas de localización EAD no están vinculados al contenido digital en el sitio de la Biblioteca), ésta prevé que un sistema de búsqueda con amplio alcance será establecido a mediano plazo.

² Ver definición del campo 856 MARC en el glosario de términos

Cuando la ayuda para el acceso toma forma de registro bibliográfico, el nivel de perfección catalográfica puede variar. En algunos casos, miembros de la Biblioteca del Congreso producen catalogaciones completas apegadas a las reglas y autoridades más relevantes; en otros, una variedad de simplificaciones, variaciones, y faltas de ortodoxia son empleadas o toleradas. Para algunas colecciones, notas en prosa por ejemplar han sido proporcionadas. En la medida de lo posible (cuando existen suficientes recursos), términos temáticos han sido establecidos para los encabezados temáticos de la Biblioteca del Congreso (Library of Congress Subject Headings LCSH) o Tesoros de Materiales Gráficos de la Biblioteca del Congreso (Library of Congress Thesaurus for Graphic Materials LCTGM).

La existencia de una diversidad de arreglos de elementos de acceso inevitablemente disminuye la efectividad de los registros bibliográficos clásicos - registros que emplean un vocabulario estructurado y ponen atención especial en el uso de las reglas para nombres y temas- dentro de grandes organizaciones. Términos de temas asociados a los Tesoros, por ejemplo, no siempre están presentes en una base de datos de trabajo que proporciona acceso a nivel de ejemplar dentro de una colección individual, en una ayuda para la localización EAD que acumula registros manuscritos tradicionales, o en búsquedas de textos completos proporcionadas por algunas colecciones de manuscritos y de material impreso, cuyo vocabulario consiste de términos y palabras en inglés ordinario.

La Biblioteca espera desarrollar ayudas adicionales para acceder como un sistema de búsqueda lo suficientemente poderoso para manejar el vocabulario estructurado y no estructurado. Un primer paso será el crear un catálogo a nivel de registros bibliográficos para las colecciones históricas. Esta base de datos de registros de colecciones servirá como punto de entrada para muchos usuarios.

Identificadores, Nombres lógicos y el Repositorio Digital

La Biblioteca está desarrollando un esquema para la asignación de nombres o identificadores para las reproducciones digitales. Estos identificadores ligan las reproducciones y las ayudas para el acceso, como una biblioteca emplea números para ligar los registros catalográficos y los libros por sí mismo.

¿Por qué los identificadores son necesarios?

En el repositorio digital de la Biblioteca del Congreso las reproducciones se mueven de un dispositivo a otro, principalmente cuando los datos son respaldados o cuando el equipo es reemplazado. Esta migración de datos requiere que los identificadores sean lógicos en vez de nombres físicos. La localización física de los archivos en el repositorio será registrada en un sistema localizador de documentos. Así, la solicitud hecha al sistema de ayuda para el acceso recuperará el identificador (nombre lógico) del ejemplar deseado. Cuando se envía al sistema localizador, el nombre lógico será traducido en el nombre físico actual, y la reproducción será recuperada.

El empleo extensivo de contratistas por parte de la Biblioteca para crear reproducciones ha conducido a practicar el empleo de identificadores como nombres físicos al momento de la entrega. Los materiales digitales producidos por contratistas (y frecuentemente también por personal de la propia Biblioteca) son entregados en CD-ROM's. De esta manera, la Biblioteca requiere que los nombres de los directorios se ajusten a los requerimientos de DOS.

El actual sistema de identificación trabaja dentro del sistema cerrado de servidores de la Biblioteca y no proporciona referencias a la institución por sí misma, es decir, a la dirección del servidor que tiene los datos (incluyendo el nombre del dominio para la Biblioteca del Congreso). Así, si el sistema de ayuda al acceso con los actuales identificadores fuera empleado, por ejemplo, para el catálogo de otra biblioteca, los datos no podrían ser empleados para vincular las reproducciones en los servidores de la Biblioteca. El desarrollo del repositorio de la Biblioteca está pensado para resolver esta circunstancia. La Biblioteca está analizando la opción de tratar los identificadores de dos partes como una unidad individual, agregando la referencia perdida, y formalizando el resultado como un URN (Uniform Resource Name). Si se adopta, el empleo de URN cambiará la estructura del identificador y su colocación en el registro bibliográfico. Los lectores reconocerán que el empleo propuesto de URN parte del mismo impulso que conduce a la definición de PURL's (Persistent Uniform Resource Locators), que pueden servir como identificadores de archivos.

Ejemplares, Agregados e Identificadores

La Biblioteca está asignando los identificadores descritos en la sección anterior para ejemplares y grupos de ejemplares, llamados agregados. Esto es, en efecto, las dos partes de los identificadores.

Cuando son archivados en el repositorio digital, los ejemplares de reproducción son colocados en grupos llamados agregados. Colecciones más pequeñas, individuales y medias pueden ser representadas por un agregado individual, mientras colecciones complejas y más grandes pueden ser representadas en múltiples agregados. Una práctica reciente de la Biblioteca es colocar ejemplares similares en agregados separados, es decir, un agregado contendrá reproducciones de fotografías, otros libros y grabaciones de sonidos.

El establecimiento de agregados separados para agrupamientos de ejemplares en formatos originales permite a la biblioteca regularizar la ejecución de rutinas de presentación y despliegue. La presentación en línea y despliegue de una fotografía, es diferente de la presentación y despliegue de un libro. Cuando la biblioteca presenta el registro bibliográfico de una fotografía, por ejemplo, el sistema acompaña el registro con el despliegue de una versión mínima de la imagen. En el caso de un libro, el sistema de ayuda para el acceso proporciona vínculos al conjunto de imágenes de las páginas. Teniendo fotografías en un agregado y libros en otro, el sistema conoce el tipo de ejemplar y puede activar las rutinas de entrega necesarias. En los registros bibliográficos, los agregados son referenciados en el subcampo f del campo 856 MARC.

Como se mencionó anteriormente, la biblioteca investiga la posibilidad de emplear un URN (Uniform Resource Name) para la vinculación; si se adopta, el URN puede ser colocado en un subcampo del campo 856 MARC aun por determinar.

Referencia única, diversas partes o versiones

Debido a que un ejemplar de una colección puede ser multiparte, la presentación en línea debe ofrecer al usuario diversos medios para recuperar todas las partes tomando como base un identificador único proporcionado por el sistema de ayuda para el acceso. Aunque ejemplos como carpetas de manuscritos multi-imagen salta a la mente, vínculos de uno a muchos existen frecuentemente cuando el ejemplar descrito es un original singular. *Un cartel teatral grande*, por ejemplo, puede ser reproducido por (1) versiones en alta y moderada resolución (dos reproducciones) o (2) una imagen conteniendo un segmento del original (siete reproducciones). De manera similar, un libro puede ser descrito en un único registro bibliográfico pero debido a su longitud, ser presentado en línea como un conjunto de capítulos.

Una diversidad de dispositivos pueden ser empleados para proporcionar a los usuarios varias reproducciones (o parte de reproducciones) que están vinculadas vía una referencia. La forma más simple pero menos elegante es el crear una tabla de contenidos o un menú (por ejemplo la actual versión de la Memoria Americana a través de la colección de los libros de notas de Walt Whitman). Aquí, un archivo HTML lista las páginas por número y vincula cada una con su reproducción en imagen. Como una opción alterna, algunos observadores han sugerido considerar archivos PDF multipáginas (el Portable Document Format fue desarrollado por Adobe Corporation y los archivos PDF son desplegados en la aplicación Acrobat).

Mientras tanto, otras organizaciones están desarrollando modelos para paginar un documento empleando conjuntos de imágenes que son mostradas “en línea” (en el visualizador). Un grupo de está desarrollando un encapsulado, mientras que un grupo en la Biblioteca de la Universidad de California Berkeley, están desarrollando lo que ellos llaman Ebind. Un enfoque de paginación a sido aplicado a un conjunto de publicaciones seriales en el proyecto de la Universidad de Cornell y la Universidad de Michigan “Making of America”. Estas presentaciones proporcionan tablas de contenidos, conjuntos de imágenes de páginas (en formato GIF, a resolución media), herramientas para la revisión (paginación), y en algunos casos, acceso a versiones de alta resolución de las imágenes para su impresión o entrega.

Cada una de estas presentaciones tienen sus propias ventajas y desventajas; la biblioteca prevé que se emplearan las tres todo el tiempo, haciendo su selección después de analizar las necesidades de la colección en particular. Mientras se implementan enfoques adicionales para la paginación o selección de las partes de un ejemplar multipartes, la Biblioteca continuará confiando en los archivos HTML con tabla de contenidos.

Produciendo los elementos, ensamblando la colección

Personal de la biblioteca produce marcos de referencia, sistemas de ayuda para el acceso y programas introductorios. Hablando en forma general, los marcos, los sistemas de ayuda para el acceso y programas introductorios son producidos por personal de la Biblioteca del Congreso, típicamente especialistas en la división de Colecciones Especiales. Una gran variedad de aplicaciones son empleadas: el sistema de catalogación basado en equipo mainframe (MUMS), aplicaciones MARC (Minaret, en ejecución en sistemas UNIX y DOS), y diversos programas de bases de datos y procesamiento de palabras.

Cuando los registros de datos destinados para servir como registros bibliográficos han sido creados en un procesador de palabras o aplicación de base de datos, son identificados con etiquetas y delimitadores de campos, y son introducidos subsecuentemente en Minaret o en otra aplicación MARC. Adicionalmente, algunos registros creados en MUMS son editados en Minaret, al tiempo que ciertos campos o datos que no son manejados por MUMS pueden ser agregados. Mientras tanto, la Biblioteca está muy interesada en adelantar la simplificación de la catalogación representada por el desarrollo del conjunto Dublin Core de campos de información bibliográfica y se espera iniciar la implantación de este enfoque en un futuro cercano.

La Biblioteca también está explorando otras opciones para los sistemas de ayuda para el acceso. Las actuales actividades de producción sugieren el considerar las bases de datos relacionales, el contenido de las cuales pueden tener algunas de las características de los registros bibliográficos. Mientras tanto, la Biblioteca está produciendo el primer sistema de ayuda para la localización con contenido vinculado en línea EAD y está investigando aplicaciones opcionales para una eficiente producción.

Los contratistas y el personal de la Biblioteca *elaboran reproducciones*. En ambos casos, equipo y técnicas altamente especializadas son empleados. Los requerimientos son bosquejados en las diferentes etapas del trabajo (para ser colocadas en línea en un futuro muy cercano), se encuentran en las solicitudes de propuestas (RFP) y son asociados con los contratos de digitalización de la Biblioteca del Congreso.

Múltiples reproducciones. Los proyectos de la Biblioteca frecuentemente producen múltiples reproducciones digitales. Por ejemplo, una fotografía histórica puede ser digitalizada como:

1. imagen digital sin compresión de alta resolución (excelente calidad)
2. imagen digital comprimida de alta resolución (*muy buena calidad*)
3. imagen digital comprimida de resolución media (buena calidad)
4. imagen reducida (para presentación con el registro bibliográfico)

Para documentos convertidos a formas de texto plano para las búsquedas:

1. texto con etiquetas SGML
2. texto con etiquetas HTML
3. conjunto de imágenes de páginas que apoyen a los textos manejados por la computadora

Se espera que en el futuro cercano cuando los visualizadores manejen versiones SGML o HTML para que puedan ser generados en tiempo real.

En algunos casos, un proyecto producirá reproducciones analógicas y digitales. Por ejemplo, el deseo de crear copias de alta resolución y gran duración para la preservación ha permitido a los conservadores de fotografías recomendar un procedimiento de producción que, primero genera una película intermedia y, segundo una reproducción digital. Debido a que el futuro de los actuales formatos digitales de imágenes en movimiento es incierto y, a que la mayoría de los formatos de video digital son derivados del analógico, los planes de la Biblioteca para colecciones de video siempre incluye, primero, la producción de una cinta maestra y, segundo, la producción de las versiones digitales.

La Biblioteca archiva todas las versiones de las reproducciones pero generalmente selecciona uno o dos para el acceso en línea. En los ejemplos mencionados, la Biblioteca puede proporcionar acceso en línea a las imágenes digitales comprimidas de una fotografía y archivos de video digitales.

Contribuciones a la preservación

La presente discusión resalta las incertidumbres que rodean el papel de las reproducciones digitales en las actividades de preservación de la Biblioteca. ¿Cómo y cuándo puede el proyecto de digitalización contribuir a la preservación?. El personal del NDPL cree que el proyecto de digitalización puede hacer tres contribuciones:

1. Para algunos materiales, las reproducciones digitales reproducen fielmente al original y, si perdura, competirá o sobrepasará reproducciones analógicas como copias de preservación. El NDLP cree que éste es el caso de manuscritos, material impreso, y grabaciones sonoras. Además también puede ser el caso de fotografías hasta cierto tamaño, es decir, negativos de 35 mm.
2. Relacionado a lo anterior, pero cuando la reproducción es de menos fidelidad, las copias pueden servir a los usuarios típicos. Por ejemplo, imágenes digitales de fotografías con resolución espacial de 4,000 x 3,000 píxeles no pueden capturar cada uno de los matices del original pero son capaces de ser útiles virtualmente a las necesidades de cada usuario. Así, para una colección de imágenes históricas con valor artístico moderado, como es un conjunto de imágenes digitales, si perduran, proporcionarán el mismo servicio que un conjunto de copias de negativos. En una época de recursos limitados, tales copias pueden ser suficientes para algunas colecciones. Este enfoque será especialmente atractivo cuando los materiales originales pueden ser conservados, es decir, proporcionando almacenamiento frío para evitar el deterioro de los negativos antiguos.
3. Para algunos proyectos, el esquema de producción creará copias analógicas de preservación y reproducciones digitales. Por ejemplo, un proyecto para preservar la colección de negativos de nitrato se puede crear un conjunto de interpositivas analógicas que son luego digitalizadas para obtener la reproducción digital.

La frase “si perdura” en los puntos 1 o 2 revela la principal ansiedad sobre la información digital: ¿hará inservible a la información digital la obsolescencia de las aplicaciones, del equipo y de los traductores de medios?. La Biblioteca comparte la ansiedad pero al mismo tiempo espera que el repositorio de la información digital referida en cualquier parte de este documento - con sus respectivos procedimientos de respaldo y migración de datos- tendrá éxito en el mantenimiento de la información digital.

Ensamblaje de los elementos

Una vez producido, el marco, es decir el sistema de ayuda para el acceso, las reproducciones y programas introductorios son almacenados en la aplicación de recuperación: una vez almacenado, los vínculos son actualizados y la colección ensamblada está lista para usarse.

El vehículo para el acceso principal a la Biblioteca es el World Wide Web (WWW). Las colecciones son actualmente almacenadas y consultadas a través de los vínculos de hipertexto WWW; el sistema de ayuda para el acceso esta indexado por aplicaciones de búsqueda y solicitudes. Mientras tanto, la Biblioteca recibe de buen grado iniciativas del sector privado que propiciarán otras formas de diseminación, es decir, publicación de CD-ROM o servicios en línea comerciales.

2.2.6. Estándares para imágenes digitales

Asociación para el Manejo de Información e Imágenes

Los estándares para publicaciones que son desarrollados por científicos, ingenieros, y administradores de negocios, a través de los comités de estándares de la Association for Information and Image Management (AIIM), pueden ser aplicados al manejo de imágenes electrónicas, cuyo proceso incluye la transferencia de documentos (en imagen), recuperación y evaluación, digitalización de documentos y almacenamiento en discos ópticos, diseño de documentos y conversión. Cuando se combinan los sistemas de planeación y operación, los estándares pueden apoyar a la generación de bases de datos de imágenes que son intercambiables entre una gran variedad de sistemas. Las aplicaciones con distintos enfoques de etiquetado de imágenes, indización, compresión y transmisión, frecuentemente provocan inconsistencias en la compatibilidad de sistemas, calibración, desempeño y compatibilidad a futuro, hasta que los parámetros estándares de implantación son establecidos. Los estándares de AIIM que están en desarrollo para este tipo de aplicaciones pueden reducir considerablemente la incertidumbre e incompatibilidad, integrando exitosamente los procesos sobre las imágenes, y promocionando los "sistemas abiertos". AIIM es un desarrollador de estándares acreditado por la American National Standards Institute (ANSI), con más de 20 comités compuestos de 300 representantes voluntarios de usuarios, vendedores y productores. La publicación de los estándares que son desarrollados en tales comités gozan de aceptación y proporcionan la base de la estandarización internacional en desarrollo por la International Organization for Standardization (ISO).

El desarrollo de estándares por parte de AIIM incluye proyectos de bases de datos como un marco de referencia amplio para la industria del manejo de imágenes incluyendo captura, grabación, procesamiento, duplicación, distribución, despliegue, evaluación y preservación. La Base de Datos de Imágenes de AIIM contemplará estándares de formato de imágenes en desarrollo por organizaciones en todo el mundo. Contendrá además formatos estándares de fechas de publicaciones, orígenes, proyectos nacionales e internacionales relacionados, estado actual, palabras clave y resúmenes.

Características de los estándares para el manejo de imágenes digitales

- La naturaleza de la tecnología digital implica volatilidad continua
- El establecimiento de estándares en forma precipitada no es posible y probablemente no deseable
- Los estándares constituyen un problema complejo que involucran el medio, el equipo, las aplicaciones, y la capacidad técnica para la reproducción fiel y clara
- El potencial significativo y atractivo de la tecnología digital es el verla como medio de preservación y mecanismo de acceso
- El uso productivo de imágenes digitales para la preservación requiere de una reconceptualización de los principios de preservación en un mundo volátil sin estándares
- Importancia del concepto de manejo continuo de acceso en ambiente digital en lugar de enfocarse en la permanencia del medio y estándares de largo plazo desarrollados principalmente para el mundo analógico.

Periodo de transición: ¿Cuanto tiempo y qué hacer?

- Redefinir acervo
- Eliminar el peso de "copias del acervo" de los artefactos de papel
- Uso de tecnología digital para almacenamiento, desarrollo de estrategias de mantenimiento y administración para el renovación del medio, equipo y aplicaciones
- Crear copias en papel libre de ácido como respaldo para el período de transición hasta desarrollar procedimientos que aseguren el acceso continuo a archivos digitales

2.2.7. Herramientas de acceso y políticas

- Mecanismos de uso y recuperación: Las herramientas de acceso a la información electrónica deben ser soportadas por la Biblioteca Digital (por ejemplo, los catálogos en línea y los OPACs, FTP, gopher, Worl Wide Web y sus múltiples clientes), además del empleo de nuevas herramientas. Un punto importante al respecto es el hecho de "granular" los documentos: ¿Cómo puede el usuario recuperar sólo una parte de un documento muy extenso?. Muchos de los documentos deben ser pre-codificados (o pre-marcados) para permitir tal acceso granular? ¿Pueden estar disponibles mecanismos de tiempo de acceso?

El campo del Almacenamiento y Recuperación de Información se enfoca al estudio de todos los aspectos de información, proporcionando fundamentos para la constitución de bibliotecas digitales, así como de las tecnologías clave para la información en red. Varios modelos permiten a los desarrolladores construir sistemas de información y a los usuarios tener un marco conceptual para la búsqueda, revisión y otras operaciones.

1. Enfoque funcional y temático
2. Modelo Booleano

3. Modelo PAT
4. PAT y arboles PAT
5. Algoritmos en arboles PAT
6. Modelo Vectorial Espacial
7. Indización Automática
8. Constelaciones

Las técnicas para actualizar documentos y consecuentemente archivarlos y etiquetarlos deben ser desarrollados tomando en cuenta el hecho de que estos ya son obsoletos o indicando el estado en que se encuentran con una versión autorizada; un ejemplo de esto es que en los estándares ANSI mensualmente se reportan o liberan versiones. Una forma de SGML puede ser apropiada en algunos casos, por ejemplo el formato propuesto por TEI (Text Encoding Initiative).

- **Catalogación:** El acceso a un gran volumen de información es un problema intelectual que ha sido solventado por las bibliotecas tradicionales, lo han hecho con la abstracción e indización de información, así como por la catalogación con reglas y procedimientos para asegurar la consistencia de la misma. Ahora estas herramientas deberán trabajar con información electrónica y se tendrán que realizar correctamente. Deberán adoptar nuevos mecanismos de recuperación, de tal suerte que el usuario pueda igualmente navegar desde la localización de la información hasta recuperarla sin tener que cambiar su modo de uso. Los mecanismos probablemente ligarán el catálogo de registros a los documentos asociados empleando herramientas tales como WWW, el Indicador de Recursos Uniforme (y Localizador), o URI/URL, el recientemente propuesto campo MARC 856. Los SGML pueden ofrecer otras posibilidades para ligar ciertos documentos a través de sus técnicas de descripción. En cualquier caso se deben de tomar en cuenta la representación de los registros de la bibliografía de la localización electrónica física (discos optomagnéticos, discos magnéticos, etc.) y la representación de la localización virtual (red de computadoras).

Si los sistemas de catalogación de las Bibliotecas Electrónicas trabajan bien, los usuarios estarán en la posibilidad de buscar la información, localizar los registros bibliográficos y emplearlos para obtener la información deseada en su lugar de trabajo. Cuando se autentica se tiene que incluir y probar la certificación de ésta.

- **Acceso remoto:** Una Biblioteca Electrónica debe desde el principio tener el propósito de ser accedida desde muchos lugares remotos. El acceso a través de Internet puede ser posible. En las primeras implementaciones piloto que se tienen se puede observar que algunas bibliotecas tienen en la mira la planeación y desarrollo de catálogos y mecanismos de acceso que integran la colección de cada una de las bibliotecas a la Biblioteca Electrónica. Se deben tener procedimientos para la diseminación de dichos registros de catalogación; no sólo serán cuestiones técnicas, si no que también serán cuestiones políticas entre las bibliotecas asociadas con la Biblioteca Electrónica para proveer un acceso no local y tomar en cuenta los patrones locales. Las herramientas bibliográficas ayudaran mucho a esta labor.

- **Derechos y libertad:** Este tipo de elementos frecuentemente representan ciertos problemas, los estándares y técnicas empleadas en este aspecto cubren los siguientes puntos:
 - * **Preservación de la privacidad para los usuarios.**
 - * **Protección a los derechos de los autores intelectuales de la información.**

2.2.8. Derechos de Autor

Las leyes de Derecho de Autor protegen los trabajos creativos. La protección brindada a los autores de trabajos incluyen el derecho a autorizar o realizar: reproducciones del trabajo, trabajos de preparación derivados, distribución al público de los trabajos, y muestra pública de los mismos. Además, los poseedores de los derechos de grabaciones sonoras y programas de cómputo controlan la renta de sus trabajos. Estos derechos no son ilimitados, existen una buena cantidad de limitaciones y excepciones.

En un ambiente electrónico, el sistema Derechos de Autor implica muchas cosas. Los poseedores de los derechos desean controlar el uso de su trabajo y que se pague por cualquier empleo que se le dé; por otro lado el público desea acceso rápido y fácil a un bajo o nulo costo. El mercado está trabajando en esta área. Contratos, lineamientos en el uso electrónico y licencias colectivas están siendo refinados.

Problemas concernientes a la posibilidad de cambiar trabajos sin detección es un asunto complicado. Preguntas concernientes a la integridad del trabajo y del estado de las versiones modificadas bajo la ley de Derechos de Autor tienen que ser resueltas. Hay problemas con las políticas públicas que requieren ser consideradas a través de un diálogo informado.

2.3. TECNICAS DE AUTOMATIZACIÓN Y ADMINISTRACIÓN DE BIBLIOTECAS ELECTRÓNICAS

Hemos entrado a una era de una sociedad basada en el conocimiento, por lo tanto es el momento crucial para las bibliotecas y los centros de información para dar a conocer sus características más importantes, emplear sus recursos e integrar las tecnologías más eficientes.

El espejo electrónico es una herramienta muy poderosa para recuperar información y un eficiente interfaz con el usuario para un sistema de información integral. Propiamente diseñado, un sistema de espejo provee ventajas de costo muy significativo y un ROI (retorno de la inversión) sobre los sistemas basados en manuales y microfilmación. Sin embargo, los costos de la implantación pueden ser considerables tomando en cuenta que los sistemas de reflejo representan un nuevo camino que están empleando los negocios. De otra forma, esto puede ser benéfico para una biblioteca o un centro de información para garantizar su consulta dentro o fuera de sus instalaciones para tener éxito.

Los procesos e implementaciones para automatizar la Biblioteca Digital incluidos son:

- **La Intranet**
- **La re-ingeniería de procesos en los negocios**
- **Las directrices del gobierno federal.**

Las tecnologías enfocadas para automatizar y administrar la Biblioteca Digital permiten establecer una organización adecuada, una arquitectura abierta y aportan estándares flexibles a las bibliotecas, algunas de ellas son:

- Espejo de documentos
- Administración electrónica de imágenes (EIM)
- Reconocimiento óptico de caracteres (OCR)
- Sistemas de administración electrónica de documentos (EDMS)
- Documentos de flujo de trabajo
- Aplicaciones de administración de bibliotecas

Las tecnologías de integración de documentos electrónicos revolucionan la forma en que nosotros administramos los procesos de información. Mucho más que el reemplazar un gabinete de consulta, o aumentar la productividad simple, estas tecnologías permiten procesar, administrar y almacenar documentos electrónicos a través de su ciclo de vida, ofreciendo un nuevo nivel de acceso y control. Correctamente implementadas, las tecnologías de integración de documentos electrónicos pueden reducir los costos de captura, almacenamiento, administración, recuperación, seguridad, distribución y destrucción de información en un 30 a 70 por ciento, a pesar del constante incremento en la demanda de volumen y velocidad de proceso.

2.3.1. Intranet

La atención se ha enfocado en las formas de usar la tecnología WEB para proveer una comunicación interna e incrementar la productividad en un negocio o en este caso en la Biblioteca Electrónica. El término "Intranet" es empleado para describir los sistemas WEB en los que su contenido es dirigido a una audiencia específica en contraste a la Internet, en donde el contenido del World Wide Web (WWW) es dirigido al mundo entero. Un WEB de organización interna, o Intranet tiene un impacto inmediato y substancial en ROI cuando se usa por comunicación interna, o centros de información de cualquier organización moderna. En este momento los negocios, la educación y el gobierno descubren las ventajas de las tecnologías de Internet y WEB para uso interno; una visión de colaboración está comenzando a ser estructurada. Las Intranets representan un nuevo modelo para la administración interna de información, cómputo distribuido y de colaboración, y ofrece una simple pero poderosa implantación del modelo de computación de cliente/servidor.

Los cambios en las organizaciones y la búsqueda de eficiencia ha resultado en una demanda significativa en bibliotecas y centros de información para automatizar y administrar el flujo de información. La forma efectiva de compartir información y la habilidad de intercambiarla entre diferentes grupos de organizaciones, juegan un papel importante en cuanto a lograr un nivel cooperativo importante. Cualquiera que pueda estar haciendo algo para favorecer el compartir información ayuda a la organización a implementar servicios más eficientes.

Necesidades de usuarios y de organizaciones

Una gran parte del nuevo modelo de computación implica simplemente colocar una interfaz tipo navegador con el usuario dentro de una organización o biblioteca, además de algunos servidores Web como soporte, y así se puede crear una Intranet útil y productiva. Esta perspectiva refleja el presente estado del Web: la actual generación de navegadores y la infraestructura de la Internet y Web se han desarrollado para una oleada de recreación. Los usuarios de la organización tienen un conjunto diferente de requerimientos de aplicaciones (por ejemplo, una Intranet es absolutamente un ambiente diferente al de la Internet, de la misma forma esta puede emplear la misma tecnología. Las tecnologías de la Internet pueden ser realmente muy útiles en un conjunto organizacional, pero ellas deben ser integradas en un estilo que observe realidades organizacionales. Hay dos requerimientos primarios de una Intranet: eficiencia individual y administración de información en grupo), accesos, autoridad de colaboración, y distribución; una administración de documentos que cumpla costo-eficiencia; y un costo administrativo.

Los fundamentos de una Intranet

Como los sistemas de información organizacionales, las Intranets demandan más que simples navegador y servidores Web, los cuales son dos estructuras importantes de la Intranet. En lugar de esto los fundamentos de una Intranet caen dentro de una infraestructura de cómputo de una organización. En muchos casos, algunas de las infraestructuras ya están puestas. Muchos de los

componentes restantes pueden ser fácilmente adicionados mientras se avanza en la infraestructura tecnológica, y sin incrementar significativamente complejidad en la red. Las tecnologías de Internet y Web son de igual importancia en la construcción de una Intranet, y el diseño exitoso de ésta requiere un cuidadoso peso de ventajas y desventajas de cada pieza tecnológica.

2.3.2. Reingeniería Organizacional

Nosotros estamos en una era donde los negocios y las profesiones son reingeniería y reestructuración como una consecuencia de un brinco exponencial de la tecnología de la información. Esta reingeniería y reestructuración (llamada BPR o reingeniería de procesos en negocios) dejará un gran número de mejoras en productividad evaluada por los compradores, en optimización de negocios, y competitividad. Ahora, los usuarios finales de la biblioteca esperan que la información sea liberada rápidamente, con la disponibilidad de obtener cualquier información, 24 horas al día.

En este tipo de ambientes, todas las bibliotecas, incluyendo las corporativas especiales, deben estar observando muchos caminos para optimizar sus servicios empleando reingeniería y reestructuración. Los administradores de bibliotecas corporativas, para permanecer en sincronía con este tipo de desempeño, deben de hacer reingeniería en sus bibliotecas y centros de información para entregar información empleando las herramientas electrónicas con mayor relación costo eficiencia y productos disponibles en la industria.

En conclusión, la biblioteca o centro de información debe tratar de existir como un servicio sin límite, un lugar donde la información pueda ser solicitada donde quiera que ésta exista y ser usada inmediatamente por los clientes locales o remotos. Los centros de información de negocios deben hacer el ajuste, excepto en casos especiales, de información en tiempo real la cual será valuada como una herramienta competitiva.

Ajuste a estándares

Las organizaciones están incrementando su dependencia a la transferencia de información impulsada por las tecnologías avanzadas; la administración de la red llegará a ser más importante aún en las organizaciones. Los productos que trabajan con los estándares de la industria son críticos para la administración de las redes; los productos de información que las bibliotecas usan necesitan ser desarrollados para cumplir con tales estándares. Los profesionales de la información deben asegurar las relaciones con proveedores que están desarrollando la tecnología la cual será aplicada a sus servicios. De esta forma, ellos están preparados para aceptar los estándares que requieren para efectuar sus responsabilidades de forma efectiva.

2.3.3. Directrices federales

El gobierno federal de algunos países donde se están instituyendo las Bibliotecas Electrónicas, está tomando medidas muy importantes.

En el documento “El Reto de Internet del gobierno federal de Canadá: El reporte final del Consejo Asesor de Internet” estipula algunas importantes recomendaciones concernientes a la digitalización de bibliotecas:

- Rec. 7.9. El gobierno puede tomar medidas para fortalecer la digitalización de pertenencias digitales con una visión para promover y facilitar el uso del contenido Canadiense en aplicaciones de multimedia intermedia.
- Rec. 11.29 La sociedad interinstitucional puede ser fortalecida con bibliotecas públicas y escolares, con oficinas culturales regionales, etc., para acelerar la producción del acceso universal a información electrónica.
- Rec. 11.32. El Acta de la Biblioteca Nacional de Canadá y el Acta del Archivo Nacional de Canadá pueden ser reformadas para requerir que todos los depositarios de publicaciones basados en texto y registros gubernamentales, normalmente hagan sus depósitos en formato digital.

Las bibliotecas y centros de información en Canadá están actualmente en el proceso de digitalizar sus colecciones. Si las recomendaciones del gobierno federal son adoptadas, la necesidad de digitalizar será cada vez mayor.

2.3.4. Imagen de documentos

“Imagen de documentos” es el nombre genérico para cualquier tecnología que capture la imagen de un documento en papel y almacene ésta en otro medio. El sistema de procesamiento de imagen de documento brinda tecnologías que permiten a los usuarios capturar, almacenar, administrar, recuperar, desplegar, distribuir e imprimir la información que origina un papel, una película, o que es originada en computadora. La administración de documentos es la aplicación que provee la solución que incluye todo lo anterior.

Las imágenes y discos ópticos son herramientas o partes del sistema de administración de imagen electrónica. Generalmente la automatización de Bibliotecas Digitales para imágenes de documentos sigue tres procedimientos, dependiendo de la complejidad del flujo de papel:

- Administración de imagen electrónica
- Administración de documentos
- Flujo de trabajo de documentos

Dependiendo del proveedor y el producto, los tres componentes pueden traslaparse. Este es el punto donde surge la confusión en las tecnologías de documentos electrónicos. Además, el flujo de

trabajo es una tecnología poderosa que no siempre es apropiada para ciertas organizaciones; en lugar de usar esto, el trabajo en grupo puede ser apropiadamente empleado para la distribución serial de documentos.

2.3.5. Sistemas de Administración de Imágenes Electrónicas (EIM)

Los Sistemas de Administración de Imagen Electrónica (EIM) permiten almacenar, compartir, comunicar y administrar la información en formas que no son posibles con un sistema basado en papel. Con un sistema EIM, se puede emplear accesos inmediatos y más flexibles a documentos, rápido procesamiento, y proveer servicios a clientes o acceso a usuarios.

EIM en la biblioteca

Los procedimientos de las bibliotecas son diferentes a otras organizaciones. El poder de la EIM radica en la habilidad para cumplir con la única necesidad de las personas administradoras de las bibliotecas, que es el de integrar los nuevos procedimientos con sus sistemas de información existentes. Con el crecimiento en la organización, sus aplicaciones de procesamiento de información vienen a ser más lentas por el volumen y complejidad del procesamiento de papel. Cuando se planea un sistema EIM, se debe de tomar en cuenta los actuales métodos de procesamiento y determinar qué elementos pueden ser requeridos para hacer esto más rápido y eficiente.

Integrando imágenes en una aplicación existente

Si se desarrolla un sistema EIM sin integrar el equipo o programas de computadora existentes, se puede emplear con facilidad un diseño considerable con pocos contrastes operativos y de plataforma. Sin embargo, empleando un nuevo sistema independiente de las plataformas existentes de información, se corre el riesgo de crear una información aislada de otras aplicaciones y esto es pocas veces apropiado. Una EIM hace su más poderosa contribución en la biblioteca o centro de información cuando se integra con otras aplicaciones.

2.3.6. Reconocimiento Óptico de Caracteres

El Reconocimiento Óptico de Caracteres (OCR) y la EIM son dos tecnologías que algunas veces son confusas.

En un sistema EIM, un documento es capturado como una imagen (como tomar una fotografía de una página). La imagen de una página es tratada como una unidad de información única; el texto y gráficos de una imagen son fijos, y no pueden ser editados con un procesador de palabras o editor de gráficos.

La tecnología OCR brinda la posibilidad de interpretar los caracteres en una imagen de página reconocidos por la computadora y ser convertidos a un archivo de texto que se puede editar. El OCR puede también indizar documentos automáticamente “leyendo” la imagen para una información específica ya que se tiene información en texto. Sin embargo, la tecnología ha sido estimada para tener un 97 % de exactitud. De esta forma, la prueba manual de documentos es aún requerida, es decir, verificar que la información en texto sea la que está en la imagen original.

2.3.7. Sistemas de Administración de Documentos Electrónicos (EDMS)

Un Sistema de Administración de Documentos Electrónicos (EDMS) indiza y recupera documentos que se crean en su forma original, como son los archivos de computadora. Un EDMS usualmente no incluye proceso de imagen (EIM), flujo de trabajo, fax electrónico o reconocimiento de caracteres; en lugar de esto puede integrarse con otros sistemas que proveen estas funciones. La decisión de integrar estas tecnologías depende de la organización y necesidades de los usuarios de la biblioteca en cuestión.

Se debe considerar la integración EIM con un EDMS si se tienen entradas de documentos los cuales deben de ser archivados y recuperados, así como también promover el uso del mismo sistema empleado en los documentos electrónicos generados internamente. El reconocimiento de caracteres puede ser empleado para convertir estos documentos digitalizados electrónicamente a texto, permitiendo búsquedas en texto completo del EDMS.

2.3.8. Flujo de trabajo de documentos

La integración de flujo de trabajo con un EDMS es apropiado si los documentos son revisados por algunas personas. Se puede entonces desarrollar reglas acerca de cómo hacer las rutas basadas en las respuestas de los recipientes de cada conjunto de flujo de trabajo. Para simples aplicaciones de definición de rutas, el sistema de correo electrónico existente puede proveer esta capacidad.

Con la tecnología de flujo de trabajo de documentos en la biblioteca, el proceso en papel debe ser primero claramente predeterminado y bien definido. Cada documento seguirá una serie de procesos fijos y pasos aprobados, y cada paso debe ser completado antes de que el documento se mueva. A través del análisis del flujo de trabajo, el jugar un papel indefinido y tareas imprecisas se eliminan, y métodos más eficientes de trabajo completo son identificados e implementados.

Los flujos de trabajo permiten digitalizar documentos para moverse electrónicamente a través de procesos de negocios y en cada área de la biblioteca. La digitalización, cuando se combina con el flujo de trabajo llega a ser una tecnología de procesamiento, en lugar de capturar documentos, almacenarlos y recuperarlos.

Algunos sistemas EIM permiten a un flujo de trabajo adecuado que automatice el flujo del documentos en fax y papel de cada día. Esto es útil si se está manejando una variedad de documentos que requieren diferentes tipos de procesamiento. Con un flujo de trabajo adecuado, se puede determinar qué documentos necesitan ser vistos por quién y en qué orden. El sistema puede también permitir seguir la pista de la marcha del documento.

2.3.9. Aplicaciones de administración de bibliotecas y la Intranet

Hay cientos de paquetes de cómputo de administración de bibliotecas. Una tendencia muy común es que los profesionales de la información y de las bibliotecas están enterados del OCLC, la red de información de bibliotecas más grande del mundo. El OCLC es la organización que dio origen a las “Dimensiones de Información”, la cual tiene productos de computadora en más de 3,000 lugares en el mundo.

Software SiteSearch (TM) de OCLC

La familia SiteSearch OCLC de productos que ayudan a hacer recursos de información locales, trabajan en un ambiente de red muy grande. Los sistemas Servidores Z39.50 de SiteSearch dan soporte para construir recursos de información local, incluyendo texto completo, imágenes, vídeo y sonido, y habilitan el acceso en red de estos recursos a través de Z39.50.

En adición, los servidores WebZ(TM) de SiteSearch ofrecen acceso WWW a recursos de Z39.50 remoto y local, así como también herramientas para construir interfaces basadas en HTML. Con estas bases, las universidades y organizaciones con red muy grande pueden ligar servicios de información local, servicios OCLC como el servicio FirstSearch (TM) y otros recursos Web o Z39.50 dentro de una costumbre que se difunde de Biblioteca Electrónica.

El OCLC tiene algunos productos de soporte a digitalización que permiten construir bases de datos de imágenes localmente desde diferentes tipos de materiales, incluyendo documentos impresos y fotografías. Los productos proveen herramientas para crear colecciones de imágenes electrónicas más complejas para aplicaciones tales como colecciones de fotografías y reservas.

El estándar Z39.50

La compatibilidad es importante para la Biblioteca Digital y Z39.50 es el estándar que emerge para automatizar los sistemas de bibliotecas. El estándar Z39.50 soporta un ambiente de cómputo cliente-servidor en el cual una interfaz única de usuario (cliente) puede referenciar información desde múltiples fuentes (servidores). Este sistema es ventajoso para las bibliotecas y centros de información que necesitan una interfaz de usuario común para buscar sus catálogos, referencias a bases de datos localmente, y bases de datos comerciales populares localizadas en servidores remotos. Tanto el OCLC como las “Dimensiones de Información” se apegan al estándar Z39.50.

Soluciones Intranet de Dimensiones de Información (IDI)

A diferencia de muchos proveedores de programas de administración de bibliotecas que ofrecen un sistema de biblioteca integrado, el enfoque de la “Dimensión de Información” es como el de una compañía de aplicaciones de cómputo para la administración de texto y documentos. Los

proveedores de sistemas de biblioteca tradicional están comenzando a desarrollar soluciones de documentos electrónicos. Los productos de “Dimensión de Información” demuestran como la información electrónica se extiende más allá de una biblioteca y es empleada para las operaciones dentro de una organización.

TECHLIBplus(TM) es un sistema integrado que facilita las funciones de administración de una biblioteca, tal como la catalogación, circulación, control serial, y adquisiciones. Cuando se da el acuerdo sobre el Gateway WEBServer BASIS, IDI provee un único recurso de conocimiento organizacional integrado. IDI se especializa como un centro cooperativo de recursos de información de biblioteca, tiene 400 instalaciones (incluyendo TECHLIB) en América del Norte, 40 de los cuales están en Canadá.

Las Dimensiones de Información proveen amplias soluciones de administración de documentos para la Intranet (cuando se emplea ésta para la administración de la Biblioteca Electrónica). El Administrador de Documentos BASIS(TM), el software base para TECHLIB, es empleado para organizar los documentos internos y compra de recursos electrónicos en el centro de información; además permite administrar y proveer acceso bibliográfico para materiales publicados. El sistema BASIS está compuesto por robustos servicios (incluyendo servicios de biblioteca, recuperación de texto completo, liberación, seguridad y autenticación). Un administrador de almacenamiento de documentos habilita una base de datos entidad relación de documentos; y cuando es necesario optimizar las aplicaciones. La Red de Información de Herencia Canadiense (CHIN), un programa del Departamento federal de la Herencia Canadiense, emplea el software administrador de documentos BASIS de Dimensiones de Información.

Netscape Communications y Dimensión de Información

La Netscape Communications Corp. fue fundada por James H. Clark, fundador de Silicon Graphics, y Marck Andreesen, creador del prototipo de investigación de NCSA Mosaic para Internet. En septiembre 27 del 95, Las Dimensiones de Información y Netscape Communications anuncian una asociación para liberar las soluciones para la administración de documentos. La integración del Administrador de Documentos BASIS con el Servidor de Comunicaciones Netscape(TM) y el Servidor Comercial(TM) provee un acceso amplio a las empresas para administrar las colecciones de documentos usando una Intranet privada, de tal modo que se está habilitando a las organizaciones para transformar zonas departamentales de información dentro de una base amplia de conocimiento potencial.

2.3.10. La gente: la clave para la implantación exitosa de la tecnología

La EIM, EDMS y las tecnologías relacionadas, especialmente el flujo de trabajo, proveen una ganancia en la productividad. Ellos también requieren trabajo de rediseño (o BPR), nuevas formas de trabajo y nuevas formas de pensar acerca del trabajo. También frecuentemente, el elemento humano es supervisado. El error que comenten los responsables en el diseño de un sistema EIM, es la negligencia en los requerimientos dados, la indiferencia en los temores de los responsables, y otros factores humanos son las contribuciones que se dan para que se presenten fallas en un sistema EIM.

Una persona que ha colaborado exitosamente en un EIM en el pasado debe ser incluido en el equipo actual. En el equipo que trabaja en el desarrollo de EIM deben ser considerados los usuarios de la biblioteca. Sus conocimientos aseguran un sólido entendimiento de los procedimientos existentes y la integración de las aplicaciones finales. Es importante trabajar con los usuarios que se resisten al cambio; se debe asegurar que el progreso se dé de forma tranquila pero exitoso en los proyectos EIM.

2.3.11. Costo y justificación de las tecnologías

Con un EDMS, el ROI (recuperación de la inversión) puede ser estimado en un 10 % de decremento en el tiempo invertido en distribución y recuperación de documentos. Para usuarios con un salario de \$40,000, esto se transforma a \$4,000 por usuario. Para un sistema de 100 usuarios, esto se traduce en salvar \$400,000 por año. Este ROI es conservador por que:

- la productividad es incrementada en mas de 10 %,
- no incluye beneficios como el incremento de empleos, clientes y satisfacción de clientes.

El ROI puede ser reducido desde el costo del EDMS (incluyendo el tiempo invertido en selección del sistema, integración y entrenamiento del mismo) para estimar el período de reembolso.

2.3.12. Conclusión dentro de la automatización de bibliotecas digitales implementadas

“El Web ya es una parte integrada a la infraestructura de la computación en muchas empresas, haciendo de esto una plataforma ideal para las aplicaciones donde se concentren los documentos,” dice Bill Forquer, presidente y oficial ejecutivo de Dimensiones de Informaciones. “La administración de Bibliotecas Electrónicas es una aplicación ideal de la Intranet”. Con la integración de administración de documentos con las características necesarias para la administración diaria de una Biblioteca Digital, una única plataforma es creada para soportar las necesidades del manejo de información total de una organización.

Los límites de la información que puede ser proporcionada por una organización cuando se emplean las tecnologías apropiadas pueden no existir. Los servicios de información pueden ser soportados desde un lugar central, y de esta forma se logra que una organización tenga la oportunidad de moverse a un sistema de información centralizada cuando se adquiera información de todos tipos, producidos interna y externamente.

2.4. ALCANCES Y LIMITACIONES ACTUALES DE LA BIBLIOTECA ELECTRÓNICA EN EL TRATAMIENTO DE LA INFORMACIÓN

Los subsidios de cómputo en la biblioteca tradicional (catálogos, bases de datos, etc.) han favorecido la investigación o búsqueda de información con una lista de referencias bibliográficas. La biblioteca debe entonces tomar esta lista para proporcionar el acceso al material físico o para tener los préstamos entre bibliotecas. Por supuesto estas dos opciones pueden traer consigo tardanzas, esperando que otro usuario regrese un libro o a que otra biblioteca lo libere. Este modelo, en el cual los datos electrónicos son solo referencias bibliográficas al material físico, está ahora cambiando con la disponibilidad de recursos de texto completo almacenados electrónicamente. El decremento en el precio del poder de cómputo significa que el costo efectivo para almacenar texto completo electrónicamente, provee nuevos beneficios y desafíos para el usuario. Las ventajas del texto completo se refieren inicialmente al acceso, aunque el acceso universal e inmediato al material puede traer consigo tardanzas. Pero también el mejor beneficio es el acceso al texto por sí mismo: la habilidad para buscar un cuerpo electrónico completo en forma instantánea; la oportunidad para el análisis textual automático y la facilidad para que los documentos electrónicos puedan ser ligados directamente a otros recursos relevantes.

Los beneficios del texto completo electrónico es tal, que estos recursos están proliferando exponencialmente a través del mundo de la biblioteca.

2.4.1. Vías para crear recursos de texto completo

Los beneficios para los investigadores conducen a dificultades para los bibliotecarios, o al menos dificulta las decisiones. Existen diferentes vías que proporcionan recursos de texto completo, pero no son muy claras dentro de un escenario donde las técnicas empleadas son las mejores.

Archivos de texto

El método más simple para proporcionar recursos de texto completo electrónico es simplemente poner disponible un archivo de texto. En el caso en que el documento original fue creado electrónicamente, este es un proceso trivial y de bajo costo para ponerlo disponible, tanto para obtenerlo como para buscar en él. Este método también tiene la ventaja que el usuario está observando el texto original que fue escrito por el autor sin ningún proceso de conversión. En este sentido, aunque existan algunos elementos de presentación que se tienen que ajustar (como la pregunta de que si se está viendo el documento con el estilo de letra correcto), la integridad del texto actual está asegurado. Las dificultades con la digitalización comienzan cuando se pone disponible cierto material que no fue creado electrónicamente. El proyecto ILEJ, por ejemplo hace frente a esta dificultad en su enfoque en los textos de los siglos XVIII y XIX. Ante esta situación, la solución más directa es capturando la información dentro de un archivo de computadora. Esto, sin embargo, puede ser un proceso caro, y tal vez, más importante, es estar expuestos a tener

errores en la captura de datos: al usuario que consulte la información no se le garantiza la integridad del texto (tal vez lo que se está consultando no es lo que el autor escribió).

Es posible incrementar la exactitud de la captura de datos si esta labor la llevan a cabo varias personas y se comparan los resultados electrónicamente, corrigiendo entre los diferentes textos. Aunque este método es eficiente no garantiza la perfección de la captura del texto original. El capturar el documento varias veces es un proceso caro que está abierto a errores, sin embargo mucha información de material viejo ya está disponible para la comunidad usuaria, en CD ROM o vía Internet empleando esta técnica.

2.4.2. Imágenes Digitales

Un beneficio que trae consigo el descenso del precio del poder de cómputo, es que el proceso de crear imágenes de computadora tiene ahora un costo efectivo y puede proveer un método alternativo de entrega de recursos de texto completo. En lugar de capturar el documento en un texto de computadora, la página del documento original es digitalizada y al usuario se le presenta una imagen de computadora mejor que un archivo de texto. La ventaja es que es mucho más rápido y económico que tener que capturar cada página del documento. También la integridad del texto es asegurada por que el usuario ve una imagen de la página original. El problema fundamental con este método es que casi siempre trae consigo una desmembración y aportación de adornos a los documentos originales los cuales son colocados en un digitalizador.

Existe un cierto margen de que estos documentos originales puedan ser recuperados, pero esto adiciona un costo significativo a la operación. En muchos casos el volumen original es simplemente descartado con las leyes electrónicas cuando se convierten en recursos primarios. El descartar los volúmenes originales trae muchos problemas costosos. El original es extraviado, así que el valor del original sobre y por debajo de la versión electrónica debe ser contabilizada y tomada en cuenta. Por mucho que se eleve la calidad de una digitalización, la información disponible en el original se perderá. No sólo hay una pérdida inevitable de información, existe también el costo de la versión electrónica primaria del documento. El factor importante que determina que la versión electrónica es ahora el recurso primario, es que se incrementa el costo al hacer una copia electrónica, ya que ésta debe ser de alta calidad para que puede satisfacer el potencial de los usos futuros del material. Este requerimiento incrementa el costo de almacenamiento, ya que las imágenes de alta calidad implican archivos de gran tamaño, y esto implica discos o cintas de gran capacidad, o más CD-ROM en donde almacenarlos. Los costos de mantenimiento de archivos electrónicos para grandes periodos en el futuro serán considerables. Simplemente respaldar la información en cintas magnéticas y almacenarlo en un medio no será suficiente. Surgen ante esto algunas preguntas, ¿se tendrá el correcto formato para la cinta disponible en los próximos 10 años?, ¿se tendrán los programas de computadora correctos para leer las imágenes?, ¿se tendrá el correcto sistema operativo o equipo de cómputo para ejecutar estos programas?, ¿es la integridad del medio de almacenamiento confiable para cierto periodo de tiempo, o se necesitará actualizar los datos en otro tipo de medio después de algunos años?

El costo del mantenimiento de los datos electrónicos es considerable y debe ser tomado en cuenta cuando se considere la destrucción del original. En proyectos grandes de digitalización, este costo puede ser compensado con los recursos liberados por la destrucción de los originales. Con todo esto se puede decir que destruir los originales y almacenar los datos en forma electrónica es todavía

un método viable de creación de recursos de texto completo, particularmente donde el original ya está deteriorado. Una variación de este método es el almacenar la copia primaria (del original) en microfilmación en lugar de hacerlo electrónicamente. Esto tiene el beneficio de que no se depende de los cambios dinámicos de la tecnología de la computación; una microfilmación puede simplemente ser almacenado en una gaveta y recuperarla en 10 años, el equipo para leer (o digitalizar) microfilmación aún permanecerá. Esta aproximación híbrida de proporcionar una copia electrónica para acceso y una copia de microfilmación para archivos elimina gastos y preocupaciones del mantenimiento de imágenes digitales.

Indizando Imágenes Digitales

Asumiendo que todas estas dificultades han sido exitosamente superadas, existen aún problemas asociados con un recurso de texto completo de imágenes digitalizadas. Primeramente, los archivos de la imagen son significativamente más grandes en tamaño que archivos de texto que han sido digitalizados. Esto no solo hace incrementar el costo de almacenamiento, también incrementa el tiempo empleado para la recuperación de información, el cual puede ser significativo en una red o en Internet. En este sentido, un lector emplea grandes períodos esperando a que sea liberada la información y puede provocarle molestias. En segundo lugar, tener imágenes de páginas es muy bueno, pero nosotros no hemos perdido el segundo de los beneficios de acceso con el cual inicialmente se dio la urgencia para digitalizar: la habilidad para la búsqueda en texto completo y análisis electrónico, etc. En algunas situaciones, la disponibilidad del original es para restringir que el acceso sea provechoso por el mismo. Esto es particularmente verdad con materiales muy valiosos, tales como los manuscritos medievales, pero puede también ser útil si el material es menos valioso pero es físicamente remoto.

Pero generalmente, una colección de páginas que ha sido simplemente digitalizada es de poca utilidad para un usuario; se necesita un valor adicional proporcionado por índices electrónicos y texto completo para convertirse en imágenes útiles. Estos índices pueden simplemente ser copias hechas a mano de otros originales y contenido de las páginas, pero esto sólo cumple el primero de nuestros requerimientos de acceso: disponibilidad inmediata. En términos de la disponibilidad del texto para futuras búsquedas y análisis, este método no se ofrece sobre los volúmenes originales. Sin embargo, en algunos casos, tales índices ya han sido creados por terceras partes y pueden estar disponibles.

2.4.3. OCR

El método de obtener la búsqueda en texto completo desde imágenes de páginas digitalizadas comienza por procesar las imágenes digitalizadas con un programa de Reconocimiento Óptico de Caracteres (con sus siglas OCR), el cual creará versiones de archivos de texto de las imágenes, en las cuales pueden entonces buscar información. Este es un proceso automático y el programa OCR es relativamente económico, por lo que el costo de convertir grandes cantidades de datos es muy bajo.

Sin embargo, al emplear OCR el texto convertido nunca será 100 % exacto, y esta exactitud desciende considerablemente con textos más antiguos. Una solución es una prueba de lectura y así

corregir el texto OCR. Sin embargo, este paso final agrega una enorme carga de trabajo para que sea proceso totalmente automático. Esto no es importante considerando lo poco que se aprecia el tiempo que se consume en la prueba de lectura y la edición interactiva. A menos que el texto sea de gran exactitud, éste puede frecuentemente ser mas económico para revertir a la estrategia inicial de copias mecanografiadas del original; la edición interactiva es mucho más lento que una copia mecanografiada.

La conversión a texto resulta mucho más cara y consumidora de tiempo que la captura de la imagen solamente. La experiencia con el proceso de reconocimiento óptico de caracteres (optical character recognition OCR) está relacionada y es comparada con la recaptura del texto. ¿Qué factores afectan la exactitud del OCR?, ¿cómo debe ser producido el texto completo a partir de un OCR para que sea útil?, ¿cómo reaccionan los usuarios a textos imperfectos?. Estos son cuestiones que deberán ser exploradas.

Exactitud de OCR

Excluyendo la “pruebas de lectura y corrección de OCR” en algún proyecto, el OCR puede ser dejado con errores, sin embargo esto implica proyectos de “alto volumen, bajo costo”. Esto significa que los archivos de texto derivados de OCR no pueden ser suministrados como datos primarios, pero pueden ser usados para dar valor adicional para las imágenes de páginas, permitiendo por ejemplo búsquedas. La Biblioteca Electrónica, puede consistir de pares ligados de imagen/archivo de texto, en los cuales una búsqueda puede ser hecha de los archivos de texto, pero la imagen de la página original es proporcionada para que se consulte después de que se ha tenido el resultado de la búsqueda. La ventaja de almacenar pares de texto/imagen es que, aunque el texto resultado del OCR pueda no ser correcto, el usuario puede consultar la imagen de la página original y así íntegramente del original es garantizada la información. Este modelo de hecho es empleado por sistemas de administración de documentos de oficinas comerciales en el cual las cartas y otros documentos son digitalizados en su llegada y sus imágenes recuperadas después por una búsqueda a través de texto de OCR.

El por qué no hubo la intención de corregir el OCR desde el principio, es por el hecho de que la búsqueda exhaustiva del material no es posible; siempre hubo términos en los que se les aplicó incorrectamente el OCR y que una búsqueda no puede encontrar. Esto es un defecto que dan a los recursos de texto completo de muchos periódicos (que han ofrecido su información electrónicamente como el Early Journals de Inglaterra) un sesgo diferente a recursos comparables (otras asociaciones que prestan servicios de información en Internet). La utilidad de este tipo de material digitalizado requiere un cuerpo grande de información para que se reduzca el porcentaje de error en las búsquedas, y dentro de restricciones de esfuerzos fijos este criterio ha causado una reducción asociada en la recuperación. Se piensa que la posición en la matriz de costo/recuperación es apropiada para periódicos como el Early Journal por ejemplo.

2.4.4. Uso del CD-ROM y DVD para almacenar información digital

A finales de los 80's, los editores de los documentos de George Washington, John Adams, Thomas Jefferson, James Madison, y Benjamin Franklin fueron contactados por la Fundación Packard

Humanities con el propósito de producir una edición en CD-ROM de los documentos completos de cada uno de los personajes mencionados. Esta edición electrónica es un complemento ideal para los volúmenes ya publicados, logrando que todos los documentos se encuentren disponibles para maestros y estudiantes a un costo razonable. En los próximos años CD-ROM's similares serán desarrollados y puesto a disposición. Lo importante en este proyecto no sólo es la experiencia de producir un CD-ROM, sino el impacto que la tecnología ha tenido en los proyectos editoriales más grandes. Actualmente ya se están editando los volúmenes con la mira puesta en que el material estará pronto disponible en CD-ROM. La edición electrónica completa proporcionará inmensas posibilidades en la búsqueda de documentos de información que antes parecía imposible. El tipo de innovaciones técnicas que están actualmente disponibles y que están en desarrollo, revolucionarán la investigación histórica y la producción de documentos históricos. Desafortunadamente, mucha de esta nueva tecnología no está siendo empleada en las etapas de planeación de los proyectos históricos, debido simplemente a que muchos historiadores están preocupados en la corta etapa de su existencia. Al menos dos grandes proyectos históricos están considerando la edición de microfilmes, simplemente porque no están preocupados de las posibilidades de alternativas electrónicas y las ventajas de las nuevas tecnologías en términos de flexibilidad y potencial de investigación comparadas contra el microfilm. En efecto, mucha de nuestra historia y literatura están aún en la etapa de luchar con PC's. Existen actualmente diversos proyectos de edición histórica en progreso, y un número similar de proyectos de literatura. Mientras los dos campos tienen diferentes enfoques para la edición de texto, hay formas en que la tecnología puede ser de utilidad para ambos.

Debido a que la mayoría de los editores involucrados en la edición de CD's son técnicos expertos en cualquier sentido, es necesario explorar cuántas de las innovaciones electrónicas pueden ser empleadas con éxito por los eruditos que no tienen experiencia en el mundo de la nueva tecnología. Una de las mayores preocupaciones de los patrocinadores de multitud de ediciones avanzadas es la limitada audiencia alcanzada por los volúmenes editados. La mayoría de estas ediciones están siendo publicadas en pequeñas cantidades y el precio que el editor pone en el producto lo coloca fuera del alcance de individuos y sólo a disposición a través de bibliotecas públicas y de grandes instituciones educativas. Sin embargo, muy poca atención está siendo dada a las formas en las que la tecnología puede evitar la publicación convencional para lograr que los documentos literarios e históricos estén ampliamente a disposición del público.

Lo más atractivo de la edición en CD-ROM de los documentos de George Washington fue el hecho que se colocaron 135,000 documentos coleccionados en un formato barato que puede ser colocado en bibliotecas públicas, pequeños colegios, y escuelas de nivel superior. Esto proporcionará una audiencia mucho mayor a la actual. Debido a que el CD no contiene ninguna nota explicativa que aparece en los volúmenes impresos, se espera que el empleo de este no afectará el hecho que los investigadores busquen los volúmenes impresos.

Además de la profunda ignorancia de los avances técnicos, es frecuente encontrar que muchos editores, historiadores y eruditos literarios se resisten y aun muestran hostiles ante la sugerencia de que la tecnología electrónica puede ayudar en su trabajo. Al discutir los argumentos tradicionales que se esgrimen para resistir al uso de la tecnología, estos varían desde la desconfianza por la rapidez en los cambios hasta la suspicacia por el lenguaje técnico empleado para describir los desarrollos electrónicos.

También se tiene la tendencia de emplear la nueva tecnología, DVD (Digital Versatil Disk); dentro de las características más importantes del DVD están su capacidad, interoperabilidad y compatibilidad. Los primeros DVD tienen una capacidad de 4.7GB, lo equivalente a 7 CD-ROM.

Tales capacidades serán de gran ayuda para las aplicaciones tales como, bases de datos, programas de computadora, etc. También cambian el contexto de los desarrolladores de sistemas: las aplicaciones pueden ofrecer video hasta de dos horas, además de pistas de datos adicionales y audio. Algo importante es que se están desarrollando controladores DVD-CDROM compatibles con CD-ROM, lo cual permitirá emplear los DVD sin perder la inversión realizada en sistemas y bases de datos en CD-ROM.

Con todas estas ventajas, el empleo de DVD en las Bibliotecas Electrónicas es una tendencia actual que permitirá seguir con éxito su camino.

2.4.5. Problemas en el uso de imágenes electrónicas en la preservación de documentos

En el uso de las imágenes electrónicas para la preservación documental, es necesario considerar diversos problemas, como el asegurar la calidad de las imágenes, mantenimiento sustancial de tasas de conversión, proporcionar identificación única para el acceso automático y la recuperación, y el acomodo de volúmenes grandes y material frágil.

Para mantener la alta calidad de las imágenes, las funciones de proceso deben corregir las deficiencias en la digitalización de las imágenes. Algunos sistemas comercialmente disponibles incluyen estas funciones, mientras que otros no lo hacen. La digitalización de imágenes debe ser procesada para corregir las deficiencias de contraste - ya sea por la pobreza en éste, resultado de impresiones muy tenues, y/o por fondos muy oscuros y, por la variación del contraste resultado de la decoloración. Más aún, la densidad del digitalizador debe ser adecuada para permitir la legibilidad de la impresión y la suficiente fidelidad en el material a medios tonos de gris. Los bordes o efectos de los contornos de las páginas deben ser removidos para asegurar su estética. Páginas oblicuas deben ser corregidas por razones estéticas y para asegurar la exactitud en el reconocimiento de caracteres. Las imágenes compuestas de texto en dos tonos e ilustraciones en tonos de gris deben ser procesadas apropiadamente para mantener la calidad de cada una de ellas.

2.4.6. El papel que juega la red de datos

Bases de Datos Multimedia en Red

¿Qué es necesario considerar al construir y distribuir una base de datos de material visual en un ambiente multiusuario?. En el pasado no era factible implementar bases de datos de material visual en ambiente de usuario compartido debido a las barreras tecnológicas. Cada uno de los dos modelos básicos de bases de datos multimedia han mostrado sus propios problemas. El modelo de almacenamiento multimedia analógico ha necesitado una infraestructura increíblemente compleja y cara. La economía de escala que hace las configuraciones multiusuario más baratas por usuario atendido no aplica a un ambiente que requiere de una estación de trabajo, reproductor de videodiscos, y dos dispositivos de despliegue para cada usuario.

El modelo de almacenamiento multimedia digital ha necesitado de grandes cantidades de espacio de almacenamiento (algo así como un gigabyte por cada treinta imágenes). En el pasado, el costo de tales cantidades de espacio en disco provocaron que el modelo fuera prohibitivo. Pero la caída en los costos de almacenaje finalmente ha hecho esta alternativa viable.

Si el almacenamiento ya no es un impedimento, ¿qué se necesita considerar al construir una base de datos multiusuario digital de materiales visuales?

El problema clave es el tamaño de los documentos multimedia, y cómo esto afecta no sólo al almacenamiento sino al tiempo de transmisión por la red. Cualquier medio de transmisión por debajo de T-1 (donde T es el tiempo) no es práctico para archivos de 1 Mbyte o mayores (que es el tamaño de un documento multimedia pequeño). De esta manera, aun en una línea de 56 Kbytes puede tomar hasta tres minutos el transferir un archivo de 1 Mbyte. Y este escenario toma condiciones ideales, y no toma en consideración a otros usuarios compitiendo por ancho de banda, tiempo de acceso al disco o tiempo necesario para el despliegue remoto. Las tasas de transferencia telefónica pueden ser completamente no prácticas.

Esto necesita compresión, que por sí mismo soluciona gran número de problemas. Para lograr reducir los tamaños de archivos significativamente debemos emplear algoritmos de compresión eficientes. Pero ¿qué tanta calidad estamos dispuestos a perder?. Para determinar esto sólo se ha hecho un estudio significativo de las necesidades de calidad de imagen para un grupo de usuarios en particular, y el estudio no revela una pérdida excesiva por resultado de la compresión. Sólo después de identificar las necesidades de calidad de imagen, se pueden determinar las necesidades de almacenamiento y de ancho de banda.

Experiencias en aplicaciones basadas en ambiente X-Window (como Imagequery, una base de datos de imágenes de la Universidad de California en Berkeley) demostraron la utilidad de una topología cliente - servidor, pero también hacen evidentes las limitaciones de las aplicaciones actuales en un ambiente distribuido.

Es indispensable examinar los efectos en el tráfico al desplazar documentos multimedia en la red. Se necesitan considerar diversos esquemas que ayuden a evitar los cuellos de botella entre los servidores y los puentes. Sin embargo las experiencias con este tipo de aplicaciones aun presentan muchas preguntas. ¿Qué tan cerca debe ser almacenado un documento multimedia de la aplicación empleada para visualizarlo?. ¿Puede ser accedido y visto por otras aplicaciones?. Las experiencias con el formato MARC (y más recientemente con Z39.50) muestran qué tan útil puede resultar el almacenar documentos de forma tal que puedan ser accedidos por una amplia gama de aplicaciones.

Finalmente, desde un punto de vista de acceso intelectual, es necesario resolver problemas en cuanto a dichos documentos multimedia en ambientes interdisciplinarios. Es necesario revisar la terminología y estrategias de indexación que permitan proporcionar acceso a este material en una forma interdisciplinaria.

Direcciones en Redes de Alto Desempeño para Bibliotecas

El paso al que la tecnología de cómputo ha avanzado en los últimos cuarenta años no muestra signos de abatimiento. Estrictamente hablando, cada periodo de cinco años ha traído consigo un

mejoramiento en orden de magnitud en precio y desempeño de los equipos de cómputo. No hay obstáculos fundamentales que eviten que este paso continúe por al menos la siguiente década. Tan sólo en los pasados cinco años, el cómputo ha estado omnipresente en las bibliotecas, afectando a todo el personal y sus patrones, directamente o indirectamente.

Durante esos mismos cinco años, las tasas de transferencia en Internet, han crecido enormemente. El crecimiento en el número de usuarios y el volumen en el tráfico en la red ha continuado creciendo geométricamente, a tasas aproximadas de 15% mensuales. Este flujo de capacidad y uso, crea inmensas oportunidades y retos para las bibliotecas. Las Bibliotecas se deben anticipar a las implicaciones futuras de esta tecnología, participando en su desarrollo, adoptándolas para asegurar el acceso a las fuentes mundiales de información.

La infraestructura para la era de la información está puesta. Las Bibliotecas enfrentan decisiones estratégicas sobre el papel en el desarrollo, adopción y uso de la infraestructura. La infraestructura que está emergiendo es mucho más que computadoras y líneas de comunicación. Es más que la posibilidad de realizar procesos en un sitio remoto, enviar correo electrónico a una máquina del otro lado del país, o transferir un archivo de una biblioteca a otra. Durante los próximos años seremos testigos de un desarrollo substancial de la infraestructura de información en la red.

Para proporcionar un adecuado liderazgo, los profesionales de las bibliotecas deben alcanzar un entendimiento fundamental de la apreciación y enfoque de las redes de cómputo desde las locales hasta las más grandes a nivel mundial.

2.4.7. Biblioteca Electrónica. Visiones y Realidades

La Biblioteca Electrónica ha sido una visión deseada por unos y rechazada por otros desde que Vannevar Bush acuñó el término memex para describir un sistema de información personal, automatizado e inteligente. Variaciones de esta visión incluyen a Xanadau de Ted Nelson, Dynabook de Alan Kay y "biblioteca sin papel" de Lancaster, con la más reciente encarnación del "Navegador del Conocimiento" descrito por John Scully de Apple. Pero la realidad sobre los servicios bibliotecarios ha sido menos visionaria y el salto hacia la Biblioteca Electrónica ha eludido a universidades, editoriales y la tecnología de la información.

El Instituto de Investigaciones Memex, una organización independiente no lucrativa de desarrollo e investigación ha creado un programa de Biblioteca Electrónica de desarrollo e investigación compartidos para lograr una visión colectiva más concreta. El programa trabaja en la creación de grandes colecciones de imágenes electrónicas indizadas y disponibles al público de documentos publicados en bibliotecas públicas, especiales y académicas. El plan estratégico es el resultado del primer nivel del programa, que ha consistido de investigación en las tecnologías de información disponibles para apoyar tal esfuerzo, los parámetros económicos de los servicios electrónicos comparados con las operaciones de las bibliotecas tradicionales, y los factores políticos y financieros que afectan la transición de distribuciones impresas a accesos electrónicos en red.

El plan estratégico contempla una combinación de accesos a bases de datos abiertas al público para búsquedas, colecciones de documentos en imagen y texto almacenados en servidores de archivos con acceso a red local y remoto, y un sistema de control de propiedad intelectual. Esta combinación de tecnología e información es definida en el plan como la colección de la Biblioteca Electrónica.

Algunos patrocinadores están desarrollando proyectos basados en las direcciones recomendadas por Memex.

La estrategia marcada en este plan visionaria al grado tal que puede propiciar cambios radicales y mejoras en los servicios bibliotecarios especiales, públicos y académicos. Esta visión puede llevarse a cabo con la tecnología actual. Al mismo tiempo, retará a la estructura política y social dentro de la cual la biblioteca operará: en bibliotecas académicas, el tradicional énfasis en las colecciones locales, extendiendo los problemas de acreditación; en las bibliotecas públicas, el potencial del ramo electrónico y las bibliotecas centrales totalmente disponibles al público; y para las bibliotecas especiales, nuevas oportunidades de compartir colecciones y redes.

El ambiente en el que este plan ha sido desarrollado es, en el momento, dominado por el sentido de los límites de la biblioteca. La expansión continua y rápido crecimiento de las colecciones locales de las bibliotecas académicas están claramente ante un final. Bibliotecas corporativas, y aun las bibliotecas de leyes, están enfrentando la situación de operar dentro de un clima de dificultades económicas, así como con una competencia muy activa por parte de las fuentes comerciales de información. Por ejemplo, las bibliotecas públicas pueden ser vistas como un servicio deseable pero no crucial en el municipio en una época en la que las inversiones en seguridad y salud están siendo recortadas.

Adicionalmente, las bibliotecas en general tienen una alta relación trabajo - costo, y los costos de trabajo se incrementan cada vez más, sin involucrar las inversiones en automatización. Es difícil para las bibliotecas obtener capital, impulso o fondos para actividades de innovación, y aquellas iniciativas intensivamente tecnológicas que ofrecen la posibilidad de disminuir los costos de labor provocan la oposición del personal de la biblioteca.

Sin embargo, las bibliotecas han llevado a cabo éxitos considerables en las pasadas dos décadas mejorando tanto los servicios como su credibilidad dentro de las organizaciones, y estos cambios positivos han sido realizados en gran medida gracias al uso juicioso de las tecnologías de información. Los avances en cómputo y tecnología de información de han dado en perfecta cronología: la caída continua de los costos de cómputo, el crecimiento de las redes privadas e Internet, el incremento explosivo de las bases de datos disponibles al público.

Por ejemplo, OCLC se ha constituido en una de las mayores organizaciones de redes de cómputo en el mundo creando una red cooperativa de catálogos de más de 6,000 bibliotecas alrededor del mundo. Catálogos de acceso público en línea ahora dan servicio a millones de usuarios. Las bibliotecas se han convertido en el cliente más grande de edición de CD-ROM's; más de 30,000 ediciones en medios ópticos como los ofrecidos por InfoTrac y SilverPlatter están suscritas en las bibliotecas de Estados Unidos.

Esta marcha de tecnología continuará en la siguiente década, lo cual resultará en mayores innovaciones que resultan difícil de predecir. Lo que resulta claro es que las bibliotecas pueden ir ahora más allá de la automatización de sus catálogos y colecciones, y hasta es posible resolver las limitaciones fiscales que existen actualmente.

Un plan estratégico recomienda un cambio de paradigma en los servicios bibliotecarios, y demuestra los pasos necesarios para proporcionar servicios de alta calidad con capacidades e inversiones operativas limitadas.

2.5. PUBLICACIONES SERIALES ELECTRÓNICAS EN WEB

2.5.1. ¿Qué es una publicación serial electrónica?

Primero es necesario establecer una definición satisfactoria de publicación serial. De acuerdo con el estándar 3297 de la International Standards Organization (ISO) sobre el International Standard Serials Number (ISSN):

“Una publicación, en cualquier medio, entregada en partes sucesivas, usualmente con identificaciones numéricas o cronológicas y que se intenta continuar con un final no predeterminado.”

La mayoría de las publicaciones seriales aparecen en uno de los siguientes formatos:

- Discos ópticos (CD-ROM)
- Discos de almacenamiento magnético

La mayoría de las publicaciones electrónicas son accedidas por alguno de los siguientes métodos:

- Comprado con el vendedor o representante
- Vía fax o computadora a través de línea telefónica
- Por redes de computadoras como es Internet

Algunas publicaciones electrónicas están a disposición a través de WWW en forma de documentos HTML como es de esperarse. Sin embargo, también es posible acceder a publicaciones seriales en formatos de correo electrónico, gopher, FTP o Archivos de servidores de listas empleando el WWW.

Debido a su efervescente popularidad y capacidad multimedia, el WWW es el medio que más rápidamente ha crecido y, el más interesante para la entrega de publicaciones seriales electrónicas.

2.5.2. ¿Qué tipo de publicaciones electrónicas seriales están disponibles en el WWW?

Existe una amplia variedad de publicaciones seriales disponibles a través del WWW, incluyendo a las siguientes:

- Gacetas escolares
- Noticias en correo electrónico
- Revistas, Periódicos
- Gacetas políticas

Sin embargo no solo se está limitado a estas.

2.5.3. Similitudes entre publicaciones seriales impresas y electrónicas

Publicaciones electrónicas en el WWW, especialmente gacetas escolares tienen frecuentemente estructura y contenido muy similar a la tradicional versión impresa. Por ejemplo:

- Son presentadas en partes o ejemplares sucesivos
- Tienen usualmente una identificación cronológica o numérica
- Pueden ostentar un código ISSN
- Algunas están protegidas por los derechos de autor
- Son producidas por un editor y otros miembros del staff
- Si son controladas, pasan por una oficina editorial
- Existen lineamientos claros para el envío de artículos a la publicación

2.5.4. Diferencias entre las publicaciones seriales impresas y electrónicas

También existen diferencias importantes entre las publicaciones seriales impresas y las electrónicas como las que aparecen en el WWW:

- Las publicaciones electrónicas son últimamente más baratas y ambientalmente mucho más limpias
 - ◊ Hay mucho menos papel involucrado
 - ◊ No hay costos adicionales de distribución por diseminación
- Los ejemplares electrónicos son mucho más dinámicos
 - ◊ Los errores pueden ser corregidos en la edición original
 - ◊ El contenido completo de los ejemplares es susceptible de ser indizado y buscado por palabra clave o frase
 - ◊ Los documentos son editados en lenguaje de hipertexto con marcas. Esto proporciona al documento mayor versatilidad y complejidad. Por ejemplo:
 - No existen páginas per se. Los lectores “navegan” en el documento siguiendo los vínculos de hipertexto desde los índices hasta el contenido específico
 - El contenido dentro de un documento puede ser vinculado al contenido con otros documentos. Por ejemplo:
 - ⇒ Una gaceta electrónica puede contener listas vinculadas a otros recursos en el WWW
 - ⇒ Los autores pueden revisar otros documentos y proporcionar vínculos a dichos documentos
- La audiencia potencial de las publicaciones electrónicas seriales es global en alcance, y, por supuesto, limitado a aquellos que tienen acceso a WWW
 - ◊ Esto es cierto para apreciar el verdadero impacto de las publicaciones seriales que requieren de capacidades multimedia. Con mayor frecuencia, las publicaciones seriales electrónicas en el WWW contienen los siguientes elementos:
 - ⇒ Texto con formato
 - ⇒ Imágenes gráficas
 - ⇒ Fragmentos de sonido

⇒ Fragmentos de video ocasionalmente

2.5.5. La opción de texto completo para publicaciones seriales electrónicas

El texto completo electrónico también empieza a aplicarse a las revistas, de tal modo que se ofrecen servicios de recuperación de información en formato electrónico. Cada vez se ve más el interés en los editores y bibliotecólogos es usar esta opción, por tal razón se han estudiado mecanismos tales como:

- la publicación paralela en papel y diskette
- la edición por demanda a partir del diskette
- la edición paralela con el CD-ROM, esta opción responde a dos necesidades: a la urgencia de contar con las novedades y la de tener acumulativos que sustituyan al papel
- la edición paralela en línea a través de Internet, liberando el último número al mismo tiempo en papel y en línea; o la liberación desfasada por uno o varios números, ofreciendo información no tan reciente y actualizada en línea
- la edición paralela en línea pero liberada sólo en redes locales
- la edición totalmente en línea

Dichos mecanismos son analizados y se busca su viabilidad.

Lo que se busca con el texto completo electrónico es una solución o una mejora en aspectos relacionados con tiempo y costo que detienen el rápido acceso a las publicaciones, como son el costo y tiempo de edición y de distribución y entrega postal; lo que quiere decir que la biblioteca tiene que hacer un análisis de costo- beneficio ante la oferta de texto completo, por que ya hablamos en este capítulo de lo que implica el texto completo.

El usuario tiene un punto de vista especial del texto completo en línea, lo ve como el acceso inmediato a la información que requiere y espera la publique el nuevo número de la revista, para adquirirla en el menor tiempo posible.

Otro factor que se debe tomar en cuenta es el punto de vista de los autores y de los editores, preocupados por la ganancias de las suscripciones y las regalías, el control de consultas y su traducción al pago por cada una de estas consultas. Sin embargo, estos temores son posibles de desaparecer si se discuten en conjunto y con todas las partes, ya que los representantes de la tecnología han ido resolviendo todas ellas; es más, algunas veces adelantándose a la petición y ofreciendo opciones tecnológicas.

Para que dichos factores garanticen el éxito es necesario replantearse el objetivo de las revistas académicas, el funcionamiento del derecho de autor, el pago de regalías, el costo de la primera copia, el costo diferenciado para las siguientes copias, para usuarios institucionales y para usuarios

personales; además del costo, la forma de cobro y pago utilizando las opciones tecnológicas y de crédito mundialmente establecidas.

Las publicaciones que en este momento se encuentran libremente en Internet, y las que todavía lo están discutiendo, generan costos tanto de edición como de mantenimiento para que circulen en los sitios WWW. Las bibliotecas tiene que conocer la existencias de estas revistas, para promover su uso y controlar su explotación. Los usuarios también tienen que conocer las reglas del servicio para adaptarse a ellas.

CAPÍTULO 3

CAPÍTULO 3

PLANTEAMIENTO DEL PROBLEMA

3.1. ANTECEDENTES

Las bibliotecas conforman un centros de referencia de información, propician la generación de nuevo conocimiento tanto en las ciencias como en las humanidades, en la tecnología como en las ciencias sociales y las artes. Se manifiesta como un detonador de la curiosidad que nos empuja a buscar más información que nos lleve a procesar conocimiento y en su caso generar nuevo conocimiento. Ponen a disposición de sus usuarios su acervo para consulta y préstamo

Desde su inicio como una red académica y de investigación, Internet ha sido un excelente canal a nivel mundial de distribución y acceso a información de todo tipo y origen. Ahora, con el fenómeno de su popularización, el público usuario de esta tecnología crece a tasas exponenciales. Todas estas personas están ávidas de información y de tener acceso a mas y más servicios.

En estos momento Internet ofrece facilidades enfocadas a soportar un gran tráfico de información, sistemas de cómputo enfocadas cada vez más a las necesidades de conocimiento del usuario, servicios en línea que proporcionen todo tipo de datos; la tecnología se está moviendo más rápido para alcanzar este objetivo.

Esto ha representado una oportunidad única para las bibliotecas de participar mas activamente en la difusión del conocimiento, ya que se cuenta con la tecnología necesaria para llevar los acervos mas allá de los muros de los edificios de las bibliotecas y colocarlos en la pantalla de la computadora del usuario, quien puede estar en otro edificio o al otro lado del mundo.

Sin embargo, no se debe olvidar que su objetivo es conservar los elementos esenciales que la van a tipificar como biblioteca: rescate, acopio, organización, conservación y difusión de los registros del conocimiento universal; sólo que la forma de hacerlo variará y el énfasis de sus procedimientos estará en lograr acceder al usuario y a la información, independiente de la ubicación y el formato.

La tecnología alrededor de Internet ayuda a cumplir con estos objetivos, en una biblioteca en Internet se deben de cumplir los siguientes objetivos:

- Contar con servicios de información remotos, a fin de acceder a la información desde los diferente entornos académicos, sociales y domésticos.
- Tener variedad de programas de ayuda al usuario para que él mismo acceda a la información que desee, se de cuenta con qué cuenta la biblioteca, que servicios vía electrónico están a su alcance, etc.
- Una oferta variada y plural de información que será impulsada por los desarrollos tecnológicos. Distinción y apoyo prioritario a la especificidad de la demanda y a la individualidad del usuario.
- El procesamiento de la información deberá aportar un valor agregado a la fuente primaria y a la ficha bibliográfica.

- Se tendrán que crear y perfeccionar los procesos de búsqueda de información y además se simplificarán, ya que en su mayoría no serán cara a cara con el usuario sino a través de la red.

La presencia de las bibliotecas en Internet es un paso que ha sido dado de diferentes maneras, desde la creación de aplicaciones y protocolos propietarios, hasta el desarrollo de interfaces con los ya existentes. Sin embargo, con el World Wide Web se presenta una combinación muy interesante entre interacción con el usuario, la posibilidad de presentar elementos de texto, imágenes, sonidos, etc., y la ventaja de emplear una interfaz sencilla y muy conocida como es el visualizador de WWW.

La relación entre acervos de diversas bibliotecas es inmediata, ya que si tienen presencia en Internet, solo es necesario proporcionar la dirección adecuada y el usuario se desplaza de un acervo a otro. Además, se tiene la posibilidad de realizar búsquedas de información en varios acervos simultáneamente, presentando al usuario la suma de la respuesta de todos los sistemas de búsqueda. Sin embargo, esto constituye que la biblioteca integre a su funcionamiento actual nuevas actividades y responsabilidades.

Como ya se mencionó en el capítulo II en la Biblioteca Electrónica los elementos básicos son los usuarios y la información; es muy importante asentar que las colecciones de una biblioteca están formadas por todo tipo de información que, independientemente de su formato y medio de presentación, papel, audiovisual y electrónico, deberán: seleccionarse, adquirirse, analizarse, organizarse, clasificarse, diseminarse, difundirse y promocionarse.

El acervo de las Bibliotecas Electrónicas no solo consta de libros, también de material hemerográfico, material audiovisual (videocassettes), etc. Por lo tanto las publicaciones seriales cobran una importancia especial, implica un concepto global y completo de la revista en medio electrónico; los sistemas en WWW pueden ofrecer servicios donde puede ser consultado el título de la revista, su correspondiente índice y resumen, hasta artículo por artículo a través de Internet.

3.1.1. Biblioteca de la Dirección General de Servicios de Cómputo Académico

La Biblioteca de la Dirección General de Servicios de Cómputo Académico es una biblioteca universitaria especializada en computación y sus aplicaciones en informática y en telecomunicaciones; coordina las dos bibliotecas que se encuentran en los dos centros de extensión de DGSCA, la de Centro Nuevo León y Centro Mascarones.

Como biblioteca perteneciente a una dependencia universitaria su principal objetivo es obtener información y documentación relevantes a los programas de docencia e investigación de la Dependencia, hacerla asequible a la comunidad, orientarla en su manejo; así como proporcionar los servicios bibliotecarios necesarios para transmitir y acrecentar el conocimiento de la comunidad universitaria y del público en general.

Las funciones de la biblioteca son:

- a) Seleccionar y adquirir, en forma permanente, el material bibliohemerográfico especializado en computación con el fin de mantener el acervo actualizado.
- b) Organizar y mantener las colecciones que conforman el acervo.
- c) Proporcionar a los usuarios los servicios en forma precisa y oportuna, bien sea por medios manuales o automatizados.

- d) Vigilar el cumplimiento del Reglamento General del Sistema Bibliotecario de la UNAM y del presente Reglamento.

El objetivo de cualquier biblioteca es satisfacer la necesidad de sus usuarios, consciente de esto la biblioteca de la Dirección General de Servicios de Cómputo Académico presta servicios orientados a cumplir esta meta.

3.1.2. Servicios que ofrece la Biblioteca

Los servicios que brinda la Biblioteca de DGSCA son:

- a) Préstamo en sala.
- b) Préstamo a domicilio.
- c) Préstamo Interbibliotecario.
- d) Hemeroteca.
- e) Videoteca.
- f) Consulta a Bases de Datos.
- g) Acervo de libros en línea.
- h) Sistema en línea de gestión administrativa.
- i) Lector Automático para Invidentes.
- j) Programoteca.
- k) Internet.
- l) Visitas guiadas.
- m) Asesorías.
- n) Fotocopiado.

a) Préstamo en Sala.

Se ofrece a los usuarios de la Biblioteca a través del sistema de estantería abierta, es decir el usuario busca el material en el catálogo automatizado y después lo localiza en los estantes de acuerdo a la clasificación indicada.

b) Préstamo a Domicilio.

c) Préstamo Interbibliotecario.

d) Hemeroteca

Los fascículos de las publicaciones seriadas están disponibles para los usuarios.

e) Videoteca

Los videocassetes están disponibles para consulta en sala, a domicilio o a través de préstamo interbibliotecario.

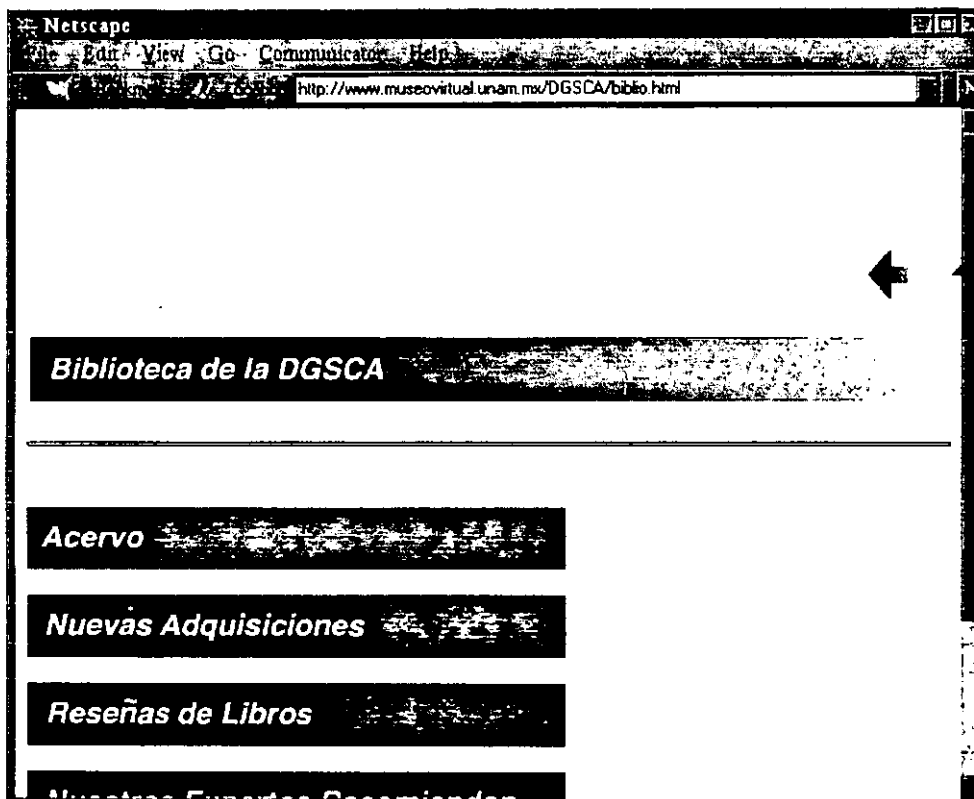
f) Consulta a Bases de Datos.

Servicio a través del cual los usuarios realizan búsquedas de información de su interés en bases de datos en discos compactos.

g) Acervo de libros en línea.

La Biblioteca de DGSCA tiene un acervo que está integrado por más de 6, 000 títulos de libros especializados en Supercómputo, Redes, TCP/IP, Telecomunicaciones, Lenguajes y Paquetería, entre otros.

Se tiene el catálogo de este acervo en línea, disponible a través de Internet mediante el sistema integral denominado ALEPH con una interfaz con WWW, el cual permite la búsqueda de información por: Autor, Título, Temas, Palabra clave en títulos y búsqueda combinada; de la misma forma da información sobre la disponibilidad del ejemplar.



Acervo Histórico

- Libros

El acervo, en total, está integrado por más de 6, 000 títulos de libros especializados en Supercómputo, Redes, TCP/IP, Telecomunicaciones, Lenguajes y Paquetería, entre otros.

El catálogo en línea se encuentra disponible mediante el sistema integral denominado ALEPH, el cual permite la búsqueda de información por: Autor, Título, Temas, Palabra clave en títulos y búsqueda combinada.


Además es posible consultar remotamente nuestro Catálogo Bibliográfico

Por otra parte, a partir de Noviembre de 1996 la Biblioteca cuenta también con el software Ariel, para el envío de documentos a través de la red Internet.

- Revistas


Welcome to ALEPH version 326 demonstration - Netscape

http://132.248.120.46:4500/ALEPH/



**UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO.**

**Dirección General de Servicios de Cómputo
Académico**



ESPAÑOL

Para acceder al catálogo haga click en el icono de la bandera.

Seleccione una de las siguientes:

información de la base Información de la base de datos.

El catálogo ofrece tres formas de búsqueda:

Busca x autores, título... Búsqueda en los índices por: Autor, Título, Tema, Clasificación o Serie

Busca x índices Búsqueda libre en los índices

Busca x palabras Búsqueda por palabras clave

Base de datos - Netscape





http://132.248.120.46:4500/ALEPH/spa/L41/L41/SCAN

Búsqueda en los Índices

Elija










- Índices para hojear
- Comience aquí =>

Escriba su petición en la casilla y presione el botón *Buscar*

	Busca x Índices 	Busca x palabras 	Regresar 
--	---	--	--






Buscar en la lista por: Titulo






Los números en la lista indican el número de documentos bajo cada término.
 Para ver los documentos relacionados, presione el icono. Use la flecha para subir y bajar en la lista.

		
	1	Microprocessor systems / M. aumiaux ; translation by arleta starza and david brailsford
	1	Microprocessors and logic design / Ronald l. krutz
	1	Microprocessors and microcomputers.
	1	Microprocessors and microcomputers : Their use and programming / Eric huggins
	1	Microprocessors : Design and applications / Andrew veronis
	1	Microprocessors in instruments and control
	1	Microprocessors/ Microcomputers/ System design / William d. simpson, staff consultant
	1	Microprocessors : Technology, architecture, and applications / Daniel r. mcglynn

Documento completo - Netscape
 http://132.248.120.46:4500/ALEPH/SPA/L41/L41/L41/FIND-ACC/0003635

Documento completo

	<i>Clasif.</i>	QA76.5 A85
	<i>Autor</i>	Aumiaux, Michel
	<i>Título</i>	Microprocessor systems / M. aumiaux ; translation by arleta starza and david bralsford
	<i>Imprim.</i>	Chichester, eng. : Wiley, c1982
	<i>Descr.</i>	217 p.
	<i>ISBN</i>	0-471-10129-x
	<i>ISBN</i>	0-471-10129-x
	<i>Tema</i>	Microprocesadores (programacion)
	<i>Copias</i>	

h) Sistema en línea de gestión administrativa

Se tiene el servicio automatizado de la gestión administrativa y de recuperación de información, a través de un sistema integral conectado a RedUNAM¹, a través del mismo sistema ALEPH y WWW. En este sistema se tienen los módulos de catalogación y prestamos para ayudar a la administración a los bibliotecarios.

i) Lector Automático para Invidentes.

Este servicio se ofrece a través de una máquina lectora, completamente integrada y adaptada para personas invidentes o débiles visuales que deseen consultar libros, artículos o documentos en español o inglés.

j) Programoteca

Este servicio consiste en la reproducción de diskettes que acompañan las obras documentales del acervo de la Biblioteca.

k) Internet.

Mediante este servicio los usuarios realizan búsquedas de información en la red Internet. Se proporciona de acuerdo a la tarifa vigente.

¹ Red de la UNAM de computadoras que se conecta directamente al backbone de Internet

l) Visitas Guiadas.

Se programan a solicitud de los interesados con el propósito de dar a conocer los servicios bibliotecarios que se ofrecen a la comunidad, nuevos proyectos, organización y programas de automatización.

m) Asesorías.

Se proporcionan a solicitud de los interesados con el fin de compartir experiencias, proyectos y propuestas en torno a la gestión de los servicios bibliotecarios.

n) Fotocopiado.

3.1.3. Tipo de Usuarios de la Biblioteca de DGSCA

La Biblioteca ofrece servicios a los siguientes tipos de usuario:

a) Usuarios internos

- Personal académico de la dependencia
- Empleados de la dependencia.
- Profesores externos que imparten cursos en la dependencia.
- Estudiantes de servicio social, de apoyo, tesis y becarios inscritos en los distintos departamentos.

b) Usuarios externos

Toda persona que solicite los servicios de la biblioteca pertenezca o no a la comunidad universitaria. Registro de usuarios externos:

- Únicamente para préstamo interbibliotecario. La Biblioteca solicitante deberá tramitar por escrito el convenio respectivo, mismo que se establecerá de acuerdo a las normas de la ABIESI (Asociación de Bibliotecarios de Instituciones de Enseñanza Superior e Investigación).
- La Biblioteca solicitante quedará registrada como tal, al ser aceptada su solicitud por escrito.
- La vigencia del convenio será por un año natural.

3.1.4. Horario de servicios de la Biblioteca

La Biblioteca brindará sus servicios durante 11 horas diarias de lunes a viernes de 9.00 a.m. a 20:00 p.m. y de 9:00 a.m. a 14:00 p.m. los días sábados.

La Biblioteca suspenderá sus servicios:

- a) Los días señalados como feriados en la Universidad.
- b) Por causas de fuerza mayor.

3.1.5. Políticas de préstamo

a) Préstamo en Sala.

Se ofrece a los usuarios de la Biblioteca a través del sistema de estantería abierta, es decir el usuario busca el material en el catálogo automatizado y después lo localiza en los estantes de acuerdo a la clasificación indicada.

b) Préstamo a Domicilio de libros

Se ofrece a los usuarios bajo las siguientes modalidades:

- Préstamo regular.- Préstamo de 5 materiales por un plazo de 5 días hábiles, prorrogables hasta 2 veces, en caso de no existir reservación alguna.
- Préstamo Interbibliotecario. Es el préstamo establecido mediante convenios entre bibliotecas de la U.N.A.M. u instituciones educativas, públicas o privadas.

c) Préstamo de videocasette

Se establecen las siguientes modalidades:

- En Sala el usuario proporcionará una identificación vigente para consulta del material.
- Préstamo a domicilio únicamente para usuarios internos.
- Préstamo interbibliotecario de acuerdo a la tarifa vigente.

d) Préstamo de Hemeroteca

Los fascículos de las publicaciones seriadas se prestan por 5 días hábiles sin derecho a renovación.

3.1.6. Personal que labora en la Biblioteca

La biblioteca está conformada por la Coordinación de las Bibliotecas y las siguientes áreas :

- Procesos Bibliográficos
- Servicios al Público
- Circulación
- Consulta
- Hemeroteca

Funciones del Area de Hemeroteca

- Seleccionar, adquirir y organizar la colección de publicaciones seriadas de la biblioteca de C.U.
- Actualizar periódicamente esta colección.
- Coordinar las adquisiciones de las bibliotecas de los Centros Nuevo León y Mascarones, y llevar el registro por separado de cada una de ellas.
- Coordinar la preservación del material hemerográfico a través de encuadernación.

3.2. PLANTEAMIENTO DEL PROBLEMA

La Biblioteca de DGSCA entró a la dinámica de ofrecer sus servicios más allá de sus instalaciones a través del sistema de consulta a su acervo de libros por medio del software comercial ALEPH y una interfaz WWW, sin embargo con el material hemerográfico aún manejaban únicamente el material impreso.

Los usuarios tienen que acudir a la biblioteca para advertir dentro de los índices de contenido de las revistas, qué artículo de qué revista contenía la información que ellos necesitaban. De acuerdo a los reglamentos de la biblioteca, mismos que ya se presentaron, el préstamo de revista a los usuarios se limita a los del tipo internos y aún en estos casos el préstamo es por 5 días hábiles sin derecho a renovación, para los usuarios externos no hay este tipo de servicio.

Todo esto dificultaba que el usuario pudiera obtener el artículo su de interés una vez que veía el índice de las revistas disponibles.

Las revistas representan un acervo de información fresca, sobre todo en ámbitos como son el cómputo y las telecomunicaciones, para el usuario es muy importante mantenerse actualizado en estas cuestiones, el acervo hemerográfico representa un gran apoyo para ellos dado que revistas de esta naturaleza son difíciles de obtener y los costos no son bajos. La demanda en este sentido es muy grande, los usuarios consultan ampliamente las publicaciones seriadas.

El éxito que se presentó con el servicio de consulta a través de Internet del acervo de libros de la biblioteca de DGSCA, nos presenta el hecho de que los usuarios de las bibliotecas tiene acceso a esta tecnología (conexión a la red Internet, uso de navegadores WWW, etc.) y que es importante presentar este tipo de opciones como servicios necesarios que debe ofrecer toda biblioteca que incursione en el mundo electrónico.

La Biblioteca de DGSCA advirtió la necesidad de llevar el acervo hemerográfico más allá de sus instalaciones así como se había hecho con el acervo de libros, sobre todo por la gran demanda de solicitudes de los artículos de revistas especializadas en cómputo. Por lo tanto se requiere el acervo electrónico de las publicaciones seriales disponible a través de la red Internet y con una interfaz amigable tal como WWW que permita al usuario advertir qué revistas se tienen disponibles, además de la posibilidad de ver las tablas de contenido de estas publicaciones y poder obtener alguno de los artículos que le interesen sin necesidad de acudir directamente a la biblioteca.

3.3. OBJETIVOS

Se busca establecer el servicio de consulta al acervo hemerográfico de la Biblioteca de la Dirección General de Servicios de Computo Académico (DGSCA), en donde se coloquen a disposición del público las tablas de contenido de las publicaciones seriales recibidas por esta biblioteca. Se plantean los siguientes objetivos:

- La meta es poner a disposición del público, título, ficha bibliográfica y tabla de contenido del acervo de publicaciones seriales que se tienen desde 1994 y brindar las opciones de búsqueda y recuperación de información sobre las tablas de contenido o índices de las revistas para advertir que artículo es de interés para el usuario y enviárselo vía un medio (por ejemplo correo electrónico).
- La búsqueda y recuperación de información tendrá dos modalidades, recuperación por orden alfabético que proporciona los años disponibles y cada año muestra a su vez las tablas de contenido de acuerdo a la periodicidad de la revista; la segunda modalidad es una búsqueda por palabra en los títulos de los artículos de todas las revistas.
- También es importante que el servicio a desarrollar sea una alternativa más para acceder al material conservando la opción de consulta del material impreso.
- Se debe contar con un mecanismo de actualización que esté al alcance del personal que colabora en el área de la Hemeroteca para alimentar el acervo electrónico de las revistas, buscando aprovechar sus aptitudes y evitar lo más posible que dicho procedimiento implique un aprendizaje muy alto para ellos.
- Se deben aprovechar los recursos humanos, de equipo de cómputo y de comunicaciones con que actualmente cuenta la biblioteca para llevar a cabo el desarrollo y brindar el servicio.

CAPÍTULO 4

CAPÍTULO 4

ANÁLISIS DEL SISTEMA

Actualmente la oportunidad que ofrecen las redes de computadoras de permitir tener acceso a grandes acervos de información, ha hecho que cada vez más organizaciones educativas, privadas y gubernamentales deseen tener disponible en red su información.

El servicio que busca ofrecer la biblioteca la Dirección General de Servicios de Cómputo Académico implica la consulta y recuperación de información hemerográfica empleando los servicios de Internet a través de una interfaz WWW.

La biblioteca desee poner en la red el servicio de consulta al acervo de sus publicaciones seriales, en donde se pondrán a disposición del público las **tablas de contenido** de las publicaciones recibidas por esta biblioteca, de tal suerte que, el usuario final pueda elegir que artículo de la publicación es de su interés y solicitarlo a través de este mismo servicio para que a vuelta de algún medio electrónico la biblioteca pueda otorgárselo.

Dentro del sistema que se desea ofrecer está la posibilidad de buscar una palabra o frase clave en las tablas de contenido de las revistas y ofrecer también la posibilidad de que si el usuario final lo desea, solicitar copia del artículo donde se encontró la frase o palabra.

El sistema a desarrollar implica desarrollo de software y procedimientos a seguir.

El ciclo de vida par un sistema exige un enfoque sistemático, secuencial y comienza con el análisis, siguiendo el desarrollo que implica el diseño, codificación y prueba.

Análisis del sistema.- el trabajo comienza estableciendo los requerimientos de todos los elementos del sistema. Esta visión del sistema es esencial cuando el software a desarrollar debe interrelacionarse con otros elementos tales como hardware, personas y datos. El principal elemento es el análisis al software a desarrollar, ayuda a comprender la naturaleza de los programas que hay que construir, tener el dominio de la información, así como la función e interfaces requeridas.

Diseño.- El diseño es realmente un proceso multipaso que se enfoca sobre tres atributos distintos del programa: estructura de datos, arquitectura del software y detalle procedural.

Codificación.- El diseño debe traducirse en una forma legible para la máquina. El paso de la codificación ejecuta esta tarea.

Prueba.- Una vez que se ha generado el código, comienza la prueba del programa, se realizan pruebas para asegurar que la entrada definida producirá los datos que se requieren.

El presente capítulo tiene como objetivo mostrar el análisis que se hizo del sistema, incluyendo:

- a) Identificación de necesidades, donde se advierte la información que se va a producir, la información que se va a suministrar y las funciones.
- b) Recursos humanos y de cómputo con que se cuentan.
- c) Estudio de viabilidad.
- d) Análisis de requerimientos del software a desarrollar.
- e) Análisis técnico

4.1. IDENTIFICACIÓN DE NECESIDADES

El primer paso el proceso de análisis del sistema implica la identificación de las necesidades, es el punto de partida en la evolución del sistema.

Después de las entrevista con el responsable de la biblioteca se plantearon las siguientes necesidades:

- Contar con un sistema que permita consultar el acervo hemerográfico de la Biblioteca de DGSCA a través de una conexión a Internet.
- Se requiere que dicho sistema incluya una interfaz lo suficientemente amigable al usuario final, el cual solo necesitará saber navegar en un sitio WWW para obtener dicho acervo.
- La obtención del acervo de revistas debe ofrecer dos modalidades:
 - a) La recuperación de información por orden alfabético de las publicaciones seriales que anota los años disponibles. Esto es equivalente al tradicional kardex manual que se utiliza en las bibliotecas por medio de tarjetas de registro. Cada año a su vez debe mostrar las tablas de contenido de acuerdo a la periodicidad de la revista. Con esto los usuarios pueden revistas las tablas de contenido como si estuvieran hojeando físicamente el índice de la revista.
 - b) Búsqueda en el contenido de los títulos de los artículos de las publicaciones seriales, es decir, en sus tablas de contenido. Esto a través de una forma WWW que solicite al usuario las palabras o términos a localizar.
- El servicio debe permitir al usuario obtener el acervo hemerográfico con las siguientes especificaciones:
 - a) Nombre de las revistas disponibles
 - b) Fichas bibliográficas
 - c) Tomos disponibles (mostrando el año, número y volumen)
 - d) Tablas de contenido del volumen
 - e) Las tablas de contenido deben incluir todos los nombres de los artículos de la revista
- Debe permitir al usuario elegir algunos de estos artículos que sean de su interés para que a través de un medio electrónico se los envíe el personal de la biblioteca.

- La interfaz debe incluir una solicitud dirigida al el usuario para recabar sus datos personales, para que posteriormente sean empleados a la hora en que se le envíe por algún medio el artículo que fue de su interés.
- Se debe integrar al sitio WWW el servicio de búsqueda de información que permita al usuario buscar palabras o terminos de su interés en los títulos de los artículos, es decir, en las tablas de contenido de las revistas.
- El servicio de búsqueda debe ofrecer la opción de hacerlo en tablas de contenido de todas las revistas o de alguna revista en particular.
- La información que arroja dicha búsqueda debe incluir el nombre del artículo, revista a la que pertenece, volumen, número y fecha de la revista.
- De la misma forma que la opción de recuperación de información en orden alfabético (revista por revista), el servicio de búsqueda por palabra debe permitir al usuario elegir algunos de estos artículos que sean de su interés para que a través de un medio electrónico se los envíe la biblioteca, también en esta opción el usuario debe dar sus datos personales.
- En la parte de captura de información del usuario final con sus datos personales se debe garantizar su confidencialidad e integridad cuando dichos datos viajen por la red para llegar a su destino (almacén de datos en el equipo de cómputo donde se prestará el servicio)
- Especificación del mecanismo que debe hacer el personal de la biblioteca para el envío de los artículos digitalizados a los usuarios finales.
- Definición del mecanismo para generar las tabla de contenido de las revistas en formato electrónico para integrarlos al servicio.
- También se requiere un mecanismo de actualización del acervo hemerográfico en donde la hemeroteca genere el material digital, dicho mecanismo debe ser lo más automático posible.
- Aprovechar los recursos de computo y humanos con los que cuenta la biblioteca; dado que ya se presta el servicio de disposición del acervo de libros a través de Internet por medio de una interfaz WWW, ya se cuentan con recursos de red y de cómputo.
- Para todas las etapas del desarrollo del sistema y la prestación del servicio se necesita trabajar con el software que implique el menor costo posible, se plantea la premisa de que se debe aprovechar todo el software libre que circula en Internet y que se apegue a las necesidades del proyecto.

4.2. RECURSOS HUMANOS Y DE CÓMPUTO

La biblioteca de DGSCA, cuenta con los siguientes recursos de cómputo que son considerados para el desarrollo del servicio y para la futura prestación del mismo:

1. Estación de trabajo Sun, SPARCstation 5:

- 32 MB en RAM
- 70 MHz
- Disco interno de 535 Mb
- Disco externo de 1.05 Gb
- Solaris 2.5.1
- Unidad externa de CDROM
- Cámara de video
- Bocinas

2. Equipo de digitalización.

- 1 Scanner HP 4c

3. Software de digitalización y de reconocimiento óptico de caracteres (OCR)

- Omni Page ver 4.0

4. Equipo PC para la digitalización y para el proceso de OCR

- 2 equipos tipo PC:

- PC ACER POWER P100
- 32 Mbytes en RAM
- 100 Mhz
- 1 disco duro: 1.2 GB
- Unidad lectora HD 3 ½
- Unidad CD-ROM
- Tarjeta de Sonido
- Tarjeta de Red

5. Conexión directa a RedUNAM Internet a través de una salida de 10 Mb por par trenzado.

Cabe aclarar que en la estación de trabajo se tiene actualmente el servicio en línea donde se ofrece el acervo de libros de la biblioteca y el servicio en línea de gestión administrativa (prestamos y catalogación), ambos a través de una interfaz WWW con el software ALEPH (Automated Library Expandable Program).

La biblioteca ya cuenta con un hardware y software de digitalización y OCR dado que otras áreas tienen este tipo de requerimientos.

Por otra parte en la hemeroteca se cuenta con un técnico académico de tiempo completo y 2 becarios de medio tiempo que trabajan con todo lo referente a las publicaciones seriales. Este personal cuenta con una experiencia en la digitalización de documentos y generación de OCR.

4.3. ESTUDIO DE VIABILIDAD

Para todo proyecto es necesario y prudente evaluar la viabilidad del mismo, para tal efecto se dividió tal labor en dos etapas que se consideran de suma importancia para el tipo de proyecto:

Viabilidad económica

Viabilidad legal

4.3.1. Viabilidad económica

La viabilidad económica es una evaluación del costo de desarrollo y puesta en marcha del servicio frente al beneficio producido.

En el momento de inicio de este proyecto el personal de la biblioteca manifestó la falta de un presupuesto alto para cualquier tipo de sistema o servicio en red que se quisiera desarrollar, por tal motivo solicitó el empleo de herramientas de dominio público o lo menos costosas posibles si es que se optaba por herramientas comerciales.

Por otro lado en el ambiente académico en que se encuentra esta dependencia y dado que nosotros laboramos para la misma instancia, el trabajo del análisis, investigación y desarrollo por nuestra parte implicó una labor de colaboración y que posteriormente fue tomado como tema del presente trabajo.

A la biblioteca de DGSCA no le representó un gasto por concepto de desarrollo del servicio o pago de honorarios por servicios profesionales al personal que trabajó en este proyecto.

El factor costo-beneficio del proyecto está dado por el tiempo que se invierte en el desarrollo y puesta en marcha, mientras se tiene el servicio tradicional donde el usuario tiene que acudir a las instalaciones de la biblioteca, la difusión de su acervo hemerográfico vía Internet se retrasaría más y esto afecta los objetivos de la biblioteca al incursionar al mundo electrónico y desear difundir el conocimiento a la sociedad. Se debe tomar en cuenta que ya se tenía el servicio en línea con el acervo de libros, por tal razón el usuario espera pronto que se ofrezcan otros acervos a través del mismo medio.

4.3.2. Viabilidad legal

En este punto se debe determinar cualquier infracción, violación o ilegalidad que pudiera resultar del desarrollo del sistema.

Por los requerimientos del sistema un punto importante de analizar es la opción de enviar vía un medio electrónico la digitalización de un artículo que el usuario final haya elegido al momento de usar el servicio en línea del acervo hemerográfico, sin embargo por el hecho de que no se cobra este servicio, no se lucra y como se hace en un ambiente académico no se incurre en algún tipo de falta, sobre todo en países como México donde por esta actividad no se tiene que pagar derechos de autor.

4.4. ANÁLISIS DE REQUERIMIENTOS

Es un proceso de refinamiento de lo establecido inicialmente en el planteamiento del proyecto, sus objetivos y necesidades. Especifica las funciones y comportamientos de los programas a desarrollar, indicar la interfaz con otros elementos del sistema en general. Permite refinar la asignación del software y representar el dominio de la información que será tratada por los programas. El análisis de requerimientos da a la etapa diseño la representación de la información y las funciones que pueden ser traducidas en datos.

En este punto se debe evaluar el flujo y estructura de la información, refinar las funciones del sistema y establecer las características de la interfaz del sistema.

4.4.1. Análisis de la Información

El software se construye para procesar datos; para transformar datos de una forma a otra; esto es, para aceptar entrada, manipularla de alguna forma y producir una salida. Esta etapa incluye:

- Contenido de la información
- Flujo de la información

4.4.1.1. Contenido de la información

En este punto se advierte el tipo de información con que inicialmente cuenta la biblioteca y que datos ellos requieren que arroje la interfaz WWW.

Identificación de las revistas

Dentro del sistema como tal cada revista debe identificarse, para ello tienen un nombre único y además la mayoría de ellas tiene un número de identificación que es el ISSN (International Standar Serial Number). Por ejemplo:

ACM Communications, ISSN: 0001-0782

Tabla de contenido de las revistas

El personal de la biblioteca tiene definido de forma general el formato de las tablas de contenido de las revistas que conformarán los datos de entrada en el flujo de información del sistema. A continuación se muestran la estructura de una de las tablas de contenido de una revista:

CONTENIDO

CONCEPTOS BASICOS

Tecnología Gigabit Ethernet: Redes de alto rendimiento con mayor velocidad

Una tecnología emergente que puede satisfacer las necesidades creadas por las aplicaciones recientes bajo redes locales. A esta nueva tecnología se le conoce como Gigabit Ethernet

Jesus Gutiérrez

14

SECCION ESPECIAL

En materia regulatoria ya no hay modelo a seguir: Casarus

-Educación y telecomunicaciones: la combinación perfecta

-La Cofetel no frena la apertura, al contrario, la promueve

-Hemos alcanzado los problemas de los países desarrollados en materia de telecomunicaciones: Casarus

Yolanda Aldaco

14

SU PLEMENTO PUBLICITARIO

Intersys

22

INTERNET

IP Versión 6: El soporte de la siguiente generación El acelerado crecimiento de Intenet y el surgimiento de nuevas aplicaciones requieren de un protocolo que sea capaz de enfrentar las grandes demandas de las redes en los próximos años. La nueva versión del protocolo IP busca resolver estos problemas

Hector Acevedo

24

Los componentes que se advierten en esta tabla de contenido son:

- a) Nombre de la revista
- b) Fecha
- c) Nombre del artículo
- d) Resumen del artículo
- e) Página de inicio

Los elementos que hacen de esta tabla de contenido única y bien identificada son los elementos inherentes a revista a la que pertenecen, y que son los elementos que el personal de la biblioteca desea que se le muestre al usuario:

- a) Nombre de la revista
- b) Volumen
- c) Mes
- d) Fecha
- f) Nombre del artículo
- g) Resumen del artículo
- e) Página de inicio

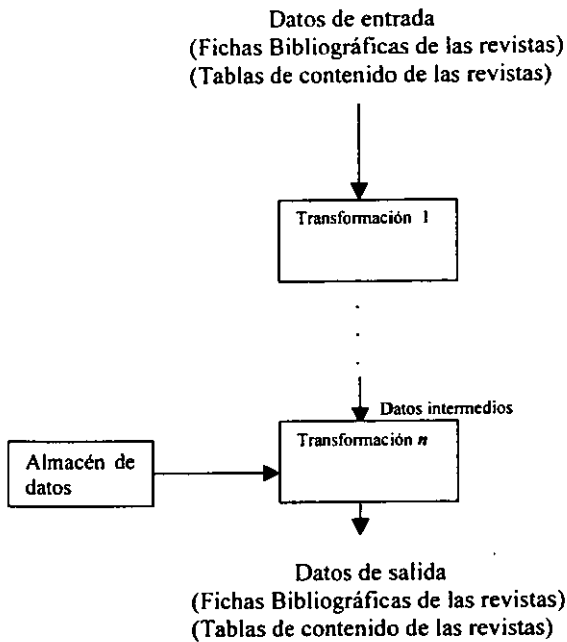
Ficha bibliográfica de la revista

Por otra parte dentro de las necesidades del personal de la biblioteca se manifestó el hecho de que el servicio logre mostrar al usuario final la ficha bibliográfica de la revista, los elementos que deben incluir esta ficha son:

- a) Título
- b) Editorial
- c) Lugar de edición
- d) ISSN/Num. Ref.
- e) Idioma
- f) Suscripción
- g) Periodicidad
- h) Tomos disponibles (generalmente, año y su volumen y número)

4.4.1.2. Flujo de la información

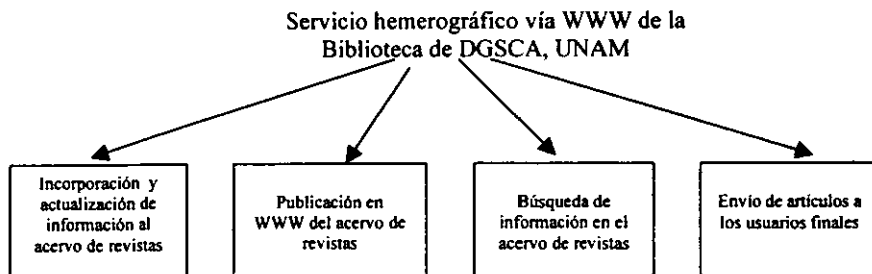
El flujo de la información representa la manera en la que los datos cambian conforme pasan a través de un sistema. La entrada se transforma en datos intermedios y más adelante se transforma en la salida. A lo largo de estos caminos de transformación, pueden introducirse datos adicionales.



El anterior cuadro muestra el flujo de la información y los posibles caminos que se tendrán para alcanzar el objetivo expuesto por la biblioteca; el servicio arrojará las fichas bibliográficas y las tablas de contenido al usuario que use un navegador WWW.

4.4.2. División del sistema y descripción funcional de cada módulo

De acuerdo al planteamiento del proyecto y a los objetivos del mismo se tiene que particionar (dividir) el proyecto en módulos que puedan ser fácilmente comprendidos, y establecer interfaces entre las partes, de forma que se realice la función global.



1. Incorporación y actualización de información al acervo de revistas

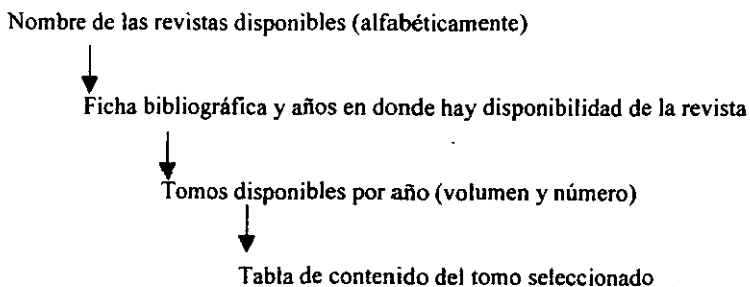
El personal de la biblioteca digitalizará las tablas de contenido de las revistas y se les aplicará el proceso de OCR para tenerlas en el formato digital e incorporarse al sistema. Así mismo se incorporarán las fichas bibliográficas de las mismas. Una vez que se tenga el material en formato digital se incorporará a un repositorio que será la entrada de datos al proceso automático que actualizará toda la información del servicio de WWW, que incluirá recuperación directa (alfabéticamente) revista por revista o por búsqueda de las tablas de contenido de las revistas.

2. Publicación en WWW del acervo de revistas

Dicho modulo se refiere a la interfaz WWW que mostrará al usuario el acervo de las revistas disponibles en la biblioteca en orden alfabético, es decir revista por revista. La información que se debe proporcionar es la siguiente:

1. ficha bibliográfica:
 - a) Titulo
 - b) Editorial
 - c) Lugar de edición
 - d) ISSN/Num. Ref.
 - e) Idioma
 - f) Suscripción
 - g) Periodicidad
2. años en donde hay disponibilidad de tomos
3. tomos disponibles por año (volumen y número)
4. tabla de contenido del tomo

La interfaz WWW debe estar dividida en niveles, así en cada nivel (o página Web) se le mostrará información anidada al usuario:



Dicho modulo también permitirá al usuario final elegir algunos de los artículos de las tablas de contenido para que por algún medio electrónico el personal de la biblioteca se lo envíe, para ello el usuario deberá a través de la misma interfaz WWW dar sus datos personales, dichos datos incluirán:

1. Nombre Completo
2. Domicilio electrónico (E-mail)
3. Teléfono
4. Dirección Postal
 - a) Calle y Número
 - b) Colonia
 - c) Delegación
 - d) Estado
 - e) País
 - f) Código Postal

3. Búsqueda en WWW de información en el acervo de revistas

Dicho servicio usa la misma interfaz WWW que el modulo de navegación por todo el acervo, pero en este caso presentará al usuario final la opción de buscar en todas las tabla de contenido (que contiene los títulos de los artículos) del acervo de las revistas, una palabra o termino específico, el servicio presentará los nombres de los artículos donde se encontró tal información, los datos arrojados son:

- a) Nombre del artículo
- b) En que revista se encuentra
- c) Volumen de la revista
- d) Numero de la revista y
- e) Fecha

Dicho modulo también permitirá al usuario final elegir algunos de los artículos para que por algún medio electrónico el personal de la biblioteca se lo envíe, para ello el usuario deberá a través de la misma interfaz WWW dar sus datos personales, dichos datos incluirán:

5. Nombre Completo
6. Domicilio electrónico (E-mail)
7. Teléfono
8. Dirección Postal
 - g) Calle y Número
 - h) Colonia
 - i) Delegación
 - j) Estado
 - k) País
 - l) Código Postal

Cuando el usuario capture sus datos personales, para evitar que algún mal intencionado usuario en la red intente interceptar esta información, se requiere que estos datos viajen en forma segura (es decir, cifrados) cuando los envíe el usuarios vía WWW.

Por un proceso posterior los datos personales del usuario y los artículos que él elija serán enviados vía correo electrónico al personal de la biblioteca para que atiendan esta solicitud.

4. Envío de artículos a los usuarios finales

Una vez que el programa de computadora envíe los datos del usuario y los artículos que éste eligió al personal de la biblioteca, procederán a digitalizar el o los artículos y posteriormente enviarán del artículo al usuario vía correo electrónico.

4.5. ANÁLISIS ORIENTADO AL FLUJO DE DATOS

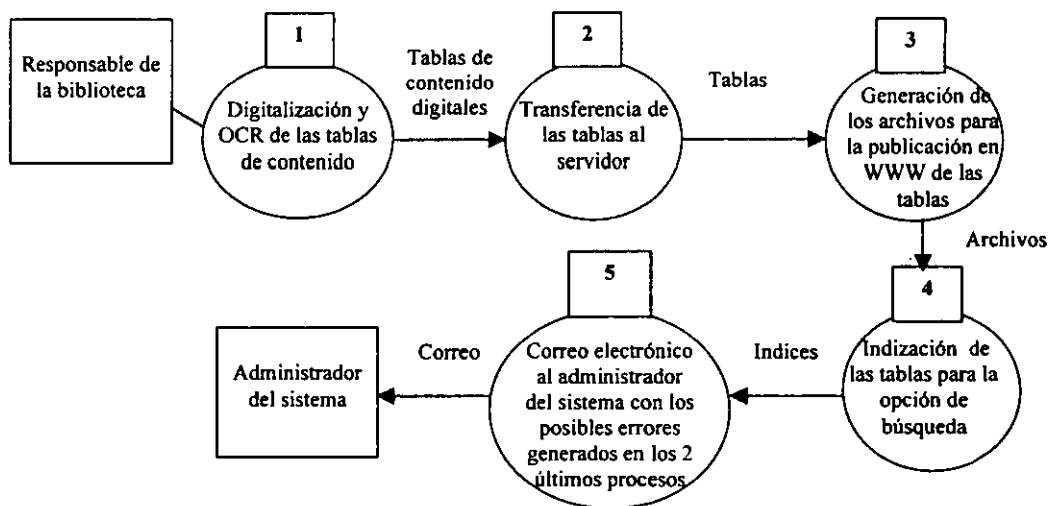
En este punto la información se transforma como un flujo a través de un sistema de cómputo. El sistema acepta entrada de distintas formas, aplica un hardware, software y elementos humanos para transformar la entrada en salida; y produce una salida en distintas formas.

A continuación se mostrará el análisis de cada uno de los módulos en que se dividió el sistema; para tal efecto nos auxiliaremos de la metodología de análisis de sistemas de Myres y Yourdon con los diagramas de flujos de datos y diccionario de datos para cada uno de los módulos.

4.5.1. Módulo de Incorporación y actualización de información al acervo de revistas

Diagramas de Flujos de datos

DFD 0 Incorporación y actualización de información al acervo de revistas



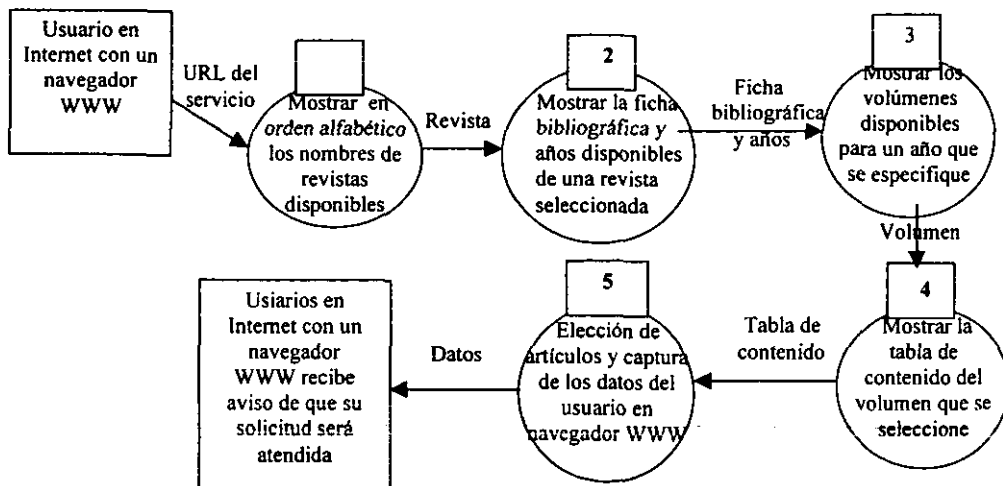
Diccionario de datos del modulo Actualización del acervo de las revistas

Tablas de contenido digitales	= Nombre de la revista + Volumen + Mes + Año + Nombre del artículo + Resumen del artículo + + página de inicio del artículo
Servidor	= Equipo donde reside el programa servidor de WWW, los programas desarrollados y el acervo digital de las tablas de contenido de las revistas.
Archivos en el servidor con las tablas de contenido	= En formato texto
Indices de las tablas de contenido para la búsqueda	= Indices generados por el software de Z39-50
Correo electrónico del administrador del servicio	= revdgsca@calmecac.dgsca.unam.mx

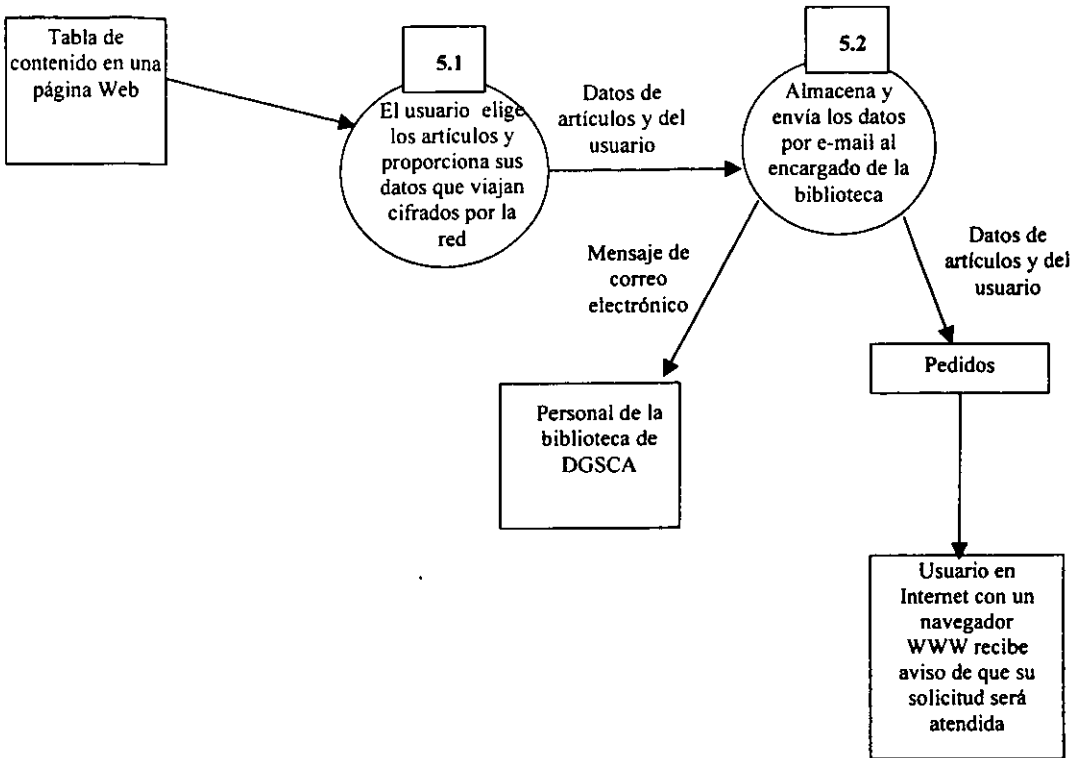
4.5.2. Módulo de Publicación en WWW del acervo de revistas

Diagramas de Flujos de datos

DFD 0 Publicación en WWW del acervo de revistas



DFD 5 Elección de artículos y captura de los datos del usuario



Diccionario de datos del modulo Navegación en las revistas disponibles

Nombre de revista	= {carácter válido}
Carácter valido	= [A-Z] -]
Ficha bibliográfica	= Título + Editorial + Lugar de edición + SSN / Num. Ref. + Idioma + suscripción + Periodicidad + Años disponibles
Año disponible para revistas	= [1994 1995 1996 1997 1998 1999]
Nombre de Volúmenes	= No. Volumen + No. Revista + Mes + Año ** Ejemplo: Vol. 0008 Num. 0001 JANUARY 1996

Contenido de las Tabla de contenido = Título de los artículos + Resumen del artículo +
Página de inicio del artículo

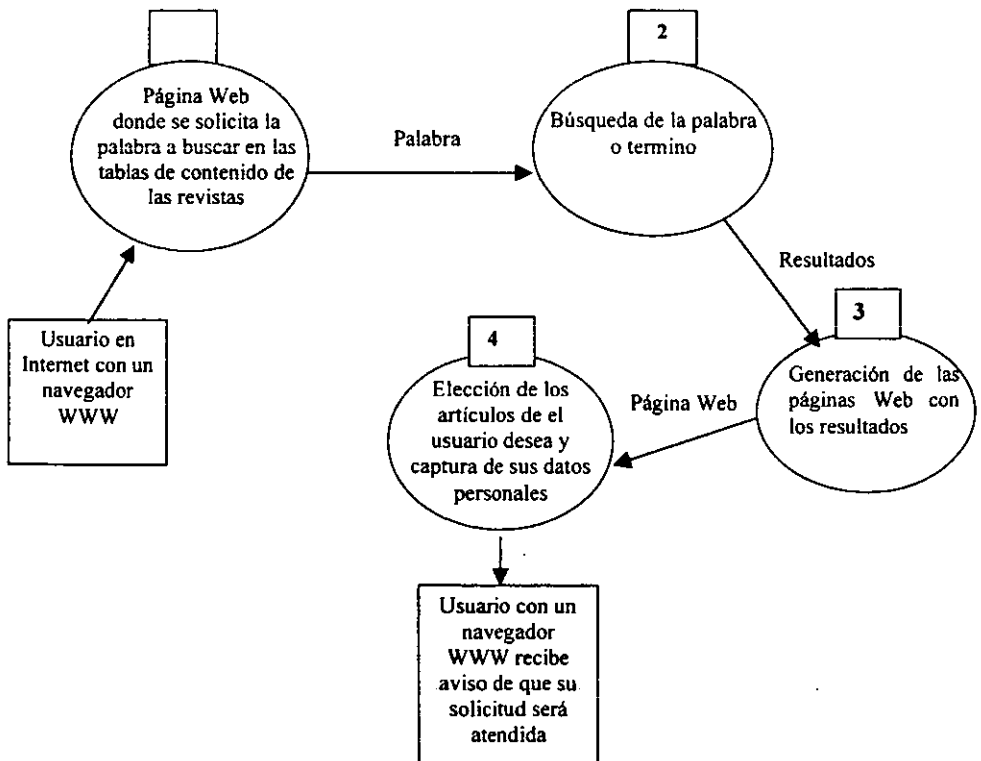
Datos del usuario al que se le
enviará copia del artículo = @ Nombre Completo + E-MAIL + Teléfono +
Dirección

Clave de correo del personal de
la biblioteca de DGSCA = biblio@servidor.unam.mx

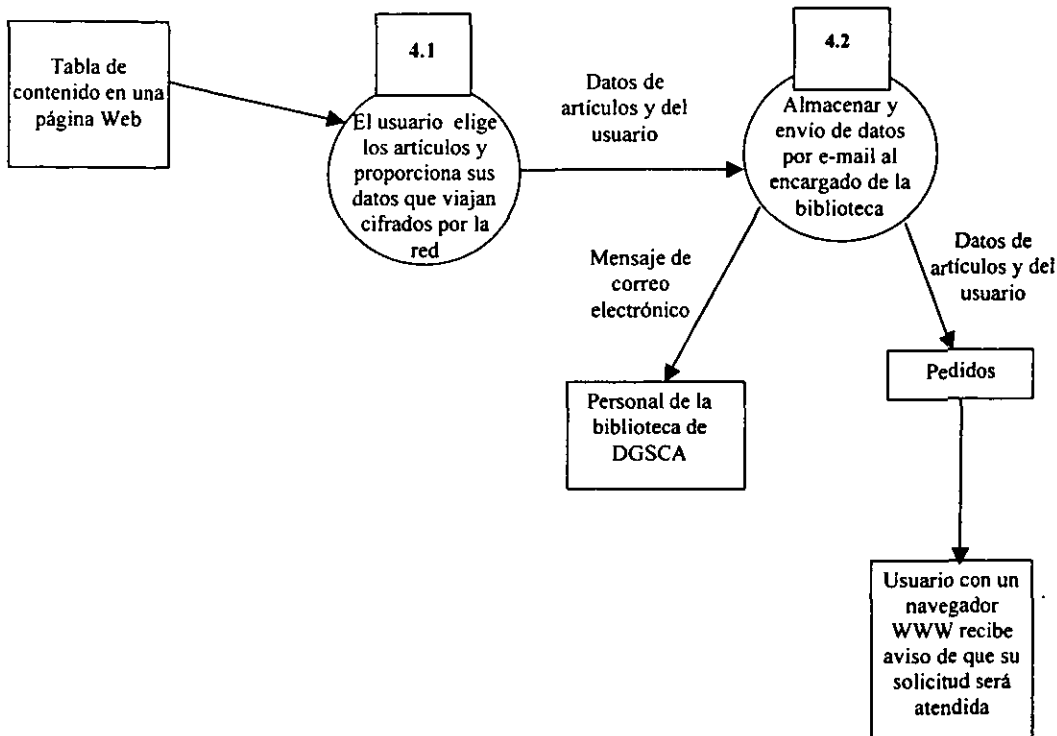
4.5.3. Módulo de Búsqueda en WWW de información en el acervo de revistas

Diagramas de Flujos de datos

DFD 0 Búsqueda en WWW de información en el acervo de revistas



DFD 4 Elección de artículos y captura de los datos del usuario



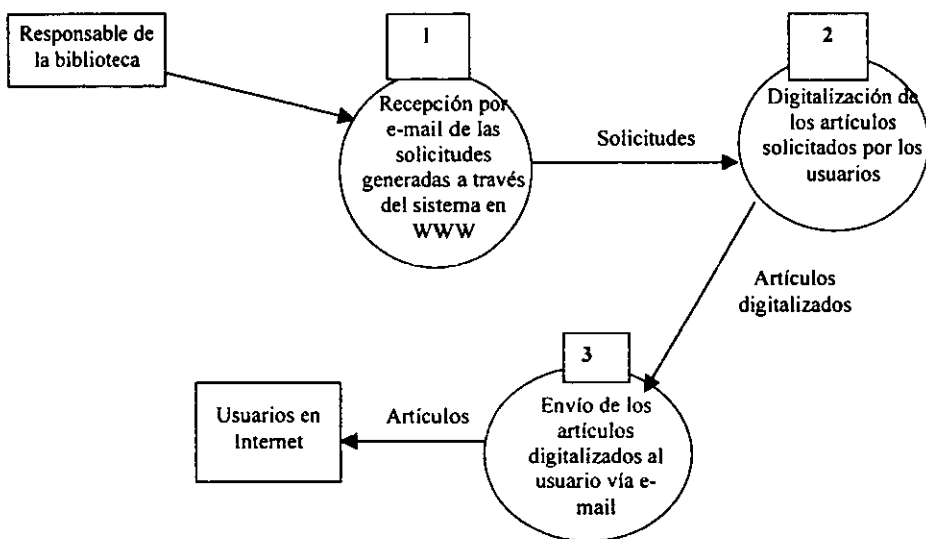
Diccionario de datos del modulo Búsqueda via WWW en el acervo de las revistas disponibles

- Palabra o termino a buscar = {carácter válido}
- carácter válido = [A-Z|a-z|0-9]
- Resultados de la búsqueda = Título de los artículo + Resumen del artículo + revista + volumen + número + fecha + opción de seleccionar el artículo
- Datos del usuario para que se le envíe copia del artículo = @ Nombre Completo + E-MAIL + Teléfono + Dirección postal
- Clave de correo del personal de la biblioteca de DGSCA = biblio@servidor.unam.mx

4.5.4. Modulo de Envío de los artículos a los usuarios finales

Diagramas de Flujos de datos

DFD 0 Envío de los artículos a los usuarios



4.6. ANÁLISIS TÉCNICO

El análisis técnico empieza con una definición de la viabilidad técnica del sistema propuesto. ¿Qué tecnologías se requieren para conseguir la funcionalidad y el rendimiento del sistema? ¿Qué nuevos materiales, métodos, esquemas o procesos se requieren para su desarrollo?

Para cada uno de las funcionalidades de los módulos del sistema, se necesita la implantación de software; dentro de los objetivos del sistema se planteó la necesidad de emplear software de dominio publico, por tal motivo se presentan a continuación las justificaciones del uso de estas tecnologías y de los esquemas empleados para los formatos, organización y representación de la información.

Módulos en los que se dividió el proyecto:

- a) Incorporación y actualización de información al acervo de revistas
- b) Publicación en WWW del acervo de revistas
- c) Búsqueda de información en el acervo de revistas
- d) Envío de artículos a los usuarios finales

Premisas:

- Como plataforma de desarrollo y de prestación del servicio se tiene un ambiente Unix con un equipo Sun Sparc Station 5.
- El servicio de “Recuperación y búsqueda de información en el acervo de las revistas” que ya presta la biblioteca usa un servidor de WWW (Apache), el software de búsqueda y recuperación de información es ALEPH (Automatic Library Expandable Program).

Dentro de los módulos a desarrollar se advierten las siguientes necesidades de software y métodos para la organización y representación de información:

- Servidor WWW para atender las peticiones del usuario a través de la interfaz WWW.
- Lenguaje de programación que trabaje en ambiente Unix, particularmente Solaris , que maneje los recursos de cómputo (archivos en disco duro, procesos, etc.) y que sea de sencillo manejo para trabajar el modulo de “Incorporación y actualización de información en el acervo”. Además que permita desarrollar programas que generen la comunicación entre el navegador de WWW y los recursos de cómputo del equipo servidor, es decir que permita desarrollar CGI's.
- Motor de búsqueda y recuperación de información tipo bibliográfica y hemerográfica que trabaje en ambiente Unix.

- Esquema de seguridad para cifrado de los datos personales del usuario que viajan por la red cuando el los captura en la interfaz WWW.
- Esquemas para formatos , organización y representación de información hemerográfica.

4.6.1. Servidor WWW

Un servidor WWW es un software que está ejecutándose en una computadora conectada a Internet. El propósito del servidor WWW es responder a los requerimientos para archivos WWW; cuando un usuario trabaja con un navegador WWW, se le está enviando requerimientos a todos los servidores WWW sobre todo el mundo. Los servidores enviarán archivos que son usados para construir páginas WWW que el usuario ve.

Cuando el servidor recibe un requerimiento (cuando el usuario elige una hiperliga de una página Web) este responde enviando el archivo solicitado. Los requerimientos y respuestas usan un protocolo llamado HyperText Transfer Protocol o simplemente HTTP. Una petición especifica la dirección del servidor, el archivo deseado y donde enviar esto.

Un servidor WWW puede usar cualquiera software disponible comercialmente o libre. Dentro del software comercial se encuentran por ejemplo el Internet Information Server IIS 3 de Microsoft, Netscape Enterprise Server, la compañía Sun ofrece su Java Web Server, existe también el Oracle Web Server, etc.

Dentro del software de dominio publico se encuentran varias opciones tales como: Servidor WWW de NCSA (NCSA HTTPD), WebReflex para Windows 3.1, 95 y NT, AOLserver, GoAhead WebServer, Roxen WebServer, Apache Web Server.

Por el hecho de que el personal de la biblioteca planteó el empleo de software de dominio publico para el desarrollo y prestación del servicio, se seleccionó Apache Web Server dada sus características y ventajas que posteriormente se mencionan.

Apache Server

El proyecto Apache es un esfuerzo de colaboración entre los desarrolladores de software para crear una implementación robusta, completa y disponible en forma libre de un servidor de HTTP (WWW).

El soporte que ofrece apache para las instituciones que lo usan es tan robusto que se pueden construir sistemas para propósitos experimentales y de misión critica.

Apache fue originalmente basado en el código e ideas encontradas en el servidor HTTP de NCSA httpd 1.3. Apache es considerado como un sistema el cual es rival y aventaja a muchos otros servidores HTTP basado en UNIX en términos de funcionalidad, eficiencia y velocidad. Apache es desde enero de 1997 el mas popular de los servidores WWW en Internet.

Apache es substancialmente más rápido que muchos otros servidores WWW libres. Apache es un servidor WWW libre muy rápido mientras que los comerciales son extremadamente rápido pero cuyo costo se va a miles de pesos. Apache se ejecuta en sitios que tienen millones de peticiones por día y no se han experimentado dificultades de rendimiento.

Apache se ejecuta aproximadamente en 3 millones de servidores en Internet (dato obtenido en Junio de 1999). Ha sido probado ampliamente por desarrolladores y usuarios. El grupo de desarrollo de Apache, mantiene rigurosos estándares antes de liberar nuevas versiones del servidor.

Por todas estas razones Apache Server fue seleccionado para implementar el sistema en WWW.

4.6.2. Empleo de CGI (Common Gateway Interface)

El Common Gateway Interface (CGI) surgió como la primera manera de presentar información dinámicamente generada en el World Wide Web. El CGI permite a la computadora generar páginas de Web instantáneamente ante la solicitud de un usuario en lugar de tener que tenerlas físicamente por adelantado. En este momento permanece como el único método estable y bien entendido para crear tales páginas. Java presenta problemas que aun no han sido resueltos y otros productos están por anunciarse pero seguramente pasará tiempo antes de considerarlos como tecnología madura.

Aunque los CGI's pueden emplearse para colocar efectos novedosos y curiosos a las páginas de Web, también permiten que Internet ofrezca el tipo de interactividad y aplicaciones dirigidas al usuario que se espera sean brindadas por las modernas computadoras.

Los usuarios de cómputo de hoy esperan obtener respuestas a preguntas particulares. Ya quedaron atrás los días en los que las personas tenían que conformarse con el reporte general y único que los administradores del centro de cómputo pasaban a todos los usuarios. En lugar de esto, cada vendedor, administrador, e ingeniero desea proporcionar preguntas específicas y obtener respuestas actualizadas. Y si una computadora puede hacer esto, ¿por qué no el WWW?

Cuando se ha navegado por el World Wide Web, seguramente se ha encontrado con documentos que lo han maravillado. Estos documentos pueden consistir de, entre otras cosas, formas que solicitan comentarios o información de registro, imágenes que permiten seleccionar cualquier parte de la misma y lo desplazan hacia otro documento, contadores que muestran el número de accesos al documento, y herramientas que permiten realizar búsquedas de información particular en bases de datos. En la mayoría de los casos, estos efectos son logrados a través del Common Gateway Interfaz, normalmente conocido como CGI.

El mecanismo CGI (Common Gateway Interface) surge como la primera forma de presentar información dinámicamente generada en el World Wide Web. Tal mecanismo permite a una computadora generar páginas de Web instantáneamente ante una solicitud del usuario, en lugar de haber escrito tal página con anterioridad.

Uno de los secretos peor guardados en Internet es que los CGI's son extremadamente simples. Son triviales en su diseño, y cualquier persona con poca experiencia en programación puede escribir programas rudimentarios que trabajen. Sólo cuando las necesidades son más demandantes es necesario dominar los trabajos más complejos del WWW.

4.6.2.1. Puentes (Gateways)

Los puentes (gateways) en WWW son programas empleados para tener acceso a información que no está directamente disponible al cliente. Por ejemplo, digamos que se tiene una base de datos en Oracle que contiene estadísticas de jugadores de béisbol del equipo de una compañía y se desea proporcionar tal información a través de WWW. ¿Cómo se puede lograr?. Ciertamente no se puede acceder con el cliente directamente al archivo de la base de datos (es decir, abrir un URL asociado a tal archivo) y esperar observar datos con algún significado.

Un CGI representa la solución al problema en la forma de un puente. Se puede emplear un lenguaje como `oraperl` o una extensión de `perl` para formar consultas SQL para leer la información contenida dentro de la base de datos. Una vez que se tiene la información, es posible darle formato y enviarla al cliente. En este caso, el programa CGI sirve como puente para la base de datos Oracle.

De la misma manera, trabaja los CGI trabajan con cualquier tipo de software que maneje la búsqueda y recuperación de información, tal como el software que implementa el protocolo Z39-50.

4.6.2.2. Programando CGI's

Ahora lo importante es determinar qué lenguaje de programación se puede usar. La respuesta a esta pregunta es muy simple: se puede emplear cualquier lenguaje de programación, aunque ciertos lenguajes son más apropiados que otros para el desarrollo de CGI's. Antes de seleccionar un lenguaje, es necesario considerar las siguientes características:

- Fácil manipulación de texto
- Capacidad para interactuar con otras librerías de programación y utilerías
- Capacidad de interpretar variables de ambiente (en UNIX)

Ahora, revisemos cada una de estas características en más detalle. La mayoría de las aplicaciones CGI involucran manipulación de texto de una forma u otra, por lo que de forma inherente, la comparación de patrones es muy importante. Por ejemplo, la información proveniente de una forma es usualmente "codificada" dividiendo una cadena por ciertos delimitadores.

La capacidad de un lenguaje para interactuar con otras aplicaciones, como son manejadores de bases de datos, es también muy importante. Esto incrementa el potencial de WWW permitiendo el desarrollo de puentes (gateways) hacia otras fuentes de información, como motores de bases de datos o librerías de manipulación de gráficos.

Finalmente, el último atributo que debe ser tomado en cuenta es la facilidad con la que el lenguaje puede acceder a las variables de ambiente. Estas variables constituyen la entrada del programa, por lo que son muy importantes.

Algunos de los lenguajes de programación para CGI's son AppleScript, C/C++, C Shell, Perl, Tcl, y VisualBasic. Veamos un repaso rápido de sus ventajas y desventajas.

AppleScript (Sólo Macintosh)

Desde la aparición del System 7.5, AppleScript es parte integral del Sistema Operativo Macintosh. Pensando en la debilidad de AppleScript inherente en cuestión de operadores de comparación de patrones, ciertas extensiones han sido desarrolladas para hacer más sencilla la manipulación de cierto tipo de datos. AppleScript además tiene la posibilidad de interactuar con otras aplicaciones Macintosh a través de AppleEvents. Por ejemplo, un programador de CGI's en MACintosh puede desarrollar un programa que presente una forma al usuario, decodifique el contenido de la forma, y realice una búsqueda en una base de datos empleando Microsoft Fox Pro directamente a través de AppleScript.

C/C++ (UNIX, Windowx, Macintosh)

C y C++ son muy populares entre los programadores, y muchos lo emplean para programación de CGI's. Estos lenguajes no son recomendados para programadores novatos, ya que imponen reglas estrictas para la declaración de variables y uso de memoria, así como de chequeo de tipos. Además, estos lenguajes presentan grandes debilidades en lo que se refiere a extensiones para bases de datos, habilidades para la comparación de patrones, aunque módulos y funciones específicas pueden ser escritas para desarrollar tales tareas.

Sin embargo, C y C++ tienen la ventaja de que una vez compilado, se genera un ejecutable binario, que emplea menos recursos que los empleados al usar un intérprete (como Perl y Tcl) al ejecutar el CGI.

C Shell (Sólo UNIX)

C Shell es débil en operadores de comparación de patrones, por lo que otras utilerías como sed y awk deben ser empleadas al momento de manipular cadenas de información. Sin embargo, hay una herramienta llamada `ucgi` escrita en C, que decodifica los datos de la forma y las almacena en variables de ambiente, que pueden ser utilizadas de forma muy sencilla. Evidentemente, la comunicación directa con una base de datos es imposible, a menos que sea realizada a través de una aplicación externa. Finalmente, C Shell tiene serios problemas y limitaciones que lo hacen peligroso para principiantes.

Tcl (Sólo UNIX)

Tcl está ganando popularidad como lenguaje de programación para CGI's. Consiste de un shell, `tclsh` que puede ser utilizado para ejecutar los programas. Como Perl, `tclsh` consiste de estructuras simples, pero es un poco más difícil de aprender y utilizar. Como Perl, Tcl cuenta con extensiones para bases de datos y manejo de gráficos. Además soporta expresiones regulares, pero es ineficiente en el manejo de tales expresiones, sobre todo al momento de la compilación.

VisualBasic (Sólo Windows)

VisualBasic es a Windows como AppleScript es a Macintosh en lo que a CGI se refiere. Con VisualBasic, se puede interactuar con otras aplicaciones como son bases de datos u hojas de

cálculo. Esto convierte a VisualBasic en una herramienta muy poderosa para el desarrollo de CGI's en PC, y es muy simple de aprender. Sin embargo, VisualBasic carece de operadores poderosos para la manipulación de cadenas.

Perl (UNIX, Windows, Macintosh)

Perl es por mucho el lenguaje más empleado para la programación de CGI's. Contiene características muy poderosas, y es muy fácil de aprender. Las ventajas de Perl incluyen:

Es muy portable y de disponibilidad inmediata

Contiene operadores de manipulación de cadenas muy poderosos, así como funciones para trabajar con datos binarios

Consiste de construcciones muy simples y concisas

La ejecución de comandos de shell es muy simple, y proporciona equivalentes de funciones de shell y UNIX muy útiles

Existen numerosas extensiones construidas sobre Perl para funciones especializadas, por ejemplo, existe `oraperl` (o extensiones DBI), que contienen funciones para la interacción con bases de datos Oracle.

Debido a esto, nosotros decidimos diseñar y desarrollar todos los programas en Perl.

4.6.2.3. Consideraciones para la programación de CGI's

Una vez que se ha elegido un lenguaje de programación, es necesario observar una serie de consideraciones que son necesarias tener en mente para crear documentos dinámicos efectivos.

Primero y lo más importante, es necesario entender qué tipo de información será presentada. Si es texto plano o HTML no es el problema. Sin embargo, si los datos no son interpretados por el cliente, será necesario desarrollar un puente para traducir los datos.

Esto nos lleva a otra importante consideración: Los datos originales deben estar organizados de tal forma que sea fácil para el puente leerlos y escribirlos. Una vez que se tiene el puente y se está en posibilidad de recuperar la información, ésta puede ser presentada de diversas maneras. Por ejemplo, si los datos son numéricos, se pueden crear gráficas. En otro sentido, si los datos consisten de objetos gráficos, se pueden modificar empleando diversas herramientas de manipulación.

En resumen, es necesario pensar sobre lo que se desea presentar y cómo afecta esto a la implementación del programa CGI. Esto asegura la creación de documentos dinámicos efectivos.

En el *apéndice A* se muestra un completo documento sobre CGI.

4.6.3. Protocolo Z39.50 para información bibliográfica y hemerográfica

4.6.3.1. Breve historia de Z39.50

El Protocolo ANSI/NISO Z39.50 es empleado para la Recuperación de Información en la Infraestructura y Servicios de Información. Las computadoras y las redes de telecomunicaciones ofrecen un potencial enorme para compartir datos en formato electrónico. El reto ha sido desarrollar mecanismos que puedan estandarizar la comunicación entre los sistemas de cómputo existentes. NISO responde a este reto estableciendo el Comité de Estándares en 1979 para trabajar en un protocolo de recuperación de información. Este trabajo culminó con Z39.50-1988. También, a finales de los 70's se funda el Consejo en Recursos Bibliotecarios (CLR por sus siglas en inglés) llevando a cabo el desarrollo de un protocolo experimental como parte del proyecto Linked System Project (LSP) para búsquedas en bases de datos bibliográficas y transferir registros entre la Biblioteca del Congreso, the Online Computer Library Center (OCLC), el Grupo de Desarrollo en Bibliotecas (the Research Libraries Group RLG), y la Red de Bibliotecas Washington (ahora Red de Bibliotecas Western (Western Library Network WLN). Este Protocolo de Sistemas Ligados proporcionó las bases para el protocolo Z39.50.

Poco después de la aprobación del estándar en 1988, un grupo de desarrolladores de Z39.50 comenzó a trabajar en el mejoramiento y expansión de la utilidad del estándar. NISO lanzó una versión revisada de Z39.50 en 1992. Mejoras continuas en el estándar por parte de sus desarrolladores ha dado como resultado una tercera versión del estándar (versión 3 borrador).

Los trabajos en ISO en estándares de recuperación de información resultaron en la aprobación por parte de ISO en 1991 de dos estándares internacionales que son compañeros de Z39.50: ISO 10162, Definición de un Servicio de Aplicación de Búsqueda y Recuperación (Search and Retrieve Application Service Definition) y el ISO 10163-1, Especificación del Protocolo de Búsqueda y Recuperación (Search and Retrieve Protocol Specification). El ANSI/NISO Z39.50-1992 armonizaba el estándar Norteamericano y los estándares internacionales, ya que el estándar norteamericano era un superconjunto de los estándares internacionales. A inicios de 1994, los desarrolladores de estándares internacionales tomaron una decisión crucial al comenzar el proceso de adopción de Z39.50 como estándar internacional. A partir de ese momento, no hubo más diferencias entre estándares nacionales o internacionales, ya que estaban armonizados. Además, Z39.50 proporciona las bases para el estándar internacional de recuperación de información.

4.6.3.2. ¿Qué es el protocolo Z39.50?

Z39.50 es el Estándar Nacional Americano para la Definición del Servicio y Aplicación y Especificaciones de Protocolo para la Interconexión de Sistemas Abiertos. La Organización Nacional de Estándares de Información (NISO por sus siglas en inglés), un desarrollador de estándares acreditado por el Instituto Nacional Americano de Estándares (ANSI por sus siglas en inglés) que atiende a las comunidades de bibliotecas, editoriales e información, aprobó el estándar original en 1988 (referido como Z39.50-1988 0 Versión 1). NISO publicó una versión revisada del estándar en 1992 (referida como Z39.50 o Versión 2).

El Z39.50 ANSI/NISO define la manera estándar en la que dos computadoras se comunican con el propósito de recuperar información. Z39.50 facilita el uso de bases de datos enormes estandarizando los procedimientos y características empleadas para la búsqueda y recuperación de información. Específicamente, Z39.50 soporta recuperación de información en ambiente distribuido cliente - servidor donde una computadora tomando el papel de cliente envía una solicitud de búsqueda a otra computadora actuando como servidor de información. La aplicación en el servidor realiza la búsqueda en uno o más bases de datos y genera un conjunto de registros resultantes que cumplen con el criterio de búsqueda proporcionado. El servidor responde el conjunto de registros para el procesamiento por parte del cliente. El poder de Z39.50 es que separa la interfaz del usuario (en el lado del cliente) de los servidores de información, aplicaciones de búsqueda y bases de datos. Z39.50 proporciona una vista consistente de la información de una gran variedad de fuentes, y ofrece a los desarrolladores en la parte del cliente la capacidad de integrar la información de un gran número de bases de datos y servidores.

El Z39.50 de ANSI/NISO puede ser implementado en cualquier plataforma. Esto significa que Z39.50 permite que distintos sistemas de cómputo - con distintos sistemas operativos, elementos físicos, aplicaciones de búsqueda, sistemas de administración de bases de datos - puedan trabajar en conjunto de manera transparente. La implementación de Z39.50 permite que una interfaz sea capaz de consultar múltiples sistemas proporcionando al usuario final el acceso más transparente posible a todos esos sistemas. Los usuarios consultan múltiples sistemas con un conjunto de comandos familiares y mostrando los resultados en su máquina local. No es necesario aprender nuevas técnicas de búsqueda y nuevos comandos. Los resultados de la búsqueda son presentados en el sistema local en los formatos y estilos a los que el usuario está acostumbrado o familiarizado.

Otra de las fortalezas del Z39.50 de ANSI/NISO es que es un Estándar Norteamericano que NISO desarrolla y mantiene, empleando procedimientos aprobados por consenso en ANSI, el principal coordinador de estandarización voluntaria en los Estados Unidos de Norteamérica. Z39.50 no es un estándar propietario y continuará respondiendo a las necesidades de los desarrolladores que empleen el estándar y a los usuarios de la información que se beneficien de sus implementaciones.

4.6.3.3. ¿Qué hace Z39.50?

Z39.50 es un protocolo de la capa de aplicaciones con respecto al modelo básico Open Systems Interconnection (OSI) desarrollado por la Organización Internacional de Estándares (ISO por sus siglas en inglés). Los protocolos de la capa de aplicación soportan los requisitos de comunicación de interacción directa con los programas residentes en las computadoras cliente y servidor y que desarrollan actividades específicas. Además de esto, como es un protocolo de la capa de aplicación, ha probado su total independencia del mecanismo de transporte que se encuentra debajo de él. Actualmente Z39.50 es ampliamente utilizado en el ambiente Internet sobre los servicios de transporte TCP/IP.

La arquitectura cliente/servidor proporciona el fundamento para las especificaciones de comunicación y recuperación de información de Z39.50. La aplicación para la interacción y despliegue con el usuario final (el cliente) está separado de la aplicación que maneja la información, realiza la búsqueda y proporciona el resultado (el servidor). Z39.50 no establece una interfaz con el usuario (es decir, no tiene una apariencia establecida), pero hay especificaciones y procedimientos dentro del protocolo que conciernen al cliente de Z39.50 (referido en el estándar como "origen") tales como la forma de iniciar una solicitud de recuperación de información y cómo solicitar

operaciones específicas del servidor (referido en el estándar como "objetivo") como por ejemplo la manipulación del conjunto de resultados y el formato en el que se envían los registros al cliente.

4.6.3.4. Empleando Z39.50 sobre Internet

Originalmente desarrollado para su empleo en redes de comunicación de computadoras empleando los protocolos Open Systems Interconnection (OSI), el ANSI/NISO Z39.50 es actualmente operacional en el ambiente Internet. En 1992 un grupo de desarrolladores (patrocinados por la Coalición de Información en Red o Coalition for Networked Information) establecen el marco de Interoperabilidad Z39.50 (Z39.50 Interoperability Testbed o ZIT) para lograr que su implementación funcionara en Internet, identificando y resolviendo los problemas de interoperabilidad entre sus implementaciones.

Las pruebas de interoperabilidad de diversas implementaciones de Z39.50 a través de ZIT, hacen claro que los servicios de transporte (es decir, el Transmission Control Protocol/Internet Protocol [TCP/IP]) de Internet pueden ser empleados en forma exitosa por el protocolo. Empleando Z39.50 sobre Internet es un gran paso adelante. Internet es una red de redes ampliamente difundida que conecta a millones de computadoras y a millones de usuarios alrededor del mundo. El conjunto de protocolos que soportan Internet (es decir, TCP/IP) están disponibles y son empleados; es una herramienta de conectividad omnipresente. Internet es hogar de almacenamiento masivo de recursos de información clave y conecta a millones de usuarios que necesitan acceso a tales recursos.

Por todas estas características, se decidió emplear el software de Z39-50 para el manejo de la información y las opciones de búsqueda.

En el *apéndice B* se muestra un documento sobre más especificaciones de Z39-50.

4.6.4. Introducción al SSL y certificación utilizando SSLoay

4.6.4.1. Introducción

El servicio de WWW transfiere información del servidor al cliente de forma transparente, es decir, lo que viaja en la red es la información que se está enviando tal cual del navegador al servidor y viceversa, si se desea que los datos viajen de una forma no transparente se tiene que recurrir a la cifrado de información.

Para trabajar con información cifrada se requiere trabajar con un protocolo especial llamado SSL (Secure Sockets Layer).

En este caso, se requiere que la información personal del usuario (nombre, domicilio postal, etc.) viaje segura (cifrada) cuando este capture esta información en la interfaz WWW.

4.6.4.2. Técnicas criptográficas

Entender SSL requiere una comprensión de algoritmos criptográficos, además de tener conocimiento sobre el conjunto de funciones de compresión de mensajes y también de firmas digitales. Estas técnicas son temas de libros completos y proveen las bases para la privacidad, integridad y autenticación en línea.

Algoritmos criptográficos

Supóngase que una persona requiere enviar un mensaje a su banco para transferir alguna cantidad de dinero. A esta persona le gustaría que el mensaje fuera privado, ya que incluirá información tal como su número de cuenta y la cantidad a transferir. Una solución es utilizar un algoritmo criptográfico, una técnica que podrá transformar su mensaje a una forma cifrada, nada legible excepto para quienes debe serlo. Una vez en esta forma, únicamente puede ser interpretado utilizando una "llave" secreta. Sin la llave el mensaje es inservible: los buenos algoritmos criptográficos hacen muy difícil a los intrusos recuperar el texto original haciendo que su esfuerzo no valga la pena.

Hay dos categorías de algoritmos criptográficos: convencionales y de llave pública.

Criptografía convencional, también conocida como criptografía simétrica, requiere que tanto el que emite como el que recibe compartan una llave: una pieza secreta de información que puede ser utilizada para codificar o decodificar un mensaje. Si la llave es secreta, entonces nadie, más que el que envía y el destinatario puede leer el mensaje. Si La persona y el banco conocen una llave secreta, entonces pueden enviarse mensajes privados el uno al otro. La tarea de escoger una llave antes de establecer la comunicación, sin embargo, puede ser problemática.

Criptografía de llave pública, también conocida como criptografía asimétrica, resuelve el problema del intercambio de llaves definiendo un algoritmo el cual utiliza dos llaves, cada una de las cuales puede ser utilizada para codificar un mensaje. Si una llave se utiliza para codificar entonces la otra debe ser usada para decodificar. Esto hace posible recibir mensajes seguros simplemente publicando una llave (la llave pública) y manteniendo en secreto la otra (la llave privada). Cualquiera puede codificar un mensaje utilizando la llave pública, pero únicamente el dueño de la llave privada podrá ser capaz de leerla. De esta forma, La persona puede enviar mensajes privados al dueño de una llave par (el banco), al codificar usando su llave pública. Solo el banco será capaz de decodificarlo.

Codificación de mensajes

Aunque la persona puede cifrar su mensaje para hacerlo privado, existe todavía la preocupación de que alguien pudiera modificar su mensaje original o sustituirlo con otro diferente, a fin de transferir el dinero a ellos mismos, por ejemplo. Una forma de garantizar la integridad del mensaje de la persona es crear un sumario conciso de su mensaje y enviarlo al banco también. Una vez recibido el mensaje, el banco crea su propio sumario y lo compara con el que envió La persona. Si son congruentes entonces los mensajes fueron recibidos intactos.

Un sumario como éste se denomina codificación del mensaje, o one-way hash.

La codificación de los mensajes se utiliza para crear representaciones breves y de tamaño fijo de mensajes largos y de tamaño variable. Los algoritmos para la codificación de mensajes, están diseñados para producir codificaciones únicas para mensajes distintos. También se diseñan para hacer difícil de determinar el mensaje codificado, y también hacen imposible el encontrar dos diferentes mensajes que hayan sido creados de la misma codificación. Por lo tanto se elimina la posibilidad de sustituir un mensaje por otro mientras se mantenga la misma codificación.

Otro desafío que esta persona enfrenta es encontrar una manera para enviar la estructura de codificación al banco de forma segura: Cuando éste es archivado, la integridad del mensaje asociado se asegura. Una forma de hacer esto es incluir la regla de codificación en una firma digital.

Firmas digitales

Cuando la persona envía un mensaje al banco, el banco necesita asegurarse de que el mensaje es realmente de ella, para que un intruso no solicite una transacción con su cuenta. Una firma digital, creada por la persona e incluida con el mensaje, sirve para este propósito.

Las firmas digitales son creadas al cifrar la codificación del mensaje y otra información (tal como un número de secuencia) con la llave privada del que envía. A pesar de que cualquiera puede descifrar la firma utilizando la llave pública, únicamente el firmante conoce la llave. Esto significa que ellos pueden tenerla firmada. Incluir la estructura de codificación en la firma significa que la firma es únicamente buena para el mensaje; también asegura la integridad del mensaje ya que nada puede cambiar la regla de codificación ni la firma. Para protegerse contra alguna intersección y reutilización de la firma por parte de algún intruso posteriormente, la firma contiene una secuencia única de números. Esto protege al banco de un reclamo de fraude de esta persona de que ella no envió el mensaje – únicamente ella pudo firmarlo.

4.6.4.3. Certificados

Aunque la persona pudo haber enviado un mensaje privado, firmado y asegurando la integridad del mensaje, ella todavía necesita asegurarse de que ella esta realmente comunicándose con el banco. Esto significa que necesita estar segura de que la llave pública que esta usando corresponde a la llave privada del banco. Y viceversa, el banco también necesita verificar que el mensaje firmado realmente corresponde a la firma de la persona. Si cada parte tiene un certificado el cual valida la identidad del otro, confirma la llave pública y es firmado por una institución que da fe de la autenticidad, entonces ellos estarán seguros que se están comunicando con quien cada uno piensa que es. La institución que da fe de la veracidad de sus identidades se denomina Autoridad certificadora, y los certificados son usados para la autenticación.

4.6.4.4. SSL

El SSL (Secure Sockets Layer) es una capa de protocolo la cual puede ser colocada entre una capa de red confiable orientada a la conexión (por ejemplo TCP/IP) y la capa de aplicación del protocolo (por ejemplo HTTP). SSL provee comunicación segura entre el cliente y el servidor al

permitir mutua autenticación, el uso de firmas digitales para la integridad y el cifrado para la privacidad.

El protocolo esta diseñado para soportar un rango de opciones para especificar algoritmos usados para la criptografía, la codificación de los mensajes y las firmas. Esto permite la sección de algoritmos para permitir a servidores específicos estar basados en aspectos legales, de exportación u otros conceptos, y también permitir al protocolo tomar ventaja de nuevos algoritmos. Las selecciones son negociadas entre el cliente y el servidor al iniciar a establecer una sesión por el protocolo.

Hay un número de versiones del protocolo SSL, como se muestra en la Tabla 3.

Versión	Fuente	Descripción	Navegador que lo incorpora
SSL 2.0	Publicado por Netscape.	Protocolo Original.	Netscape 3.0 Internet Explorer3.0
SL 3.0	Borrado de Internet expirado.	Revisiones para prevenir ataques específicos de seguridad.	Netscape 3.0 Internet Explorer 3.0
TLS 2.0	Borrador de IETF.	Revisión de SSL 3.0	Ninguno

Tabla 3 - Versiones del protocolo SSL

En el *apéndice C* se muestran más especificaciones de SSL.

4.6.5. EL ISSN (International Standar Serial Number) para identificación de publicaciones seriales

4.6.5.1. Aspectos generales del ISSN

El intercambio de información respecto a publicaciones seriales entre diversas organizaciones, y especialmente donde se involucra el uso de sistemas de cómputo, demanda la aplicación de códigos de identificación estándares. Las comunicaciones entre organizaciones individuales dentro del campo de las publicaciones seriales trasciende las fronteras y lenguajes de tal manera que se requiere de un código internacional numérico. El ISSN fue desarrollado para proporcionar un código breve, único y no ambiguo para este propósito. El Centro Internacional ISSN tiene la responsabilidad de la entrega y registro de ISSN y es la autoridad para la aplicación uniforme de reglas comunes y estándares. Un archivo de computadora conteniendo registros bibliográficos y el código ISSN es mantenido por el Centro Internacional, proporcionando diversos servicios y productos a partir de estos datos.

Como parte de su función como identificador único, el ISSN es empleado en aplicaciones tales como la adquisición de ejemplares de publicaciones seriales, la creación de catálogos unificados,

servicios de copiado y préstamo, control de inventarios, etc. El empleo con código de barras ha tenido un efecto significativo en aquellos países en los que se ha aplicado dentro de este contexto.

Seguramente en diversas ocasiones se ha empleado un código misterioso en las publicaciones seriales, principalmente al leer o consultar referencias concernientes a los artículos o las propias publicaciones seriales. Dicho código es el ISSN, que no es más que una clave internacional estándar que permite identificar a cualquier publicación serial independientemente del país de publicación, del lenguaje, alfabeto, frecuencia o medio.

El ISSN es un código numérico empleado únicamente como identificador: no tiene significado y no contiene información por sí mismo con respecto al origen o contenido de la publicación.

El ISSN toma la forma del acrónimo ISSN seguido por dos grupos de cuatro dígitos, separados por un guión. El octavo carácter es un dígito de control calculado de acuerdo a un algoritmo en módulo 11 con base a los 7 dígitos antecesores; este último carácter puede ser una "X" si el resultado del cálculo es igual a "10", esto con el fin de evitar una ambigüedad.

El ISSN está ligado a una forma estándar del título de la publicación serial, conocida como "título clave", el cual repite el título de la publicación, calificándola con elementos adicionales para distinguirla de otras publicaciones que tengan títulos idénticos. A continuación algunos ejemplos:

ISSN	Título clave	País, fecha de inicio y final
1140-3853	Babel (Arles)	Francia (1989-9999)
1105-0748	Babel (Atenas)	Grecia (1981-9999)
0521-9744	Babel (Bonn)	Holanda (1955-9999)
0327-6414	Babel (Buenos Aires)	Argentina (1988-9999)
0005-3503	Babel (Melbourne)	Australia (1965-9999)
1147-8306	Babel (París)	Francia (1990-9999)
1176-2893	Babel (Schorndorf)	Alemania (1983-9999)
0310-0022	Balance (Canberra)	Australia (1973-9999)
0900-4521	Balance (Kobenhavn. 1968)	Dinamarca (1968-9999)
0005-4216	Balance (Londres)	Gran Bretaña (1961-9999)
0404-6722	Balance (Madrid. 1954)	España (1954-9999)
0705-3495	Balance (Montreal)	Canadá (1978-1979)
1033-4092	Balance (New Farm)	Australia (1989-9999)

4.6.5.2. Caso de estudio en México

En México, el empleo del ISSN es controlado por el Centro Mexicano del ISSN, dependiente del CONACYT y, durante los tres últimos años ha presentado un importante crecimiento en el número de nuevos registros obtenidos, sin embargo, estamos aún muy lejos de que este código sea un elemento importante en todas las publicaciones periódicas nacionales. Según los datos del propio Centro Mexicano del ISSN, en 1992 habían 1688 registros, para 1993 el número de registros creció a 1805 y así se mantuvo hasta 1996 cuando subió a 2691 y, para 1997 el número fue de 3173. A continuación se muestran las estadísticas de crecimiento de registros en el Centro Internacional del ISSN.

4.6.6. SGML y los Sistemas de Biblioteca en Línea

4.6.6.1. Del Marc a las etiquetas, SGML y los Sistemas de Biblioteca en Línea

Etiquetas

Tradicionalmente, las etiquetas son instrucciones dadas a un tipógrafo sobre cómo disponer el texto - qué estilo y tamaño de tipo de letra emplear (es decir negrita, de 24 puntos courier), elementos de tipografía especiales (ornamentos, listas, etc.), y la presentación -. Las etiquetas no tienen nada que ver con el contenido del documento, y tiene relación con todo lo referente a su apariencia. Las etiquetas de este tipo son conocidas como etiquetas de procedimiento. La edición electrónica y las aplicaciones de procesamiento de palabras emplean las etiquetas de procedimiento definidas para tal sistema. Esto es, los sistemas tienen su propio conjunto de códigos (frecuentemente incrustado en el texto del documento) que portan el proceso - escala de tipos de letra, resaltado, etc.- sólo dentro del mismo sistema o con la aplicación específica en ejecución en una determinada plataforma. Generalmente, las etiquetas de este tipo son diseñadas con un único resultado en mente, tal como es la producción de páginas impresas en un estilo particular. Si el contenido de los documentos necesita ser rehusado con diferente estilo o formato, se pueden eliminar los códigos de etiquetas, y colocar nuevos códigos para el formateo distinto. Como cualquier bibliotecario que ha vivido la implementación de más de un sistema en línea puede atestiguar, cuando hay un cambio de aplicación (y algunas ocasiones plataforma de cómputo), la traducción de datos de un sistema a otro puede ser costoso, gran consumidor de tiempo y potencialmente un esfuerzo angustioso. El intercambio de documentos basados en etiquetas de procedimientos funciona sólo si ambas partes, el remitente y el destinatario emplean el mismo sistema.

Existe, sin embargo, otro tipo de etiquetas textuales, las etiquetas descriptivas. Estas etiquetas genéricas describen la estructura y/o contenido de un documento, en lugar de su apariencia física en la página o la pantalla. Así, el contenido de un documento es separado del estilo de presentación. Elementos dentro del documento (un capítulo, una estrofa, una nota de pie de página, una bibliografía) son organizados empleando códigos que definen qué elemento es, pero no cuál debe ser su apariencia. El contenido de documentos que son etiquetados descriptivamente puede ser rehusado para distintos propósitos y presentados en muchos estilos diferentes.

SGML

El SGML (Standar Generalized Markup Language) fue desarrollado inicialmente en 1970 como GML (Generalized Markup Language) y evolucionó en un estándar norteamericano y uno internacional. El SGML ha sido un estándar internacional desde Octubre de 1986 (estándar ISO 8879); ha sido ampliamente aceptado en los Estados Unidos de Norteamérica, Europa Occidental y Japón y, es empleado para diversos negocios y en aplicaciones industriales y académicas. El SGML es frecuentemente referido como un metalenguaje. Esto significa que el SGML no es un lenguaje individual, sino un lenguaje para describir otros lenguajes de etiquetas; esto es, el SGML son las reglas o el marco para definir lenguajes de etiquetas particulares.

El SGML proporciona etiquetas descriptivas y no de procedimientos; esto es, simplemente establece los nombres para categorizar las partes de un documento en lugar de especificar los procesos a los que son sometidos.

El SGML permite la descripción de información estructurada independientemente de su procesamiento, proporcionando una sintaxis estándar para definir las descripciones de clases de documentos. Dichas descripciones son llamadas definiciones de tipo de documentos (DTD's). Los DTD's definen los tipos de documentos y sus estructuras estableciendo qué elementos son requeridos en un tipo particular, y qué elementos pueden estar presentes en el documento. La estructura de un documento puede ser etiquetada y verificada contra un DTD - empleando un programa especial llamado analizador- para asegurar que es válido, y que es acorde a la estructura del tipo de documentos definido por el DTD. Los DTD's son en la mayoría de los casos escritos teniendo procesos o resultados específicos en mente, pero pueden ser fácilmente replanteados. Dado que las etiquetas SGML son independientes de la aplicación, significa que los documentos que cumplen con un determinado DTD pueden ser rehusados en una gran variedad de formas. Sabiendo que un documento está estructurado en una forma particular hace que los rehusos subsecuentes sean más sencillos.

Hay tres diferentes partes funcionales de un documento SGML. La primera especifica el conjunto de caracteres del documento; la segunda parte nombra el tipo de documento y, por tanto, las etiquetas específicas que pueden ser empleadas en el documento. La tercera parte de un documento SGML es el texto o contenido, marcado con las etiquetas SGML.

Por ejemplo, un "poema" puede ser definido como un tipo de documento. El DTD para el "poema" puede definir líneas y estrofas (etiquetas: <línea> y <estrofa>) como los elementos requeridos. Elementos opcionales pueden incluir <copla>, <octava>, <sexteto>, <notas de pie de página>, y <autor>; estos elementos pueden estar presentes en un documento definido como poema, pero pueden no estar. El propósito de la codificación del DTD y del SGML puede ser el imprimir todos los poemas etiquetados con las octavas en negritas y los sextetos en itálicas, empleando algún tipo de programa de traducción. Puede ser que después se necesite imprimir el mismo conjunto de programas solamente con las octavas en itálicas; pero no será necesario reetiquetar cada poema, solo es necesario alterar el programa de traducción que procesa el texto etiquetado. De esta forma, los posibles usos de los poemas etiquetados son virtualmente ilimitados ya que cualquier tipo de instrucciones de procesamiento pueden ser aplicadas. Diferentes tipos de instrucciones de procesamiento pueden ser asociadas con las mismas partes del archivo. Por ejemplo, un programa de análisis textual puede ignorar notas de pie de página en un poema etiquetado, mientras que otro programa puede imprimir al final de cada poema. El mismo programa de impresión puede extraer los nombres de los autores de los poemas para imprimirlos al inicio de cada poema; el programa de análisis textual puede emplear los nombres de los autores para crear una base de datos para búsquedas. La combinación de etiquetas descriptivas y definiciones de tipos de archivos permite que los documentos codificados en formato SGML sean procesados por distintas aplicaciones teniendo en mente diversos resultados. Debido a que el SGML es no propietario (no pertenece a una plataforma de cómputo específica o a una compañía de aplicaciones) y es un estándar internacional, los datos permanecen independientes de cualquier configuración de aplicaciones o de plataforma de cómputo, haciendo que el SGML y las aplicaciones que cumplen con el SGML sean extremadamente flexibles.

Descripción de la Codificación de Archivos

Las ayudas para la localización han servido como inventarios, registros, índices o guías para colecciones mantenidas en archivos y repositorios de manuscritos, bibliotecas y museos. Las ayudas para la localización generalmente describen una colección única, su organización intelectual, y, a diversos niveles de análisis, identificación de componentes de la colección. Ellos por sí mismos son

considerados metadatos (información acerca de la información). Las ayudas para la localización son típicamente documentos narrativos con párrafos de descripción y listas de chequeo; están estructurados jerárquicamente, desde un resumen de la colección completa hasta descripciones progresivamente detalladas de series y contenedores, carpetas o ejemplares.

Históricamente, las ayudas para la localización han tomado diversas formas, incluyendo registros en papel, *catálogo de tarjetas*, *panfletos* o *libros editados*, o documentos electrónicos. La automatización de las ayudas para la localización de acceso remoto y local se convirtieron en prioridad cuando herramientas de diseminación como Internet estuvieron disponibles y conforme se enfocaron en la digitalización de la información de herencia cultural contenida en manuscritos, colecciones fotográficas y de multiformato. El proyecto Berkeley Finding Aid Project, comenzó en 1993, explorando el empleo de SGML para codificar el texto en ayudas para la localización, proporcionando contenido independiente de la plataforma, con contenido, estructura y diseño de la presentación, permitiendo los vínculos a objetos digitales actuales, haciendo posible integrar a las ayudas para la localización en “bibliotecas digitales”. A través de investigaciones posteriores y discusiones entre profesionales de diversos ámbitos, el DTD (Document Type Definition) del FindAid desarrollado en Berkeley ha evolucionado a un nuevo DTD, el Encoded Archival Description (EAD), cuya versión alfa fue presentada en Enero de 1996 para su revisión entre los primeros desarrolladores. La versión beta fue presentada en Agosto de 1996.

La codificación SGML de la ayuda para la localización ayudará a integrar las fuentes de información más valiosas en catálogos bibliográficos como una piedra angular entre un catálogo de registros de acceso público en línea, una descripción más detallada de la colección (la ayuda para la localización), y el objeto digital. Otras discusiones entre los desarrolladores de EAD se centran en sistemas tales como la ayuda para la localización de información climática que diversas instituciones deben coleccionar o conjuntar, y de ahí, que sistemas de búsqueda proporcionaran acceso a tales datos; la relación de EAD con otras iniciativas de SGML para herencias culturales; el papel de vocabularios controlados y vínculos a archivos de autoridad; el empleo de visualizadores de WWW para manejar documentos SGML.

SGML, etiquetas y los Sistemas de Bibliotecas en Línea

¿Cuál es entonces la relación del SGML con los Sistemas de Bibliotecas en Línea?. Tres proyectos e iniciativas que están actualmente en desarrollo pueden servir de ejemplos de la utilidad de las etiquetas para la recuperación de información bibliográfica y su selección. Estos proyectos emplean SGML para proporcionar una estructura y acceso a información bibliográfica.

La Iniciativa de Codificación de Texto (<http://www.uic.edu/orgs/tei/>)

La Iniciativa de Codificación de Texto (TEI por sus siglas en inglés) es un proyecto de cooperación internacional para desarrollar los lineamientos para la preparación e intercambio de textos electrónicos para investigación escolar. TEI prepara los lineamientos para la Codificación de Texto e Intercambio (<http://etext.virginia.edu/TEI.html>) para desarrollar

“un esquema común de codificación de estructuras textuales complejas para reducir la diversidad de prácticas de codificación existentes, simplificando el procesamiento por máquina, e inspirando el intercambio de textos electrónicos.”

(Sperberg- McQueen, preface).

El encabezado TEI, es una parte obligatoria de TEI. Los textos que cumplen con este formato tienen cuatro partes principales: una descripción de archivo, que contiene una descripción bibliográfica

completa del archivo de computadora por sí mismo e incluye información sobre la fuente del texto fue. Una descripción de la codificación, que describe la relación entre el texto electrónico y su fuente; esto permite la descripción detallada de si el texto fue normalizado o cómo lo fue durante su transcripción, cómo el codificador resolvió ambigüedades en la fuente y cosas similares. Una descripción del texto conteniendo información de clasificación y contextual sobre el texto, como es el tema, la situación en el que fue producido, los individuos que participaron en su producción, y cosas similares. Y una historia de revisión, que permite al codificador proporcionar una historia de cambios realizados durante el desarrollo del texto electrónico.

Los lineamientos para el encabezado TEI incluyen la discusión de la relación entre el encabezado y el registro MARC. Los lineamientos establecen que el objetivo de TEI de la creación del encabezado es:

“Asegurar que la información requerida para un registro catalográfico pueda ser obtenido del encabezado del archivo TEI, y más aun que el mapeo de uno a otro sea lo más simple y directo posible”.

(Sperberg-McQueen, 24.1).

Sin embargo, los lineamientos van en el sentido que

“La diferencia más importante entre un registro MARC y el encabezado TEI es la función de cada uno de ellos. Debido a los esfuerzos y reclamaciones de algunos miembros de la comunidad bibliotecaria, los registros MARC permanecen fundamentalmente como una versión electrónica de las tarjetas de catalogación, con las limitantes del modelo. La función primaria del registro MARC cuando fue diseñado en la primera mitad de los 60's fue el permitir la distribución electrónica de registros de catalogación en apoyo a la producción de tarjetas... LA tarjeta de catalogación es un registro unitario de un objeto físico conteniendo datos bibliográficos complejos de diversos tipos. La tarjeta de catalogación caracteriza al objeto físico. El encabezado TEI proporciona información bibliográfica completa (como la tarjeta), además de información documental no bibliográfica que apoya el análisis, tanto humano como de máquina, del texto electrónico documentado por el encabezado. La mayoría de esta información analítica, es encontrada en el perfil de descripción, descripción de codificación y la historia de revisión, tiene escasa provisión directa del registro MARC, y si es mantenida debe ser almacenada como campos de notas sin estructura (campos 5XX). Los campos de notas usualmente no tienen la estructura para soportar el análisis y recuperación automática, mientras que el perfil adecuadamente formateado y codificado y la descripción de revisiones se prestan por sí mismos a la recuperación y al procesamiento a través de máquinas (incluyendo el análisis), y apunta directamente al texto incluido en el encabezado. Más aun, el texto electrónico apunta de regreso a los elementos más relevantes en el encabezado”.

(Sperberg-McQueen, 24.3)

Apoyos para la Localización de Colecciones de Archivos
(<http://sunsite.berkeley.edu/FindingAids/>)

El proyecto Apoyos para la Localización de Colecciones de Archivos es un proyecto de colaboración para desarrollar y codificar los estándares para apoyar la localización en los archivos, museos y bibliotecas. “Los apoyos para la localización son documentos empleados para describir, controlar y proporcionar acceso a colecciones o materiales relacionados. En la estructura jerárquica

de acceso y navegación de la información a nivel de colección, los apoyos para la localización residen entre los registros bibliográficos y la fuente primaria de los materiales. Los registros bibliográficos conducen a los apoyos para la localización, y estos a su vez a las fuentes primarias de los materiales". Los objetivos del proyecto son en dos sentidos. Primero, el crear un estándar prototipo codificado para ayudas para la localización en forma de un DTD de SGML. Segundo, el construir una base de datos prototipo de ayudas para la localización. La base de datos servirá para dos propósitos:

Primero, proporcionar el estándar codificado para los desarrolladores y las aplicaciones computacionales para experimentar y, con lo cual poder refinar e informar el proceso de desarrollo. Segundo, proporcionar un medio para que el usuario final evalúe la utilidad y qué tanto son atractivas las ayudas para la localización ya codificadas, que en su momento, les permita proporcionar nuevas ideas y sugerencias para los desarrolladores de la codificación del estándar.

El SGML fue elegido sobre MARC para la codificación de registros debido a que MARC es limitado en cuanto al acomodo de la información jerárquicamente estructurada. Debido a que las ayudas para la localización son documentos jerárquicamente estructurados, la estructura plana del MARC lo hace no satisfactorio. MARC fue primariamente diseñado para capturar la descripción y acceder la información aplicado a un elemento bibliográfico discreto. Describiendo y proporcionando acceso a colecciones complejas a través de niveles descendentes de análisis, rápidamente sobrecarga la estructura MARC. Es más, un segundo nivel de análisis puede ser acomodado, pero el tipo de información que se proporciona es limitada. Una posible forma de dar vuelta al problema es el emplear múltiples registros jerárquicamente interrelacionados y entrelazados en diferentes niveles de análisis: nivel de colección, subunidad, y ejemplar. Sin embargo, el empleo de múltiples registros, además, introduce una dificultad extrema en el control de problemas inter e intra sistemas que nunca serán resueltos adecuadamente en el formato MARC o por los desarrolladores de aplicaciones basadas en MARC.

Proyecto de Acceso a Imágenes Digitales de la Universidad de Columbia (DIAP)

La iniciativa de cooperación del Grupo de Bibliotecas de Investigación para el acceso a imágenes digitales fue formulada por la Biblioteca de la Universidad de Columbia para sugerir un nuevo modelo para el almacenamiento y acceso a datos bibliográficos y analíticos en imágenes digitales. El equipo DIAP de Columbia propone que los datos pueden ser fácilmente "almacenados" en registros bibliográficos codificados en formato SGML (metadatos) que encapsule información bibliográfica resumida junto con datos jerárquicos detallados y datos relacionados con la versión, donde tales datos sean apropiados y considerados como útiles para el registro. El registro puede además incluir vínculos a los ejemplares digitales actuales, a otros registros bibliográficos relacionados o, en efecto, a diversos objetos digitales relacionados (como son publicaciones electrónicas externas, bases de datos, archivos numéricos, etc.). El trabajo de designación del Registro de Catalogación SGML (SCR) es propuesto para este nuevo tipo de registros. El SCR puede, por flexibilidad incorporar elementos de datos en forma de "constelación", permitiendo un enfoque más narrativo del registro y presentación de información bibliográfica compleja que el actual modelo AACR2/USMARC, que requiere la fragmentación de componentes jerárquicamente relacionados e información de versión en registros discretos separados. El actual modelo USMARC sirve a los usuarios y bibliotecarios de forma muy pobre - más notablemente en la catalogación de reproducciones en microfilms de textos impresos y con publicaciones seriales complejas- y esta situación ha sido ampliamente discutida. Intentos para rectificar la situación han resultado infructuosos, en gran parte debido a la estructura intrínseca plana del USMARC y de los sistemas automáticos de biblioteca que han sido diseñados en torno a ella.

La propuesta de Columbia también sugiere que las bibliotecas “pueden desde el presente e indefinidamente” continuar creando registros MARC de resumen - con vínculos hacia los SCR's - en los sistemas en línea locales y en las utilerías nacionales. (<http://www.cc.columbia.edu/cu/libraries/indiv/avery/diap.html>)

¿Qué significa esto?

¿Cuál es la relevancia de los proyectos descritos anteriormente?. Primero y más importante, todos los proyectos emplean SGML para estructurar y dar acceso a información bibliográfica jerárquica, complicada y analítica. Aunque todos los proyectos reconocen la utilidad de los registros MARC como un apoyo, también exponen fallas críticas en el formato MARC: los lineamientos TEI resaltan problemas con la incapacidad de MARC para representar estructuras analíticas, información no bibliográfica que puede ser empleada para evaluar documentos electrónicos; el Proyecto de Berkeley de Apoyo a la Localización comenta la incapacidad de MARC para funcionar con información jerárquica estructurada y para proporcionar acceso a colecciones complejas a través de niveles de análisis descendentes. El equipo DIAP de Columbia reconoce ambos problemas y adiciona aspectos sobre el manejo de las versiones de la información. Todos los proyectos lidian con recursos jerárquicamente estructurados: textos etiquetados en formato TEI van del encabezado bibliográfico a través de la información analítica, hasta el texto por sí mismo; en el proyecto Apoyo a la Localización, las ayudas son altamente jerárquicas, y son colocadas entre la información bibliográfica (registros de catalogación) y las colecciones actualmente archivadas o sus vicarios. Las imágenes digitales pueden ir de un objeto original, a través de etapas intermedias (diapositivas, impresiones, etc.) a una versión puramente digital de las imágenes. En el segundo Reporte CFCC, Michel Buckland discute los objetivos de los catálogos bibliográficos (como un conjunto externo) y la relevancia de estos objetivos para diseñar los sistemas en línea del futuro:

El enfoque de camino individual en los catálogos está aun dividido en decenas de millones de ejemplares que están siendo accedidos por la red y cómo los efectos del vincular bibliografías en línea a registros catalográficos está extendiendo el potencial bibliográfico de los catálogos más allá de los sueños de los compiladores de códigos catalográficos. Para hacer frente a la clara escala del primer objetivo (conocer si un trabajo en particular está “dentro” de una biblioteca) en un ambiente de red de una biblioteca electrónica y para atender el segundo objetivo (conocer qué trabajos están “dentro” de una biblioteca) en cualquier catálogo, un enfoque jerárquico es necesario para que los usuarios puedan fácilmente desplazarse al nivel de detalle que necesitan: trabajos, versiones variantes de trabajos; parte del trabajo; el texto completo del trabajo; tipos de trabajos relacionados, etc. Desde un enfoque de un solo camino no será útil, ya que los catálogos futuros deberán ser multicaminos, flexibles y adaptables en su operación.

(Buckland, p. C)

¿Cómo se deben realizar los futuros catálogos multicaminos, flexibles y adaptables?. MARC es claramente incapaz de proporcionar la flexibilidad, adaptabilidad operativa que los sistemas bibliotecarios del futuro requerirán. ¿Significa esto entonces que el formato MARC ya no es útil y que desaparecerá en dos, cinco o diez años?. Por supuesto que no. Hay billones y billones de registros MARC en sistemas bibliográficos en línea y en utilerías nacionales; el costo, en tiempo y recursos de cómputo solamente, para convertirlos a SGML pueden ser exagerados. Esto significa que MARC como el mundo bibliográfico lo conoce debe evolucionar para permitir a los bibliotecarios hacer el uso más efectivo de sus recursos aún muy limitados. Esta evolución bien puede significar que los registros MARC no serán más el formato individual para la codificación de datos bibliográficos. El HTML (Hypertext Markup Language) puede resultar familiar para la

mayoría de los bibliotecarios como un lenguaje de etiquetas, desde 1990 con la aparición del WWW (World Wide Web). HTML es una aplicación SGML, completa con DTD y numerosas especificaciones. El HTML es un conjunto de etiquetas descriptivas que es interpretada por procedimientos por diversos navegadores de WWW como Netscape o Mosaic (pensando en el desarrollo, los navegadores WWW apegados a SGML no son totalmente operacionales). HTML permite a los usuarios insertar imágenes, sonidos, textos y vídeo en los documentos, y proporcionar vínculos de hipertexto de un punto en un documento a otro, o separar un documento, un archivo de imagen o sonido. Existen sin embargo limitaciones para el HTML; las etiquetas son esencialmente muy simples y sólo definen la estructura de un documento a nivel muy básico. Los catálogos bibliográficos han dado un paso adelante en el WWW. Sistemas como SIRSI's WebCat (<http://www.sirsi.com/webcattoc.html>) y Data Research Associates DRAWeb (<http://www.dra.com/products/draweb/draweb.HTM>) ofrecen interfaces WWW para catálogos bibliográficos. Estos catálogos basados en WWW ofrecen clientes WWW que pueden acceder y buscar en bases de datos MARC, crear HTML dinámicamente, y proporcionar resultados al usuario en su computadora de escritorio. Algunos clientes cumplen con el estándar Z39.50.

El SGML, en combinación con otros desarrollos, ofrece soluciones adicionales. SGML es independiente de la aplicación, no propietario, y extremadamente flexible; como tal, ofrece una alternativa viable y/o conjunta para la codificación de información bibliográfica. Como los proyectos mencionados anteriormente demuestran, es posible codificar datos en un formato SGML. SGML y MARC son aún formatos separados y que no interactúan; MARC, como actualmente existe no parece ser lo suficientemente flexible para permitir a los bibliotecarios tomar ventaja total de la tecnología de recuperación de información siempre en desarrollo, especialmente en el World Wide Web. Un DTD USMARC, sin embargo, está siendo desarrollado por Jerome McDonough de la Universidad de California, Berkeley y está a disposición por FTP anónimo en <ftp://library.berkeley.edu/pub/sgml/marcdtd/>. El DTD "está diseñado para emplearse en catálogos en línea utilizando SGML como formato de registro oculto" y tiene el propósito de "identificar (por medio de etiquetas) elementos de soporte de contenido bajando al nivel de subcampos... permitiendo la conversión automática de registros USMARC al formato SGML y en sentido inverso con la menor pérdida posible del contenido original." La capacidad de transportar datos bibliográficos de SGML a MARC y en sentido inverso es similar al primer paso hacia el desarrollo de los catálogos basados en WWW, con SGML como su estructura de registros. En tales sistemas, los clientes pueden, y deben cumplir con el estándar Z39.50 y estar apegados al SGML; los clientes pueden tener la posibilidad de tomar total ventaja de las etiquetas SGML más robustas y pueden, por ejemplo, ofrecer búsquedas en los contenidos basadas en sus elementos estructurales. Un catálogo basado en SGML puede permitir desarrollos adicionales a los proyectos mencionados anteriormente, por ejemplo, una versión de textos de acuerdo a TEI puede ya no necesitar registros bibliográficos MARC para describirlos; su encabezado puede servir para los paquetes de información descriptiva y puede ser formateado y mostrado en cualquier forma que seleccione el propio usuario. Las búsquedas jerárquicas pueden considerar primero la información descriptiva de los paquetes (como los encabezados), y llevarlos a la computadora del usuario mientras mantienen un vínculo con el texto completo, la imagen o el formato digital, que también puede ser buscado, mostrado, salvado y manipulado como el usuario prefiera, y que, por supuesto, puede ser vinculado con otras versiones, imágenes o archivos analíticos.

"Sin la gran flexibilidad en la catalogación y codificación de documentos digitales, los datos bibliográficos generados por la biblioteca no serán integrados fácilmente en el ambiente de información nacional y local como inventarios efectivos de índices para los repositorios electrónicos de las bibliotecas".

(Davis)

Por todas las especificaciones antes mencionadas, para el proceso de indización con el software de Z39-50 para ofrecer la opción de búsqueda en el contenido de los títulos de los artículos de las publicaciones seriales, se empleará como entrada de información las tablas de contenido de las publicaciones dispuestas con etiquetas tipo SGML. En la etapa de desarrollo del sistema se especificará el formato de estos datos.

CAPÍTULO 5

CAPÍTULO 5

DESARROLLO DEL SISTEMA

En el análisis, una vez que se han establecido los requerimientos de sistema, la fase de desarrollo comprende tres pasos distintos:

- Diseño
- Generación de código
- Prueba

Cada paso transforma la información de forma que finalmente se obtiene el software.

A continuación se muestra cada una de estas etapas, tomando como base toda la información que se tiene del análisis del sistema.

El **diseño** es un proceso mediante el que se traducen los requerimientos en una representación de software. Dentro de esta etapa se realiza:

- a) Diseño de datos.
- b) Diseño arquitectónico y procedural.

El diseño de datos se enfoca sobre la definición de datos. El diseño arquitectónico define las relaciones entre los principales elementos estructurales del programa y define la estructura de datos y la parte del diseño procedural transforma los elementos estructurales en una descripción procedural del software.

5.1. DISEÑO DE DATOS

Las actividades primarias durante el diseño de datos son el seleccionar la representación de los datos, es decir definir como inicialmente se identificará la información.

5.1.1. Identificación de los nombres de las publicaciones seriales de la biblioteca de DGSCA.

Se cuenta con 151 títulos distintos de publicaciones seriales de la biblioteca de DGSCA, de las cuales 131 cuentan con un ISSN asignado en forma oficial (86.75 % de nuestro universo), mientras que las restantes 20 revistas no cuentan con la etiqueta asignada (13.25 %). A continuación se muestra una tabla con estos resultados.

Titulo	ISSN
ACM Communications	0001-0782
ACM Computing Surveys	0360-0300
ACM Journal	
ACM Transactions on Database Systems	0362-5915
ACM Transactions on Mathematical Software	0098-3500
American Journal of Distance Education	0892-3647
Applied Mathematics and Computation	0096-3003
Artificial Intelligence	0004-3702
Basica Revista de la Escuela y el Maestro	
Boardwatch Magazine	1054-2760
Business Week	0007-7135
Byte	0360-5280
Byte Mexico	
Cadence	0887-9141
CD-ROM Professional	1090-946X
Ciencia y Desarrollo	0185-0008
Clipper Plus	
Communications Info Disk CD-ROM	
Communications of the ACM	0001-0782
Computer	0018-9162
Computer Aided Design	0010-4485
Computer and Control Abstracts	0036-8113
Computer Architecture News	0163-5964
Computer Communications Review	0146-4833
Computer Graphics	0097-8930
Computer Graphics and Image Processing	0146-664X
Computer Graphics World	0271-4159
Computer Journal	0010-4620
Computer Languages	0096-0551
Computer Library	
Computer Networks and ISDN Systems	0376-7552
Computer Physics Communications	0010-4655
Computer Pictures	0883-5683
Computer Technology Review	0278-9647
Computer Vision and Image Understanding	1049-9660
Computer Vision Graphics and Image Processing	1049-9652
Computers and Education	0360-1315
Computers and Graphics	0097-8493
Computers and Mathematics with Applications	0097-4943
Computers and Operations Research	0305-0548
Computers and People	0361-1442
Computers and Structures	0045-7949
Computers and the Humanities	0010-4817
Computerworld	0010-4841
ComputerWorld (Client/Server Journal)	
Computerworld Mexico	0188-9370
Computing Reviews	0010-4884
Connexions	0894-5926
CVGIP : Graphical Models and Image Processing	0734-189X
CVGIP: Image Understanding	1077-3142

Cybernetics and Systems	0196-9772
Data Base	0095-0033
Data Communications	0036-6399
Data Reports Networks CD-ROM	
Database Programming and Design	0895-4518
Datamation	0011-6963
DBMS	1041-5173
Design Net	0895-4151
Dr. Dobb's Journal	
E-Media Professional	1049-0833
Electronic Business	0163-6197
Electronic Business Today	0163-6197
Electronic Design	0013-4872
Electronics	0883-4989
Electronics & Communications on Engineering Journal	0954-0695
Graphical Models and Image Processing	1049-9652
Gregory's A-Series Technical Journal	0892-4856
I/S Analyzer	0896-3231
IBM Journal of Research and Development	0018-8646
IBM Systems Journal	0018-8670
IEEE Annals of the History of Computing	0164-1239
IEEE Computer	0018-9162
IEEE Computer Graphics and Applications	0272-1716
IEEE Transactions on Circuits and Systems	0098-4094
IEEE Transactions on Computers	0018-9340
IEEE Transactions on Software Engineering	0098-5589
IEEE Transactions on Systems, Man And Cybernetics	0018-9472
Informacion Cientifica y Tecnologica	0187-6015
Informacion, Produccion, Comunicacion y Servicios	0188-5847
Information and Computation	0890-5401
Information and Control	0019-9958
Information Systems	0306-4379
International Journal of Computer and Information Science	0091-7036
International Journal of Parallel Programming	0885-7458
Internet Business Journal	
Internet World	1064-3923
Journal of Computer and System Sciences	0022-0000
Journal of Object-Oriented Programming	0896-8438
Journal of Parallel and Distributed Computing	0743-7315
Journal of Robotic Systems	0741-2223
Journal of Supercomputing	0920-8542
Journal of Systems and Software	0164-1212
Journal of Systems Management	0022-4839
Journal of The ACM	0004-5411
Journal of the Association for Computing Machinery (ACM Journal)	0004-5411
LAN: The Network Solutions Magazine	1069-5621
Local Area Networks	1051-1962
Mathematica Journal	1047-5974
Mexico en el Diseno	
Microcad News (MCN)	0895-4151

Microsoft Systems Journal (MSJ)	0889-9932
Microstation Manager	1057-9567
MPC World	
Multimedia World	1073-4759
Network Computing	1046-4468
Network World	0887-7661
Networks: an International Journal	0028-3045
Object Programmer	
OOPS Messenger	
Open Computing	1072-4044
Open Learning	0268-0513
Operating Systems Review	0163-5980
PC Computing	0899-1847
PC Magazine	0745-2500
PC Magazine en Espanol	0188-2732
PC Techniques	1053-6205
PC Tips	
PC Week	0740-1604
PC World	0737-8939
Personal Computing Mexico	0000-9339
Proceedings of the IEEE	
Quality Progress	0033-524X
Red: la Revista de Redes de Computadoras	0000-9238
Research and Education Networking	1051-4791
Revista Iberoamericana de Educacion Superior a Distancia	0214-3992
Scientific American	0036-8733
Scientific programming	1058-9244
SIAM Journal on Computing	0097-5397
SIAM Journal on Control and Optimization	0036-1410
SIGARCH: Computer Architecture News	0163-5964
SIGCOMM: Computer Communications Review	0146-4833
SIGGRAPH Computer Graphics	
Sigmetrics Performance Evaluation Review	0734-2071
Sigplan Notices	0362-1340
Siguuccs Newsletter	0163-5832
Simulation	0037-5497
Software Magazine	0897-8085
Software Practice and Experience	0038-0644
Software World	0038-0652
Soluciones Avanzadas	0188-8048
Sun Expert Magazine	1053-9239
Sun World	
Tecnologia y Comunicacion Educativas	0187-0785
Technology and Learning	1053-6728
Ulrich's Plus	
Unix Review	0742-3136
Unix World	0739-5922
Visual Developer	1053-6205
Windows Magazine	1060-1066
WordPerfect for Windows	
WordPerfect Magazine	1042-5152

Para aquellos casos en los que no hay un ISSN asignado, se decidió asignar un código arbitrario siguiendo el formato del ISSN para que en un momento dado, si alguna de estas publicaciones obtiene su registro, el cambio requerido no sea mayor. De esta manera, tenemos que los códigos arbitrarios fueron:

Título	Código
ACM Journal	XXXX-0001
Básica Revista de la Escuela y el Maestro	XXXX-0002
Byte Mexico	XXXX-0003
Clipper Plus	XXXX-0004
Communications Info Disk CD-ROM	XXXX-0005
Computer Library	XXXX-0006
ComputerWorld (Client/Server Journal)	XXXX-0007
Data Reports Networks CD-ROM	XXXX-0008
Dr. Dobb's Journal	XXXX-0009
Internet Business Journal	XXXX-0010
México en el Diseño	XXXX-0011
MPC World	XXXX-0012
Object Programmer	XXXX-0013
OOPS Messenger	XXXX-0014
PC Tips	XXXX-0015
Proceedings of the IEEE	XXXX-0016
SIGGRAPH Computer Graphics	XXXX-0017
Sun World	XXXX-0018
Ulrich's Plus	XXXX-0019
WordPerfect for Windows	XXXX-0020

Es claro que no existirá un código ISSN válido que concuerde con los códigos arbitrarios asignados, por lo que no hay riesgo de emplear un código oficial y la identificación de las publicaciones sin ISSN oficial es más sencilla.

5.1.2. Formato de las tablas de contenido de las revistas

Dentro del análisis de requerimientos del sistema se planteo el análisis de información, éste dio como resultado la necesidad de definir el formato de la fuente principal de entrada de datos al sistema, es decir, de las tablas de contenido de las revistas.

La biblioteca puede producir sus tablas de contenido o índices en texto plano realizando la digitalización y aplicando el proceso de OCR a las revistas, el paso que se debe definir es la forma en que ellos entreguen esta información al encargado del sistema.

Elementos de cada una de las tablas de contenido

Analizando la información de la tabla de contenido de cada una de las revistas se lograron detectar los siguientes elementos claves para su identificación única:

- a) Nombre de la revista

- b) Volumen
- c) Mes
- d) Año
- e) Nombre del artículo
- f) Resumen del artículo

Formato de las tablas de contenido entregadas por la biblioteca

Una vez detectados los elementos de las tablas de contenido, se solicitó al personal de la biblioteca que las proporcionará a la entrada del sistema en el formato que a continuación se muestra, la disposición de los elementos de la tabla de contenido son distintivos de las revistas y la cadena "@@@@@" indica la separación entre los datos de la revista y el índice de la misma.

NOMBRE DE LA REVISTA		
VOL. # NO. #		
MES AÑO		
@@@@@		
	<Línea en blanco>	
NOMBRE DEL ARTICULO		
Resumen del artículo		No. página de inicio
	<Línea en blanco>	
NOMBRE DEL ARTICULO		
Resumen del artículo		No. página de inicio
	<Línea en blanco>	
NOMBRE DEL ARTICULO		
Resumen del artículo		No. página de inicio

Ejemplo:

RED LA REVISTA DE REDES DE COMPUTADORAS
VOL.7 NO.84
SEPTIEMBRE 1997
@@@@@

CONCEPTOS BASICOS

Tecnología Gigabit Ethernet: Redes de alto
rendimiento con mayor velocidad

Una tecnología emergente que puede satisfacer las
necesidades creadas por las aplicaciones recientes
bajo redes locales. A esta nueva tecnología se le
conoce como Gigabit Ethernet

Jesús Gutiérrez 14

SECCION ESPECIAL

En materia regulatoria ya no hay modelo
a seguir: Casaus

-Educación y telecomunicaciones: la combinación perfecta

-La Cofetel no frena la apertura, al contrario, la promueve

-Hemos alcanzado los problemas de los países desarrollados
en materia de telecomunicaciones: Casaus

Yolanda Aldaco 14

SU PLEMENTO PUBLICITARIO

Intersys 22

INTERNET

IP Versión 6: El soporte de la siguiente generación El acelerado
crecimiento de Internet y el surgimiento de nuevas aplicaciones
requieren de un protocolo que sea capaz de enfrentar las grandes
demandas de las redes en los próximos años. La nueva versión del
protocolo IP busca resolver estos problemas

Hector Acevedo 24

INTERNET

El computo en la misión Pathfinder: la
luna quedo atrás

Rafael Fernández 32

INDUSTRIA 42

EJECUTIVOS

El servicio como vocación empresarial

Gustavo Guerrero 60

5.2. DISEÑO ARQUITECTÓNICO Y PROCEDURAL

Esta fase incluye dos elementos importantes:

- a) Estructura del programa: organización jerárquica de los componentes procedurales (módulos) del programa e implica una jerarquía de control, y
- b) Estructura de datos: representación detallada de los datos, organización, métodos de acceso, y grado de asociatividad.

El objetivo principal del diseño arquitectónico es desarrollar una estructura de programa modular y representar las relaciones de control entre los módulos. Además, el diseño arquitectónico mezcla la estructura de programas y la estructura de datos y define las interfaces que facilitan el flujo de los datos a lo largo del programa.

5.2.1. Estructura del programa

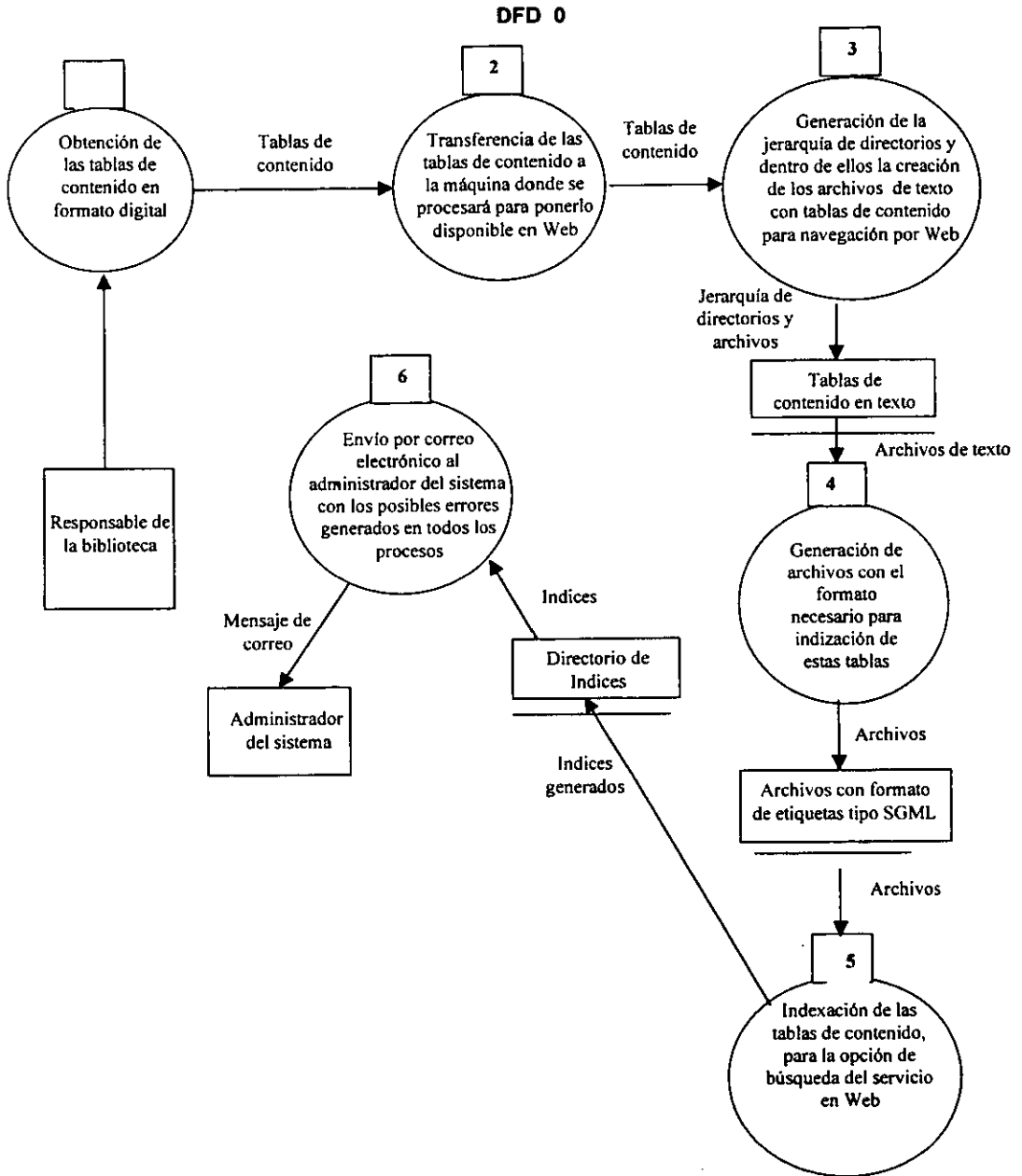
A continuación se muestran la organización detallada de los 4 módulos en los que se segmenta el sistema:

1. Incorporación y actualización de información al acervo de revistas
2. Publicación en Web del acervo de revistas
3. Búsqueda en Web de información en el acervo de revistas
4. Envío de artículos a los usuarios finales

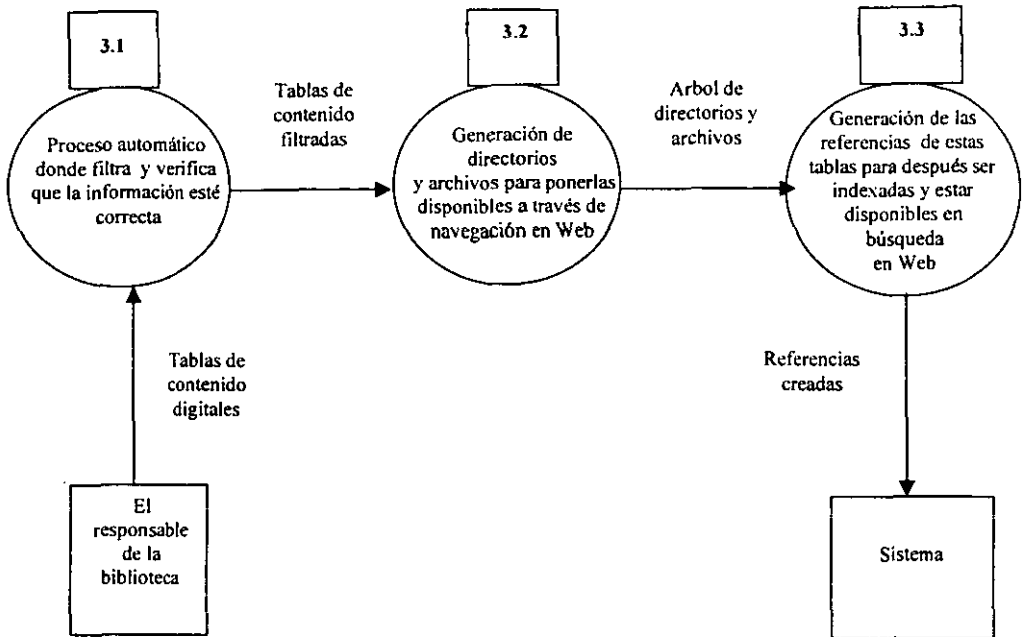
Para ello nos auxiliaremos de la metodología de Myers y Yourdon, empleando DFD (diagramas de flujo de datos). En los diferentes DFD, dentro de cada módulo y submódulo se hace la descripción procedural de cada uno de ellos.

5.2.1.1. Diagramas de Flujos de datos

Incorporación y actualización de información al acervo de revistas

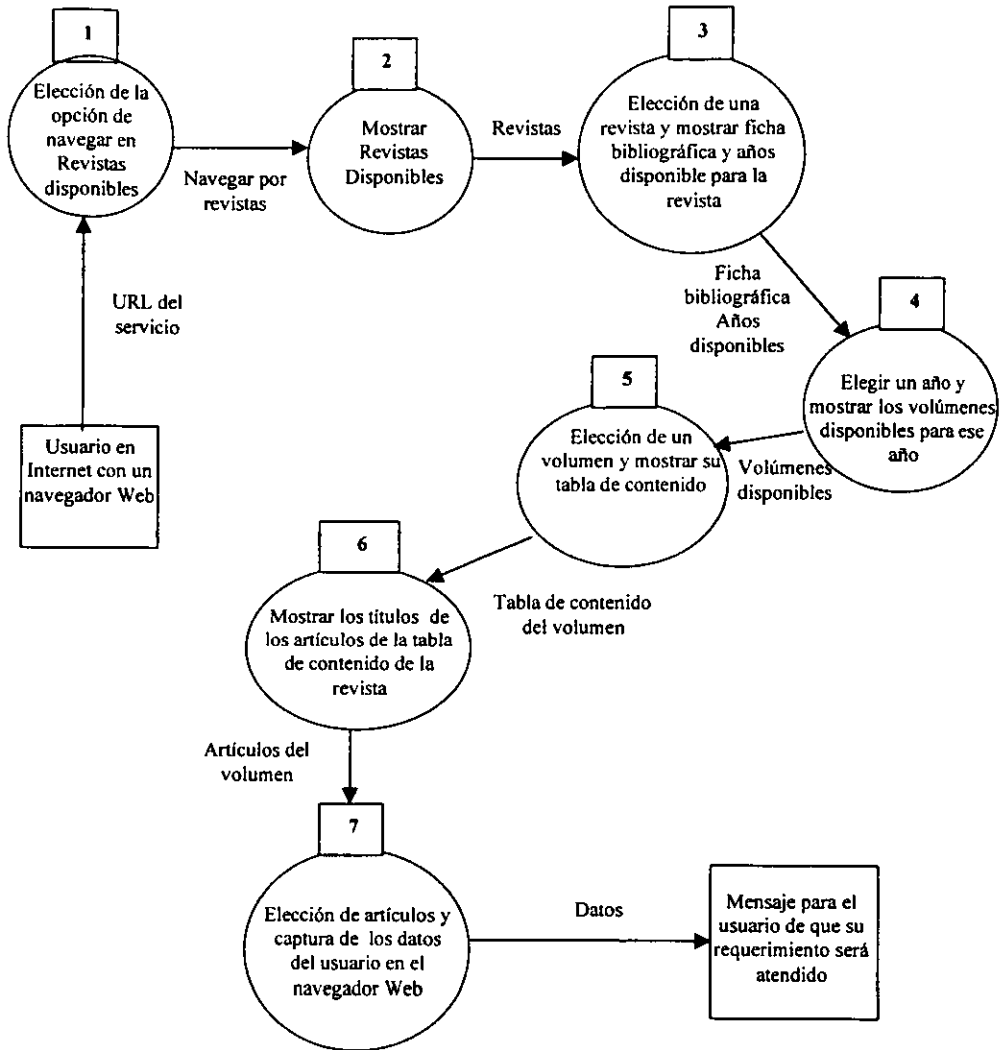


DFD 3 Generación de la jerarquía de directorios para las tablas de contenido

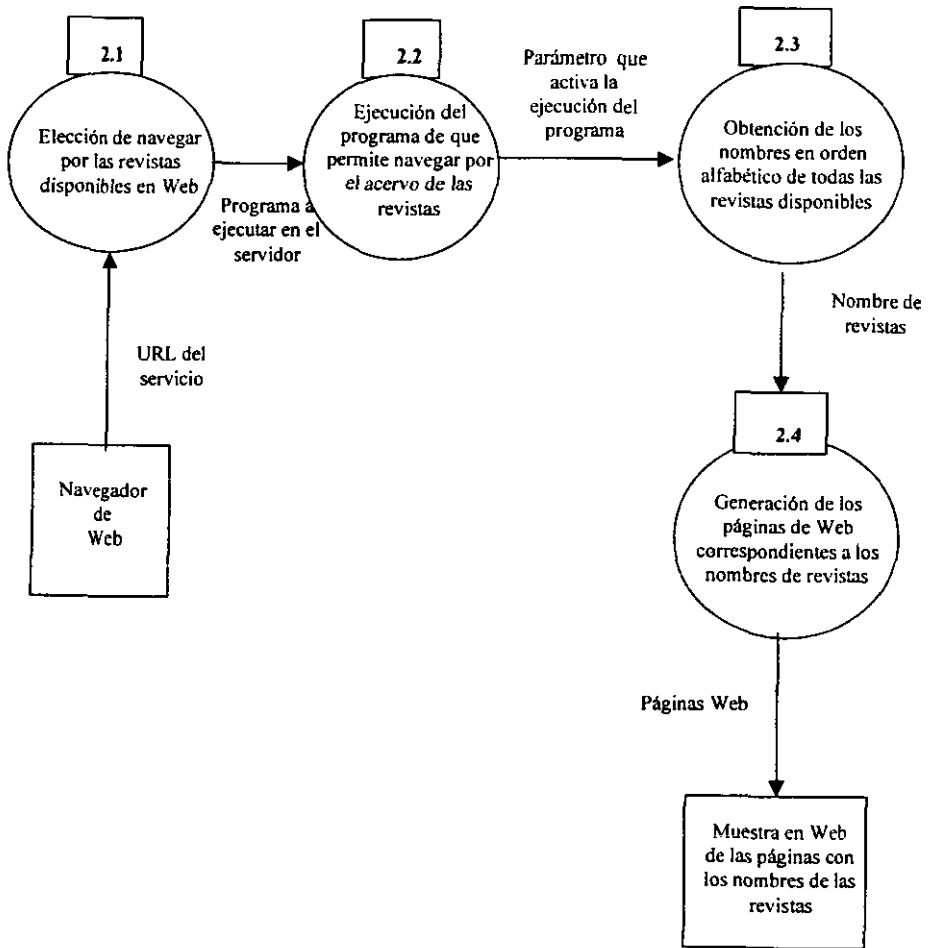


Publicación en Web del acervo de revistas

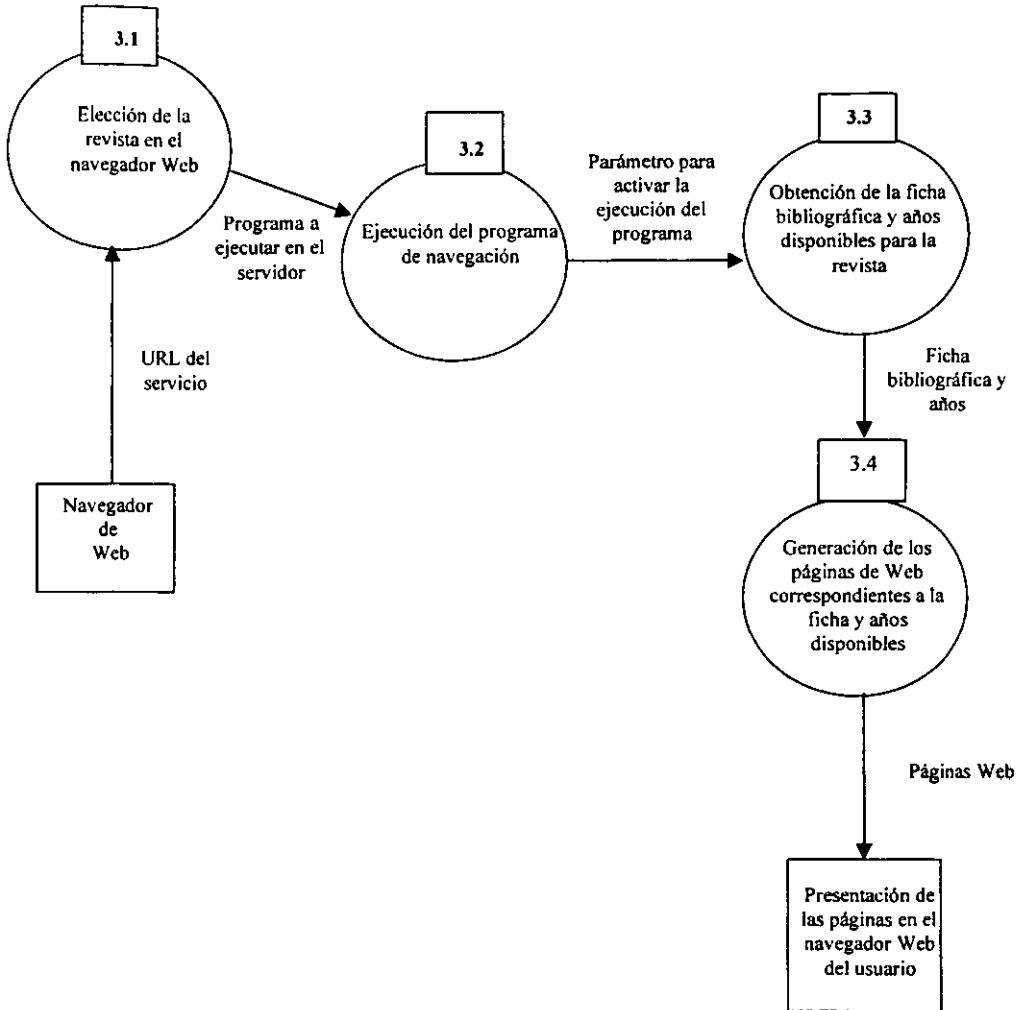
DFD 0



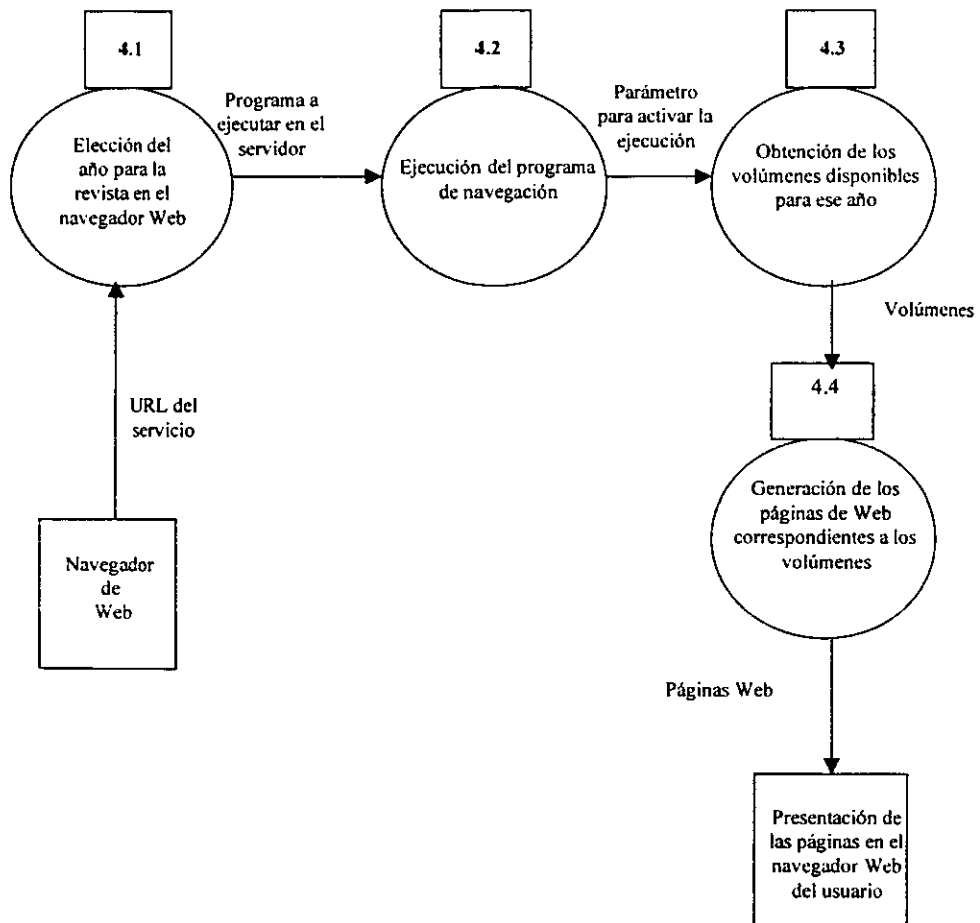
DFD 2 Muestra de las revistas disponibles



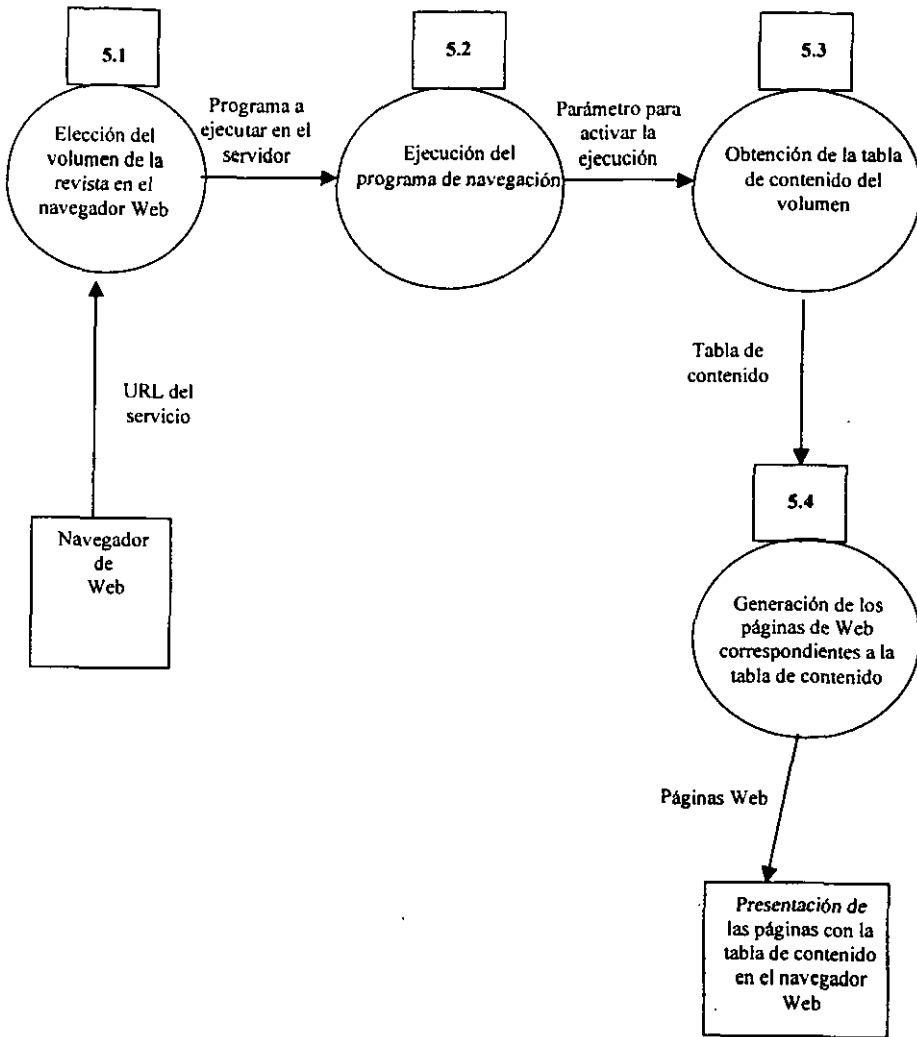
DFD 3 Muestra de ficha bibliográfica y años disponibles de la revista elegida



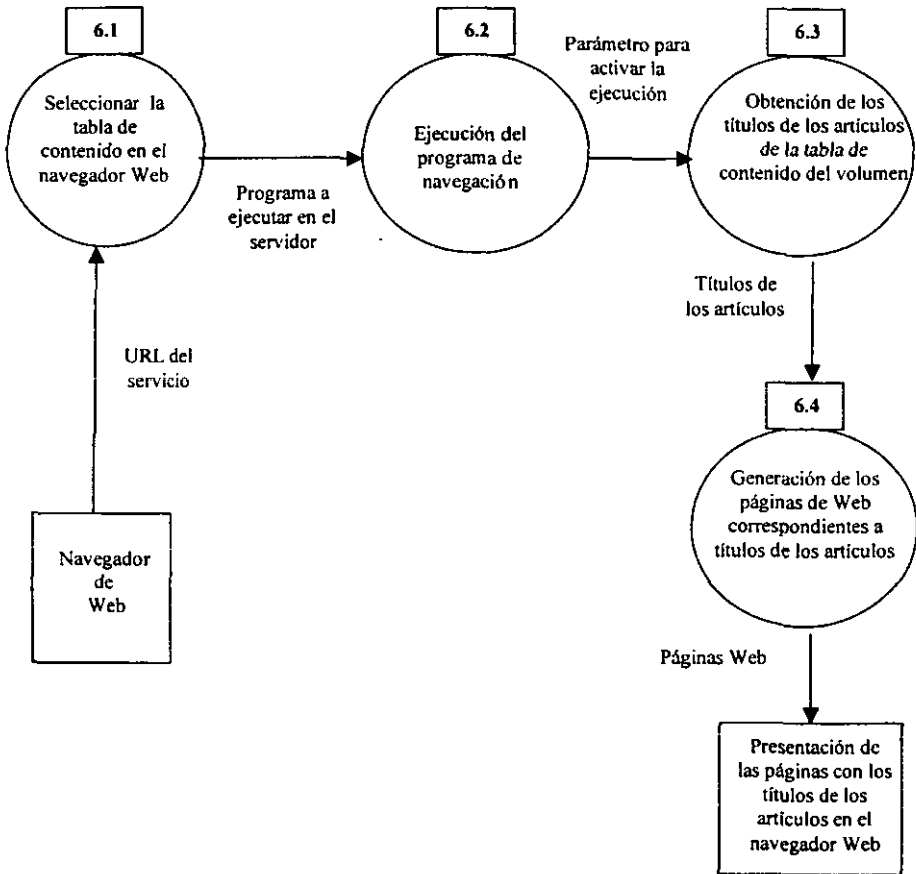
DFD 4 Mostrar los volúmenes disponibles para el año elegido



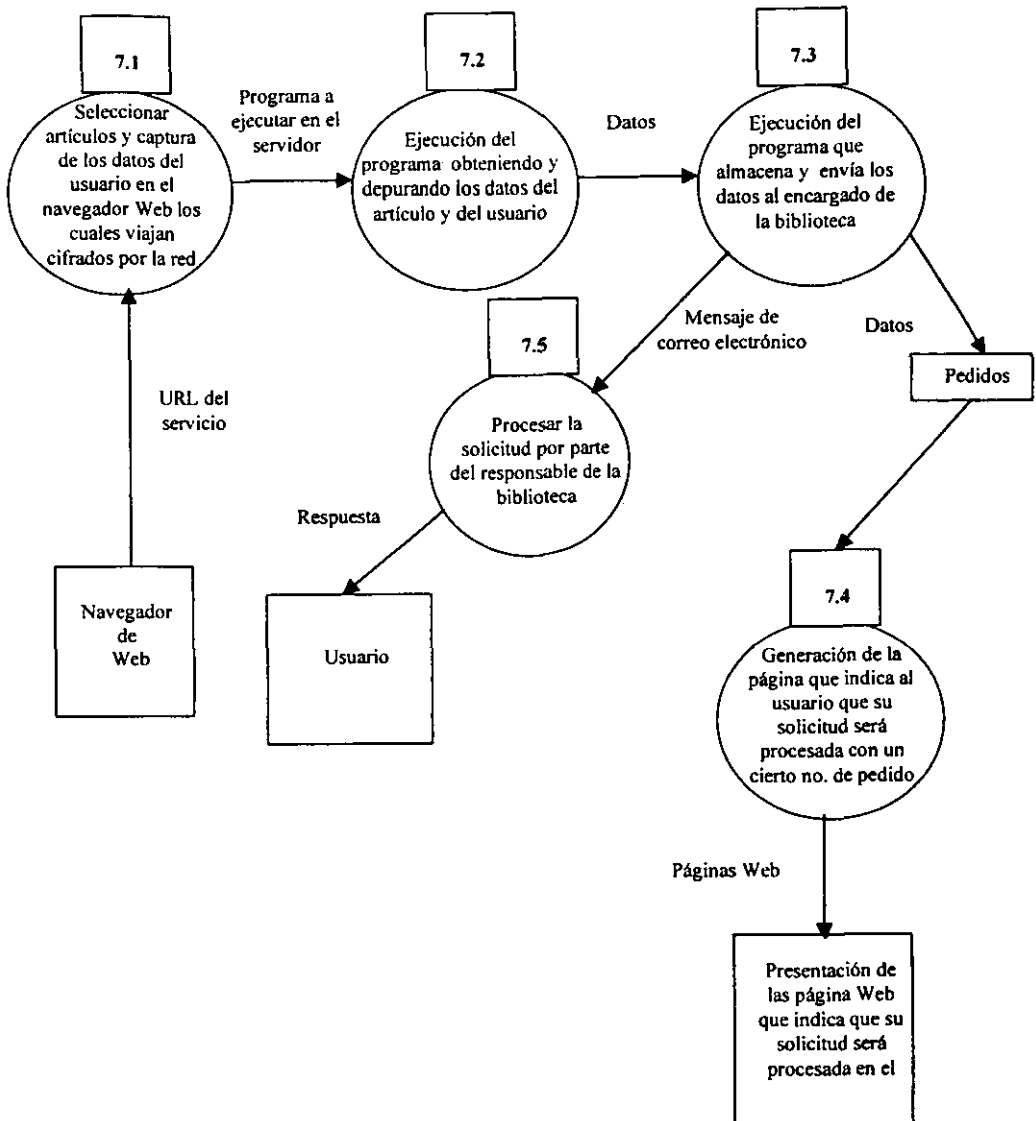
DFD 5 Mostrar la tabla de contenido del volumen elegido



DFD 6 Mostrar los títulos de los artículos de la tabla de contenido del volumen elegido

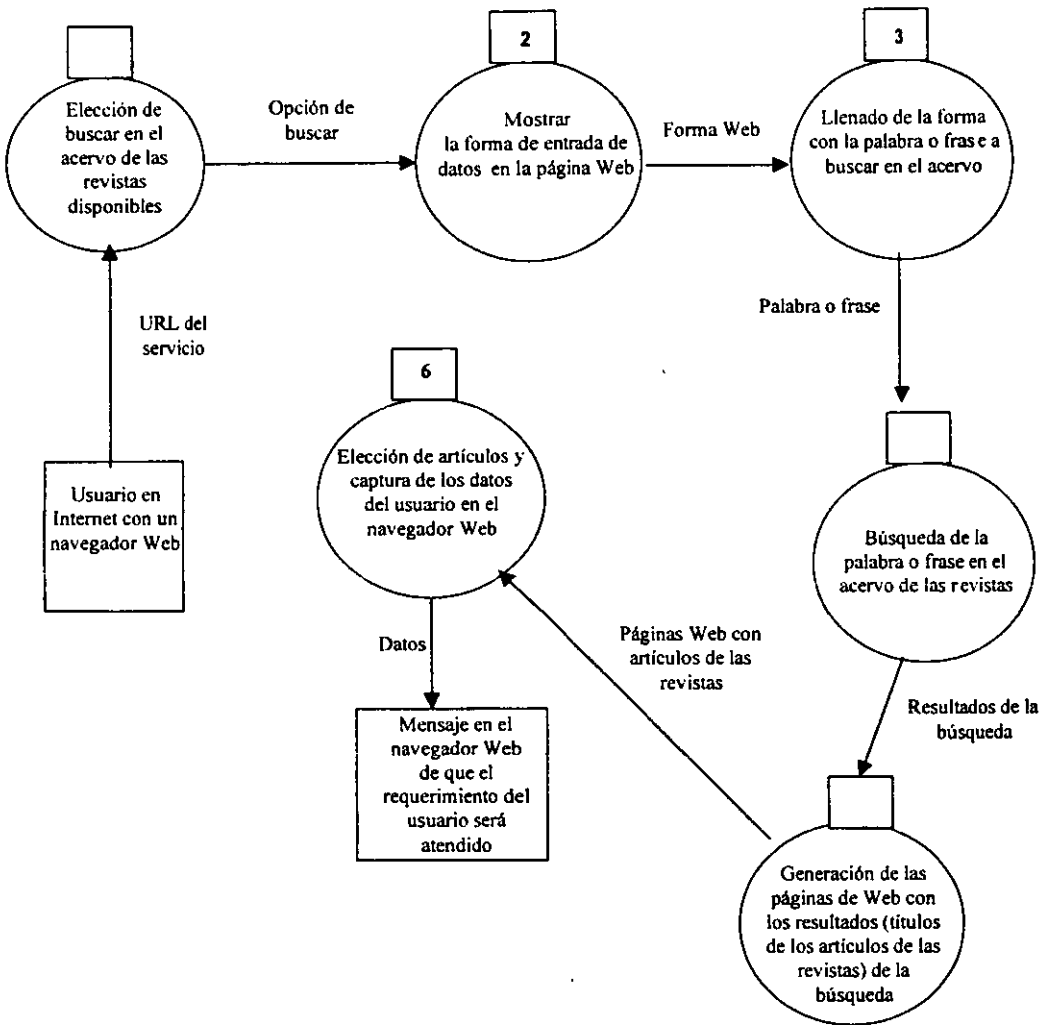


DFD 7 Elección de los artículos del volumen de la revista

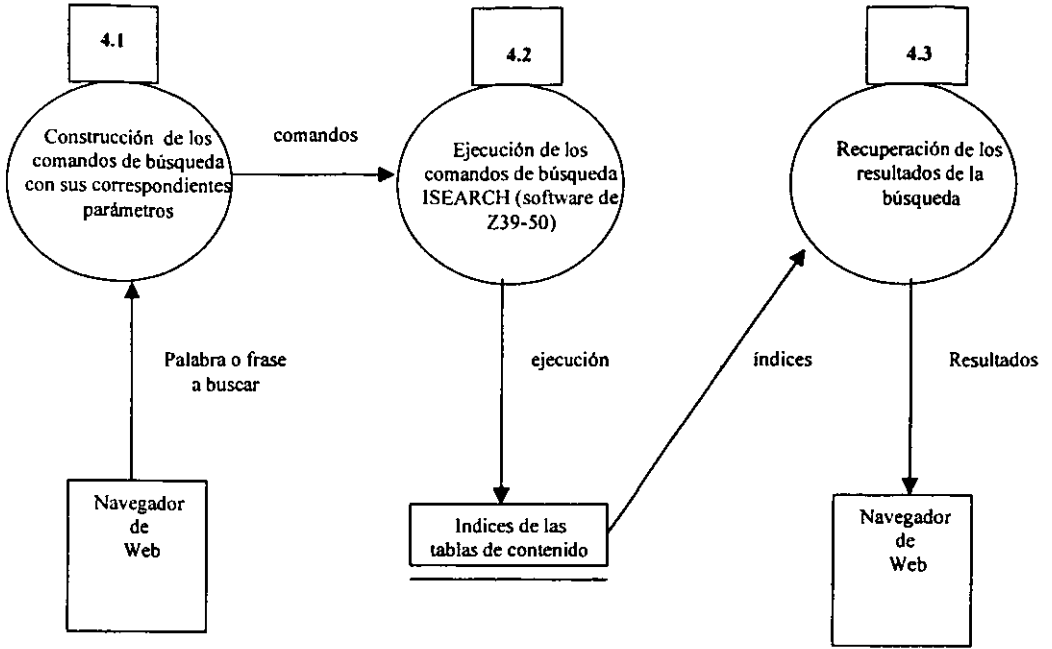


Proceso de Búsqueda en Web de información en el acervo de revistas

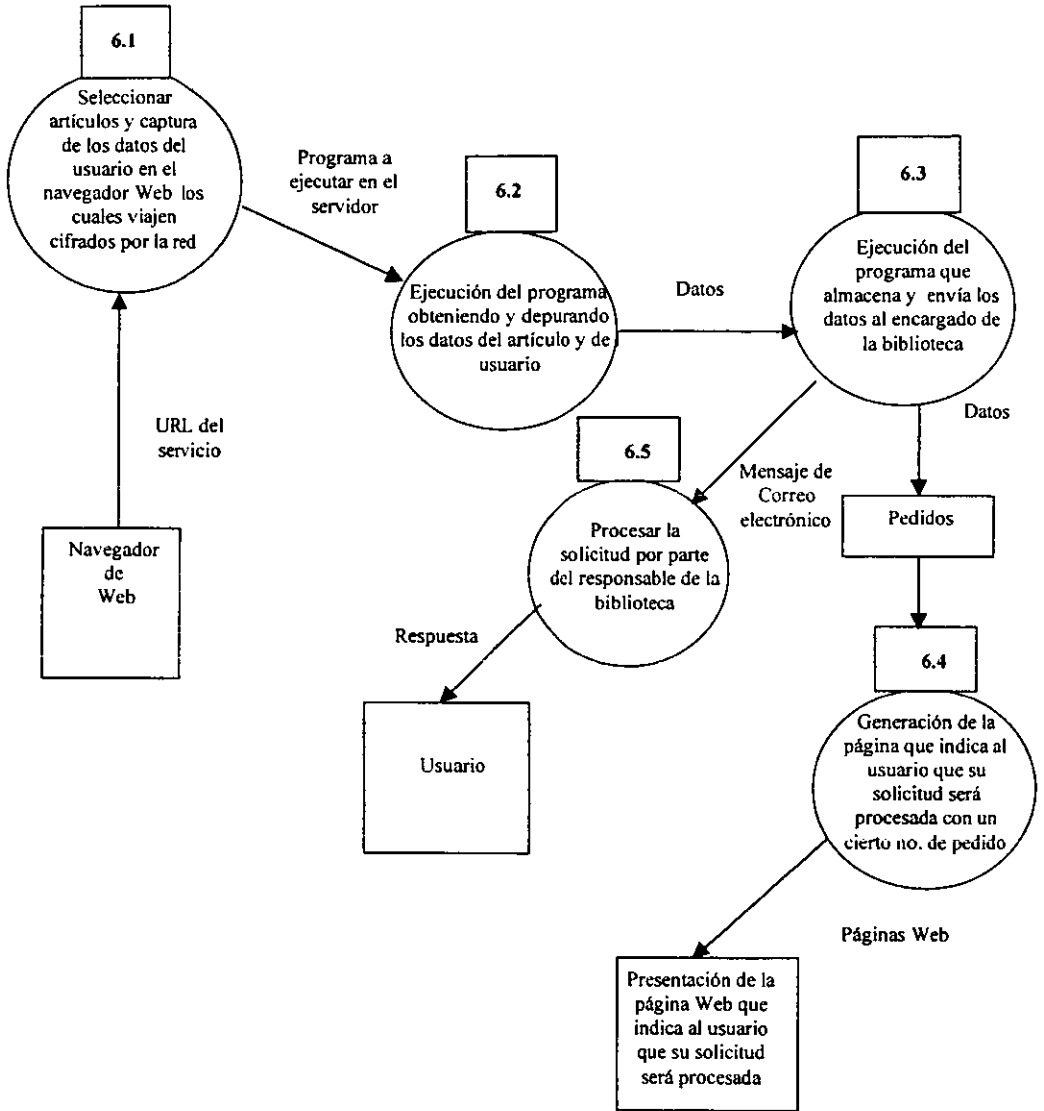
DFD 0



DFD 4 Búsqueda de la palabra o frase en el acervo

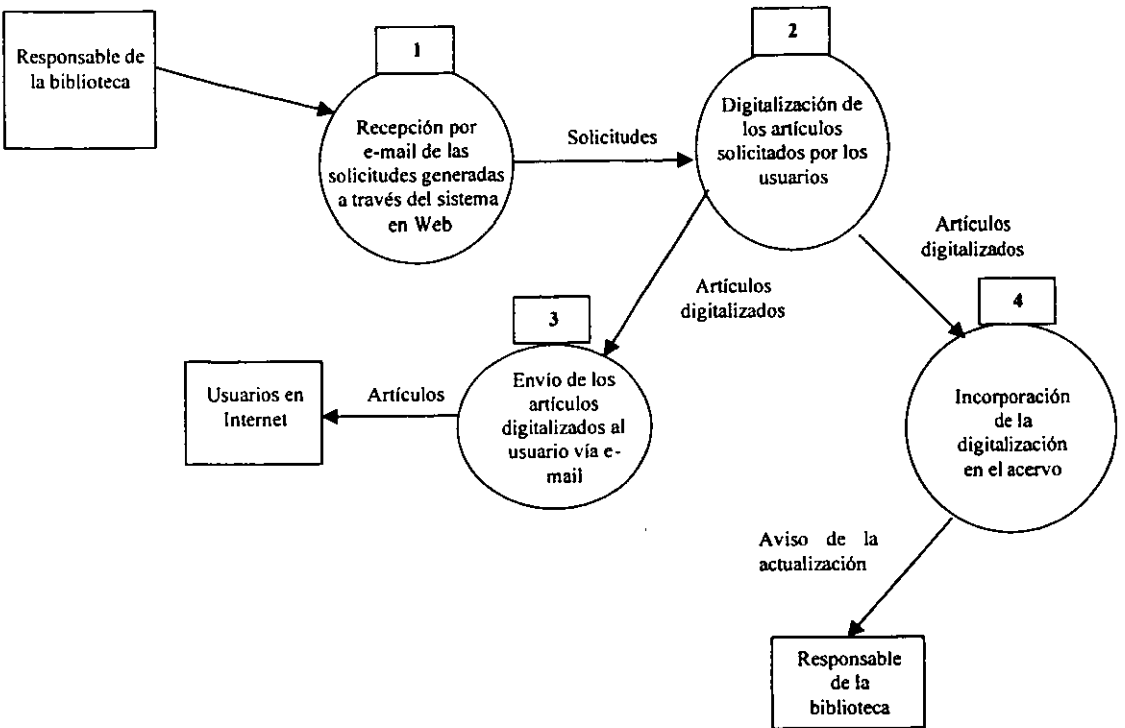


DFD 6 Elección de los artículos del volumen de la revista



Envío de los artículos a los usuarios finales

DFD 0



5.2.2. Estructura de datos

Como se había comentado con anterioridad la estructura de datos es la representación detallada de los datos, su organización y grado de asociatividad, todo esto dentro del sistema; desde la etapa de análisis se vio con que información se trabajaría, en esta fase de diseño es cuando se debe definir cómo el sistema verá a los datos.

5.2.2.1. Diccionario de datos

El diccionario de datos se creo para cada uno de los módulos de qué consta el sistema.

Diccionario de datos para el proceso de Incorporación y actualización de información al acervo de revistas

Equipo donde el encargado de la Biblioteca coloca las tablas de contenido en texto

= calmecac.dgsca.unam.mx

Directorio y archivos donde se colocan las tablas de contenido

= /directorio_base/archivos.txt
** Ejemplo: /var/tmp/textos/JUN-98.TXT

Jerarquía de directorios que se crea con las nuevas tablas de contenido

= /direcrotio_base + clave_revista + año + Volúmen-Número
*** Ejemplo: /export/sky/uno/revdgsca/datos/indices/0097-8930/1997/0017-0002

clave_revista

= El ISSN de la revista

Archivo en texto que se genera con las nuevas tablas de contenido

= @ /directorio_base + clave_revista
+ año + Volúmen-Número + archivo de texto
*** Ejemplo: /export/sky/uno/revdgsca/datos/indices/0097-8930/1997/0017-0002/tc.txt

Comando del software Z39.50 para hacer la indización.

= lindex

Rutas de los directorios donde se crearán los archivos con formato de etiquetas SGML para indexar las tablas de contenido por el software de Z39.50

= /directorio_base + clave_revista + año + Volúmen -Número
*** Ejemplo: /export/sky/uno/revdgsca/datos/indices/0097-8930/1997/0017-0002

Archivo que se genera con formato de etiquetas tipo SGML para indexar las nuevas tabas de contenido para la opción de búsqueda.

= @/directorio_base + clave_revista + año + Volúmen-Número + archivo.bdf
*** Ejemplo: /export/sky/uno/revdgsca/datos/indices/0097-8930/1997/0017-0002/tc.dbf

Rutas de los directorios donde se crearán los índices de las nuevas tablas de contenido empleados por el software de Z39.50

= /directorio_base
*** Ejemplo: /export/sky/uno/revdgsca/datos/.indices

Índices de las nuevas tablas de contenido para el software de Z39.50

= @/directorio_base ++ [ToC.dbf | ToC.dbi | ToC.mdg | ToC.imx | ToC.mdt | ToC.mdk | ToC.{0-9}]
*** Ejemplo: @/export/sky/uno/revdgsca/datos/.indices/ + [ToC.dbf | ToC.dbi | ToC.mdg | ToC.imx | ToC.mdt | ToC.mdk | ToC.{0-9}]

Correo electrónico del administrador del servicio

= revdgsca@calmecac.dgsca.unam.mx

Diccionario de datos para el módulo de Publicación en Web del acervo de revistas

URL del servicio

= http://calmecac.dgsca.unam.mx

Nombre de revista

= {carácter válido}

Carácter valido

= [A-Z] - |]

Ficha bibliográfica

= Título + Editorial + Lugar de edición + SSN / Num. Ref. + Idioma + suscripción + Periodicidad + Acervo

Año disponible para revistas

= [1996 | 1997 | 1998 | 1999]

Nombre de Volúmenes

= No. Volumen + No. Revista + Mes + Año
** Ejemplo: Vol. + 0008 Num. 0001 (JANUARY 1996)

Parámetro que active el tipo de información que se mostrará en Web, siguiendo una jerarquía en la navegación.

= nivel

Valores de los parámetro

= [1| 2| 3| 4|5]
** Ejemplo: nivel = 2

Datos solicitados al usuario
para que se le envíe copia del artículo = @ Nombre Completo + E-MAIL + Teléfono + Dirección
que se elige Postal

Dirección Electrónica del encargado de = biblio@servidor.unam.mx
La biblioteca

Primer nivel de navegación. = Títulos de revistas

Segundo nivel de navegación. = Ficha bibliográfica y años disponibles

Tercer nivel
de navegación. = Volúmenes para el año elegido

Cuarto nivel de navegación. = Mostrar el nombre de la tabla de contenido del volumen

Quinto nivel de navegación. = Mostrar los títulos de los artículos de la tabla de contenido y elegir
cuales se desean

Diccionario de datos para el módulo de Búsqueda en Web de información en el acervo de revistas

URL del servicio = http://calmecac.dgsca.unam.mx

Palabra o frase a buscar = {carácter válido}

Carácter válido = [A-Z|a-z|0-9|]

Comandos de ejecución del software
Z39.50 para la búsqueda = lsearch

Datos solicitados al usuario
para que se le envíe copia del artículo = @ Nombre Completo + E-MAIL + Teléfono + Dirección
que se elige Postal

Dirección Electrónica del encargado de = biblio@servidor.unam.mx
La biblioteca

5.2.2.2. Representación para el sistema de las revistas y sus tablas de contenido

Dentro de los diagramas de flujo de datos se advierte que el sistema tiene en todos sus módulos entradas de información, transformación de ella y finalmente salidas; todos estos datos son empleados por los módulos y submódulos del sistema. A continuación se mostrarán cada una de estas entidades de información y cuál es el papel que juegan dentro del sistema.

Identificación del nombre de las revistas

Para los diferentes procesos de cada modulo del sistema se requiere identificar las revistas de una forma sencilla, para ello se empleará el nombre de la misma y su ISSN; dado que el ISSN es un número con tamaño fijo y con menos caracteres que el nombre de la revista este es el elemento con mayor importancia dentro del sistema. Recordemos que si la revista no tiene un ISSN se le dará un numero único y consecutivo para sustituir este elemento.

En cada uno de los procesos que requieran la identificación de la revista, se llevará a cabo un mapeo del nombre de la revista y su correspondiente identificación (ISSN o número). La estructura de esta información se presenta a continuación:

ACM Communications:0001-0782
ACM Computing Surveys:0360-0300
ACM Journal:XXXX-0001
ACM SIGPLAN Notices:0362-1340
ACM Transactions on Database Systems:0362-5915
ACM Transactions on Mathematical Software:0098-3500
Advanced Systems:XXXX-0002
American Journal of Distance Education:0892-3647
Applied Mathematics and Computation:0096-3003
Artificial Intelligence:0004-3702
Basica Revista de la Escuela y el Maestro:XXXX-0003
Boardwatch Magazine:1054-2760
Business Week:0007-7135
Byte:0360-5280
Byte Mexico:XXXX-0004
Cadence:0887-9141
CD-ROM Professional:1090-946X
Ciencia y Desarrollo:0185-0008
Ciencia y Tecnologia:XXXX-0005
Clipper Plus:XXXX-0006
Communications Info Disk CD-ROM:XXXX-0007
Communications of the ACM:0001-0782
Computer:0018-9162
.
.
.
Software Practice and Experience:0038-0644
Software World:0038-0652
Soluciones Avanzadas:0188-8048
Sun Expert Magazine:1053-9239
Sun World:XXXX-0023
Tecnologia y Comunicacion Educativas:0187-0785
Technology and Learning:1053-6728
Ulrich's Plus:XXXX-0024
Unix Review:0742-3136
Unix World:0739-5922
Visual Developer:1053-6205
Windows Magazine:1060-1066
WordPerfect for Windows:XXXX-0025
WordPerfect Magazine:1042-5152

Figura 1.

Tablas de contenido

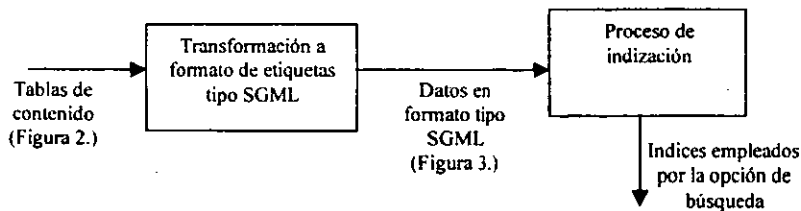
Una vez detectados los elementos de las tablas de contenido, se solicitó al personal de la biblioteca que las proporcionará a la entrada del sistema en el formato que a continuación se muestra. La disposición de los elementos de la tabla de contenido son distintivos de las revistas y la cadena "@@@@@" indica la separación entre los datos de la revista y el índice de la misma.

NOMBRE DE LA REVISTA	
VOL. # NO. #	
MES AÑO	
@@@@@	
<Línea en blanco>	
NOMBRE DEL ARTICULO	
Resumen del artículo	No. página de inicio
<Línea en blanco>	
NOMBRE DEL ARTICULO	
Resumen del artículo	No. página de inicio
<Línea en blanco>	
NOMBRE DEL ARTICULO	
Resumen del artículo	No. página de inicio

Figura 2.

Disposición de las tablas de contenido para el proceso de indización hecho por el software de Z39.50

Para que el software de Z39.50 lleve a cabo las búsquedas en las tablas de contenido de las revistas, antes se debe de llevar a cabo el proceso de indización de las mismas (en el DFD de Incorporación y actualización del acervo de las revistas); para la tarea de indización se requiere como entrada de datos una disposición específica de las tablas de contenido (Figura 3.), esta información se transformará y se producirá una salida.



%%%<DI><ARTICULO> FEATURE The Intel 4004 Microprocessor: What Constituted
Invention? William Aspray 4 </ARTICULO><REVISTA>IEEE ANNALS OF THE HISTORY
OF COMPUTING</REVISTA><VOLUMEN>0019</VOLUMEN><NUMERO> 0003</NUMERO>
<FECHA>JULY-SEPTEMBER 1997 </FECHA><CODIGO>2</CODIGO></DI>

%%%<DI><ARTICULO> Douglas Carl Engelbart: Developing the Underlying Concepts for
Contemporary Computing Susan B. Barnes

16</ARTICULO><REVISTA>IEEE ANNALS OF THE HISTORY OF COMPUTING
</REVISTA><VOLUMEN>0019</VOLUMEN><NUMERO> 0003</NUMERO><FECHA>JULY-
SEPTEMBER 1997 </FECHA><CODIGO>3</CODIGO></DI>

%%%<DI><ARTICULO> Economic Preconditions That Made Possible Appl ication of Comm
ercial Computing in the United States James W. Cortada

27</ARTICULO><REVISTA>IEEE ANNALS OF THE HISTORY
OF
COMPUTING</REVISTA><VOLUMEN>0019</VOLUMEN><NUMERO>0003</NUMERO><FECHA
>JULY-SEPTEMBER 1997</FECHA><CODIGO>4</CODIGO></DI>

%%%<DI><ARTICULO> Mathematics, Technology, and Trust: Formal Verification, Comp
uter Security, and the U.S. Military Donald MacKenzie and Garrel Pottinger

41</ARTICULO><REVISTA>IEEE ANNALS OF THE HISTORY OF
COMPUTING</REVISTA><VOLUMEN>0019</VOLUMEN><NUMERO>0003</NUMERO><FECHA
>JULY-SEPTEMBER 1997</FECHA><CODIGO>5</CODIGO></DI>

Figura 3 (Archivo con formato de etiquetas SGML, en el desarrollo se identifican como archivos con
Extensión bdf) .

5.2.2.3. Datos de salida

Esta es la información que el sistema arrojará al usuario final en la interfaz Web. En el análisis de requerimientos se definió que lo que se desea ofrecer al usuario es la información completa del acervo de revistas, a continuación se muestra la información en cada uno de los módulos de que consta el sistema:

Publicación en Web del acervo de revistas

Primer nivel de navegación: Títulos de revistas

<p>A</p> <p>ACM COMPUTING SURVEYS ACM JOURNAL ACM SIGPLAN NOTICES ACM TRANSACTIONS ON DATABASE SYSTEMS ACM TRANSACTIONS ON MATHEMATICAL SOFTWARE ADVANCED SYSTEMS AMERICAN JOURNAL OF DISTANCE EDUCATION APPLIED MATHEMATICS AND COMPUTATION ARTIFICIAL INTELLIGENCE</p> <p>B</p> <p>BOARDWATCH MAGAZINE BUSINESS WEEK BYTE BYTE MEXICO</p> <p>C</p> <p>CADENCE CD-ROM PROFESSIONAL CIENCIA Y DESARROLLO COMMUNICATIONS OF THE ACM COMPUTER COMPUTER AND CONTROL ABSTRACTS COMPUTER ARCHITECTURE NEWS COMPUTER COMMUNICATIONS REVIEW COMPUTER GRAPHICS COMPUTER GRAPHICS WORLD COMPUTER JOURNAL COMPUTERWORLD</p> <p>.</p> <p>.</p> <p>.</p>
--

Segundo nivel de navegación: Ficha bibliográfica y años disponibles

REVISTA: ACM SIGPLAN Notices

Titulo	ACM Sigplan Notices
Titulo relacionado	Sigplan Notices
Editorial	Association for Computing Machinery
Lugar de edicion	New York, USA
ISSN / Num. Ref.	0362-1340
Idioma	Ingles
Existe suscripcion (S/N)?	S
Periodicidad	Mensual
Acervo	Vol. (Numeros)
1977	12 (1-3,7-12)
1978	13 (1-7,9-12)
1979	14 (1,3-12)
1997	32 (1-12)
1998	33 (1-12)
1999	34 (1-2)

Tercer nivel de navegación Volúmenes para el año elegido

REVISTA: ACM SIGPLAN Notices

1999

Volumen: 0034

Numero: 0001 (JANUARY 1999)

Volumen: 0034

Numero: 0002 (FEBRUARY 1999)

Cuarto nivel de navegación: Mostrar los títulos de los artículos de la tabla de contenido y elegir cuales se desean

ACM SIGPLAN NOTICES	
VOL.34 NO.2	
FEBRUARY 1999	
CONTENTS Letter From The Editor A Correction and Apology	1
Activities Panel Sessions at SIGPLAN Conferences by Michael Burke	2
Conference Corner Calendar	4
Calls for Papers	16
The European Scene EAPLS and ERCIM PLT by Neil Jones	19
Functional Programming Paths between Imperative and Functional Programming by Thomas Ball	21
Practical Parsing Patterns What to Do With a Dangling Else by Chris Clark	26
Java Reflections Inner Classes by Brent W. Benson, Jr.	32
Forth Forth and the Open Terminal Architecture by Paul Frenger	36
Technical Correspondence Information for Authors	40
Letters to the Editor M. Carlisle, M. Feldman, P. Bigot, A. Syropoulos	41
Declarative Peephole Optimization Using String Pattern Matching Diomidis Spinellis	47

5.3. IMPLANTACIÓN

Lo que se ha explicado hasta ahora, el análisis y diseño del sistema, va dirigido hacia el objetivo final, traducir todos los elementos de análisis y desarrollo a una forma que pueda ser comprendida por las computadoras, es decir la codificación (proceso de transformar el diseño en un lenguaje de programación). Por otra parte en esta etapa se prepara todo el entorno de cómputo para que el sistema pueda trabajar.

5.3.1. Configuración de servicios en red

Como ya se comentó con anterioridad se eligió trabajar con el servidor de Web Apache versión 1.2.6

Para hacer las búsquedas y las indexaciones de la información de las revistas se empleó el software que implementa el protocolo Z39.50, lsearch 1.14 (para la realización de las búsquedas) e lindex 1.14 (para la indexación de las tablas de contenido de las revistas).

Se llevó a cabo la configuración del servidor de Web con establecimiento de canal de comunicación cifrado empleando el esquema de SSL (Secure Sockets Layer); para configurar el servidor Apache versión 1.2.6 con SSLey versión 0.9.0 (la implementación de SSL), ambas aplicaciones se compilaron en conjunto.

5.3.2. Codificación

El sistema está montado en un equipo Sun SPARCstation 5 con sistema operativo Solaris 2.5.1, para el desarrollo del mismo se empleó el lenguaje de programación PERL por representar un medio accesible para la interfaz de los datos de las revistas (tablas de contenido en texto plano, sus índices correspondiente para las opciones de búsqueda, etc.) con Web. Uno de sus principales ventajas de PERL es su aplicación a CGI (interfaz entre los datos de las revistas y Web), por ello se eligió para tal propósito. La versión de PERL es la 5.003.

En la siguiente sección se tendrán dos notaciones importante:

- Cliente: Navegador Web empleado por el usuario.
- Servidor: Servidor HTTP que residirá en la Sun SPARCstation 5 y que ofrecerá el servicio

Incorporación y actualización de información al acervo de revistas

Para actualizar el acervo de Tablas de Contenido de este servicio, se ha definido como procedimiento a seguir:

- Obtener la imagen de la Tabla de Contenido del ejemplo a incluir
- Obtener el archivo de texto resultante del OCR aplicado a la imagen obtenida en el paso anterior
- Dar el formato necesario al archivo de texto obtenido (Figura 2.) para poder ser procesado

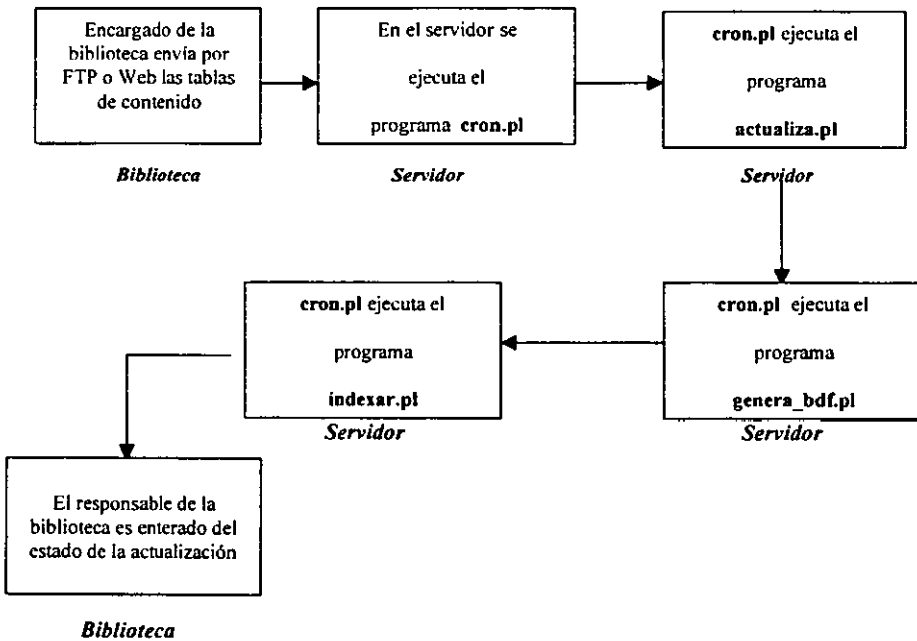
- Salvar el archivo de texto localmente en la PC donde fue generado
- Transferir el archivo de texto a la maquina calmecac.dgsca.unam.mx bajo el directorio /var/tmp/logs/procesar. Esto se puede hacer a través de la aplicación FTP o directamente con el visualizador de Web empleado.
- Repetir los pasos anteriores tantas veces como Tablas de Contenido se digitalicen y vayan a colocar

Una vez que han sido colocados los archivos con el texto de la Tabla de Contenidos, a la 12:00 p.m. es ejecutado un proceso automático (ejecución de un cron en el servidor calmecac.dgsca.unam.mx) llamado **cron.pl** que lleva a cabo las siguientes acciones:

1. verifica si el directorio /var/tmp/logs/procesar tiene nueva información o no. Si el proceso encuentra nuevos archivos, lo que hace es generar un listado completo de ellos a través de la ejecución del comando find de Unix como se muestra a continuación:

```
find parametros
```

2. ejecuta el programa **actualiza.pl**
3. ejecuta el programa **genera_bdf.pl**
4. ejecuta el programa **indexar.pl**

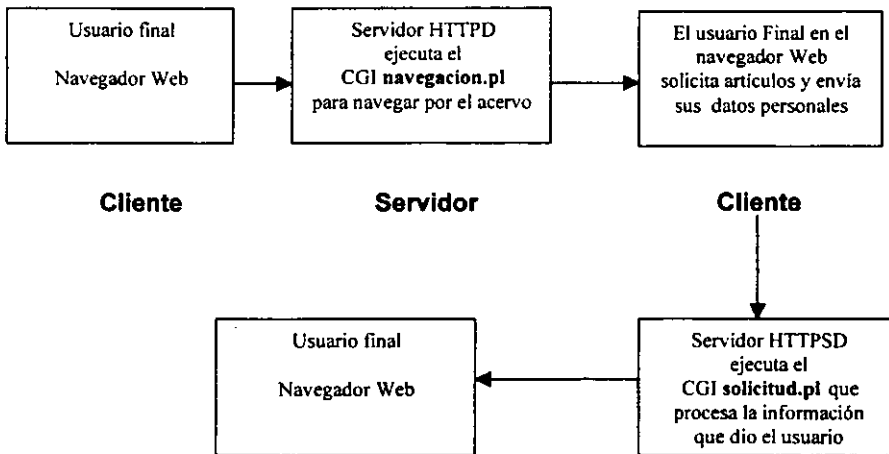


Una vez generada la relación de archivos de texto con las Tablas de Contenido a procesar, se ejecuta el programa (**actualiza.pl**) que verifica el formato de los archivos, verifica que el nombre de la revista sea válido, verifica la inexistencia de una copia anterior del mismo ejemplar y genera la estructura de directorios y archivos de definición necesarios para que sea puesto a disposición del público a través de la navegación Web o revisión por volumen, número y fecha. Además incluye la referencia en el archivo `/var/tmp/logs/indexar.txt` de tal forma que el siguiente programa pueda identificar aquellas Tablas de Contenido que fueron colocadas y que es necesario indexar en el acervo.

Una vez que tenemos colocados los archivos con la Tabla de Contenido en formato texto en la jerarquía de directorios correspondiente, es necesario llevar a cabo la generación de los archivos con el formato de etiquetas necesario para la indexación con `index`. Recordemos que debe ser un archivo con este formato por cada Tabla de Contenido. Por lo tanto, después de colocar la Tabla de Contenido, se ejecuta este programa para generar el archivo con formato de etiquetas tipo SGML. Todo esto lo lleva a cabo el programa **genera_bdf.pl**.

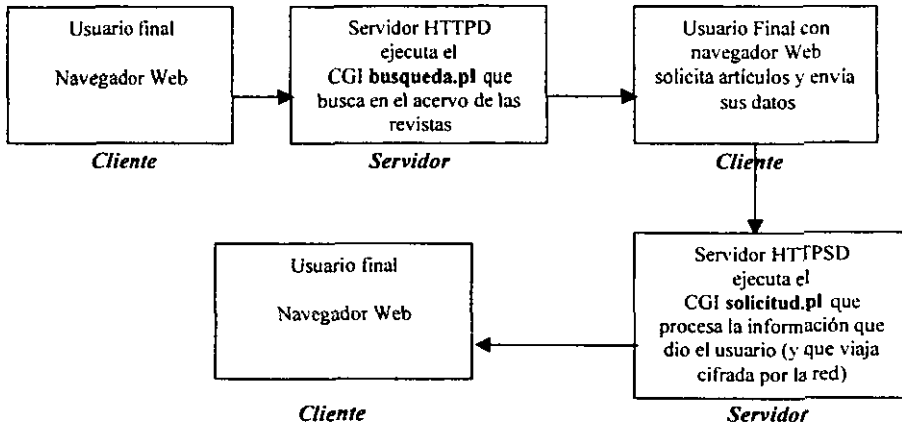
Finalmente se ejecuta el programa **indexar.pl** que ejecuta los comandos para indexar (`index` del software de Z39-50) las tablas de contenido de las revistas:

Publicación en Web del acervo de revistas



Búsqueda en Web de información en el acervo de revistas

Todas las acciones que se realizan en este proceso las lleva a cabo el CGI **busqueda.pl** que presenta la interfaz entre el usuario final (con un navegador Web) y el sistema; dicho sistema reside en el equipo donde se alberga el servidor Web con las páginas y los programas de búsqueda (donde se ejecutan los comandos de **Isearch** del software Z39-50).



El programa **busqueda.pl** nos permite realizar la búsqueda en los índices generados en formato Z39.50 que actúa con base en los parámetros recibidos por la forma de WWW empleada por el usuario.

Este programa genera la cadena con los datos a localizar en el acervo con la correcta conformación lógica; también se reflejan algunas restricciones sobre el sistema de búsqueda, de tal manera que de forma transparente para el usuario se le imposibilite realizar algunas búsquedas absurdas o demasiado generales a fin de evitar la carga innecesaria en el sistema. Permite mostrar de 20 en 20 los artículos que cumplen con la cadena que dio el usuario, los cuales serán recuperados y enviadas al visualizador Web. Además da el formato a la salida que proporciona la ejecución del **Isearch** (comando que hace la búsqueda) y se imprimen en la página de resultados directamente en el visualizador. Todo esto se hace en tiempo real y sin el uso de archivos temporales, lo que evita el uso de espacio y de tiempo. Además lleva un contador especial para determinar dentro de la página, qué número de artículo se está desplegando de tal manera que se facilite la identificación de los mismos al momento de ser seleccionados y pedidos por el usuario. Esta información es colocada en la forma que acompaña a las fichas.

Una vez que se muestran los registros de los artículos que cumplen con el criterio de búsqueda establecido, estos registros presentan un botón de selección para que el usuario pueda solicitar mayor información sobre ellos. Evidentemente, para poder solicitar tal información, se recaban los datos básicos del usuario para poder tener un registro de quién está realizando la petición. Entonces se solicita el nombre completo, la dirección postal y la dirección electrónica en donde se le puede hacer llegar los artículos solicitados, así como un teléfono de referencia. Todo esto es recabado a través de una forma HTML que se muestra en el navegador del usuario.

El CGI **busqueda.pl** ejecuta el programa **solicitud.pl** que es el que procesa los datos de los usuarios que llenan la forma de datos personales, los almacena y genera las solicitudes necesarias para el responsable de la biblioteca, quien se pondrá en contacto con el usuario para proporcionarle el artículo digitalizado. El sistema le mostrará al usuario un recibo con un número de solicitud, el cual le permitirá hacer referencia a su petición y dar un seguimiento al proceso en caso de duda o error.

5.3.3) Mantenimiento

El mantenimiento del sistema es llevado a cabo principalmente por el administrador del equipo en donde reside el servicio, ya que el acervo es almacenado y/o generado en el directorio o sistema de archivos que desde la instalación se definió para ello. Evidentemente es necesario garantizar que haya el espacio disponible para el alojamiento de la información que está siendo incorporada. Cada proceso que se lleva a cabo en el sistema genera registros en una bitácora, la cual debe ser depurada con regularidad y colocada en algún sistema de archivos distinto (según sea definido por el administrador) para facilitar su manejo y depuración. Esto es llevado a cabo a través de la definición de las variables al inicio de los programas involucrados. El manejo de estas bitácoras es similar al realizado sobre cualquier otra bitácora de servicio (Web, correo electrónico por mencionar algunos). También será necesario llevar a cabo la revisión de tales bitácoras a fin de garantizar el correcto funcionamiento del sistema y detectar oportunamente cualquier anomalía.

Del sistema, además de los programas y aplicaciones involucradas es necesario llevar a cabo el respaldo de la estructura jerárquica que se genera en el directorio de datos del servidor de Web y las tablas de índices, garantizando al máximo su integridad. En realidad, la información más importante para el servicio de búsqueda son los archivos de índices y los archivos en formato SGML (con extensión .bdf). Para el servicio de navegación, lo más relevante es la estructura jerárquica y los archivos de texto.

En conclusión, es necesario respaldar y mantener el directorio de datos del servidor de Web que brinda el servicio al público y esto se deja a criterio del administrador del sistema, ya que se ajustará este proceso a las políticas que se establezcan en este sentido. Actualmente, en el servicio que se presta en la biblioteca de DGSCA, los respaldos son realizados en forma incremental del servicio con una periodicidad semanal y un respaldo total del equipo en forma mensual.

5.3.4 Costos asociados

Desde el planteamiento del proyecto, la idea que se ha tenido en mente es llevar a cabo su establecimiento en aquellas bibliotecas que no tienen los medios económicos, tecnológicos y de personal como para adquirir una aplicación comercial, equipo especializado o contratar personal especializado para brindar este tipo de servicios al público. Por lo tanto, los costos han sido reducidos al máximo y, estos pueden ser divididos en:

Aplicaciones

Las aplicaciones empleadas en este proyecto son del dominio público (servidor de Web, herramientas de indexación y búsquedas en texto completo, herramientas de conversión y manejo de imágenes, lenguaje de programación y medio de distribución, aplicación de seguridad), lo cual implica que no existe ningún costo monetario asociado a su adquisición y puesta en funcionamiento. La inversión requerida está más bien al apoyo inicial por parte de personal entrenado para la instalación y puesta en marcha del servicio.

Al tratarse de herramientas del dominio público, se tiene por una parte la certeza de disponer de sus versiones ya compiladas y listas para usarse en diversas plataformas de cómputo y sistemas operativos y, en caso de no encontrarse disponibles, se tiene la posibilidad de generarlas al tener disponible el código fuente correspondiente. Pero por otra parte, se tiene una cierta seguridad en la permanencia de esta tecnología, ya que al estar respaldada y probada por cientos de usuarios a nivel mundial, se puede garantizar un alto grado de permanencia.

Sin embargo, es posible emplear herramientas comerciales (principalmente en el ámbito del servidor de Web y manejo de imágenes), por lo que esto puede hacer variar el costo de la solución completa.

Equipo de cómputo

En cuanto a equipo de cómputo, actualmente no se requieren grandes inversiones en este sentido, ya que con el fortalecimiento del sistema operativo Linux (versión libre de UNIX para plataforma Intel), es posible establecer sistemas de información a través de la red con alto grado de robustez y eficiencia con una inversión en equipo muy modesta. En este sentido, el factor a ser considerado de manera especial en este sentido es el espacio de almacenamiento para los índices, tablas de contenido e imágenes digitales. Sin embargo, la tendencia en el mercado es un rápido abaratamiento en el costo de este recurso de cómputo, por lo que no resulta ser un factor relevante en el costo global del sistema. El otro factor importante a considerar es el total de la memoria RAM en el sistema, factor que si es necesario considerar al evaluar alguna opción tecnológica.

Por lo tanto, el costo asociado a los equipos pueden variar de unos \$20,000.00 a \$25,000.00 pesos (costo de un equipo PC con capacidad de ejecución de Linux) hasta varios miles de dólares si se desea emplear algún equipo RISC con sistema operativo UNIX licenciado.

En cuanto a otros recursos de cómputo, evidentemente se encuentra la utilización de un scanner (o digitalizador), periférico cuyos costos tampoco son altamente significativos y que tiende a ser cada vez menor. En este sentido será importante evaluar la cantidad de digitalizaciones que vayan significando el establecimiento y desarrollo de este servicio. Un scanner puede variar entre \$4,000.00 y \$6,000.00 pesos aproximadamente.

Recursos humanos

El presente sistema está diseñado pensando en requerir un mínimo de intervención de operadores y/o administradores en forma directa con el sistema. Sin embargo, el controlar y atender un sistema en ambiente operativo UNIX requiere de contar con los servicios de una persona con cierto grado de entrenamiento especializado y experiencia. Esta actividad podría significar una erogación mensual no menor a \$6,000.00 o bien, el establecimiento de alguna estrategia de entrenamiento del personal actual en la biblioteca para tomar estas funciones. Una vez que se tiene cubierto este puesto, seguramente la misma persona puede ser entrenada adecuadamente para llevar a cabo el mantenimiento y/o modificaciones que se requieran en el servicio brindado.

Ahora bien, desde el punto de vista de la digitalización, es necesario recordar que al hacer la transformación de una imagen digital a un archivo de texto a través del proceso de conversión de OCR, si bien el grado de eficiencia y confiabilidad del proceso es superior al 95%, evidentemente se tienen aun ciertos errores. El esfuerzo y personal necesario para la revisión de este texto a fin de eliminar los errores puede llegar a ser importante. Además, si se tiene en mente a largo plazo la posibilidad de convertir a texto el cuerpo de los propios artículos a partir de sus digitalizaciones, este proceso pudiera ser aun más demandante de recursos humanos, tiempo y scanners. Así que será necesario que la biblioteca evalúe este factor y determine la inversión a ser hecha. La recomendación es que se de el proceso de revisión sólo en las tablas de contenido.

Tiempo

El tiempo invertido en el servicio es significativo y relevante sólo en la parte de conversión de imágenes a texto de las tablas de contenido, a fin de garantizar la integridad de la información. Otro proceso que puede resultar oneroso en tiempo es la digitalización de los artículos solicitados por los clientes. Sin embargo, este proceso va llenando el acervo de la biblioteca, de tal forma que en la situación ideal, la biblioteca tendría digitalizadas las tablas de contenidos mas la totalidad de los artículos que conforman su acervo. Todos los demás procesos consumen poco tiempo y se irán integrando a las actividades propias de la organización.

Programación

No es necesario contar con un programador experimentado para dar mantenimiento al sistema o para realizar modificaciones. En realidad, sólo se requiere una persona con conocimientos de programación y diseño de páginas de Web para poder modificar el sistema. Como se planteó desde un inicio, en una primera fase se puede establecer que esta actividad la lleve a cabo la misma persona que administra el sistema operativo UNIX.

5.4. PRUEBAS

La prueba del sistema generado conforma un elemento crítico para la garantía de calidad del software y representa el último repaso de las especificaciones, del diseño y de la codificación. La prueba tiene los siguientes objetivos:

- La prueba es un procesos de ejecución de un programa con la intención de descubrir un error.
- Un buen caso de prueba es aquel que tiene un alta probabilidad de mostrar un error no descubierto hasta entonces.
- Una prueba tiene éxito si descubre un error no detectado hasta entonces.

Se efectuaron las pruebas para cada uno de los módulos del sistema donde se dieron entradas de información de todo tipo, de tal suerte que se detectaran las posibles fallas y agregar el código o mejorarlo.

Publicación en Web del acervo de revistas

The image shows a Netscape browser window with the title "Netscape: Titulos de revistas disponibles". The address bar contains the URL "http://calmecac.dgsc.unam.mx/cgi-bin/navegacion.pl?nivel=0". The main content area displays the heading "Titulos de revistas disponibles" followed by a list of journal titles under two sections, A and B.

Titulos de revistas disponibles

[A B C D E G H I J K L M N O P Q R S T U V W]

A

- [ACM COMPUTING SURVEYS](#)
- [ACM JOURNAL](#)
- [ACM SIGPLAN NOTICES](#)
- [ACM TRANSACTIONS ON DATABASE SYSTEMS](#)
- [ACM TRANSACTIONS ON MATHEMATICAL SOFTWARE](#)
- [ADVANCED SYSTEMS](#)
- [AMERICAN JOURNAL OF DISTANCE EDUCATION](#)
- [APPLIED MATHEMATICS AND COMPUTATION](#)
- [ARTIFICIAL INTELLIGENCE](#)

B

- [BOARDWATCH MAGAZINE](#)
- [BUSINESS WEEK](#)

Open the web page editor

Netscape: Advanced Systems

File Edit View Go Communicator Help

Back
 Forward
 Reload
 Home
 Search
 Guide
 Print
 Security
 Stop

Bookmarks
 Netsite: <http://calmecac.dgscs.unam.mx/cgi-bin/navegacion.pl?nivel=1&dir=XXXX-0002/>

Advanced Systems

Título	Advanced Systems
Editorial	Integrated Media Inc.
Lugar de edicion	California, USA
Idioma	Inglés
Existe suscripcion (S/N)?	N Cancelada para 1996
Periodicidad	Mensual
Acervo	Vol. (Numeros)
1995	8 (1-3)

[Ejemplares de 1996](#)

[\[Inicio \]](#) [Anterior \]](#)

100%

Netscape: Advanced Systems

File Edit View Go Communicator Help

Back Forward Reload Home Search Guide Print Security Stop

Bookmarks Netsite: <http://calaecac.dgscs.unam.mx/cgi-bin/navegacion.pl?nivel=2&dir=XXXX-0002/1996/>

Advanced Systems

Ejemplares de 1996

[Vol. 0008 Num. 0001 \(JANUARY 1996\)](#)
[Vol. 0008 Num. 0002 \(FEBRUARY 1996\)](#)
[Vol. 0008 Num. 0003 \(MARCH 1996\)](#)

[\[Inicio\]](#) [Anterior](#)

Document: Done.

Netscape: Advanced Systems

File Edit View Go Communicator Help

Back Forward Reload Home Search Guide Print Security Stop

Bookmarks Go To: calaecac.dgscs.unam.mx/cgi-bin/navegacion.pl?nivel=3&dir=XXXX-0002/1996/0008-0001/

Advanced Systems

Vol. 0008 Num. 0001

[Tabla de Contenido Vol. 0008 Num. 0001 \(JANUARY 1996\)](#)

[\[Inicio\]](#) [Anterior](#)

Document: Done.

Netscape: Tabla de Contenido

File Edit View Go Communicator Help

Back Forward Reload Home Search Guide Print Security Stop

Bookmarks Go To: galmecac.dgscsa.unaa.mx/cgi-bin/navegacion.pl?nivel=4&dir=XXXX-0002/1996/0008-0001/te.txt /

Tabla de Contenido

ADVANCED SYSTEMS
VOL.8 NO.1
JANUARY 1995
@@@@@

REVIEWS
_ Artículo seleccionado

Creating order from chaos Are your projects a disorderly collection of missed deadlines? Then don't miss our comparison of three Unix-based project management packages from NIS, Digital Tools, and PSDI. By Cedric Higgins & Dalia Freeman 26
■ Artículo seleccionado

COVER STORY
_ Artículo seleccionado

Hot boxes. Three ways to serve three Pushing into the enterprise, dedicated Unix servers are all the rage. We test and rate the latest in beefy hardware from Digital Equipment (the first-ever review of an SMP AlphaServer 2100), Advanced Logic Research, and Hewlett-Packard. By Mark Cappel 36
■ Artículo seleccionado

Netscape: Tabla de Contenido

File Edit View Go Communicator Help

Back Forward Reload Home Search Guide Print Security Stop

Bookmarks Go To: galmecac.dgscsa.unaa.mx/cgi-bin/navegacion.pl?nivel=4&dir=XXXX-0002/1996/0008-0001/te.txt /

_ Artículo seleccionado

Para mayor información sobre los artículos seleccionados, favor de proporcionar los siguientes datos
Los campos con nombre resaltado son obligatorios, los demás son opcionales

Nombre Completo:

Domicilio electrónico (E-mail):

Teléfono:

Dirección Postal:

Gate y Num:

Búsqueda en Web de información en el acervo de revistas

Búsqueda en las Publicaciones Periódicas en la Biblioteca de la DGSCA - Netscape

File Edit View Go Communicator Help

Bookmarks Netsite: http://calmecac.dgsc.a.unam.mx/busqueda.html? What's Related

Universidad Nacional Autónoma de México

Dirección General de Servicios de Cómputo Académico

Biblioteca

REVISTAS ELECTRONICAS

Este servicio contiene las tablas de contenido de más de 100 títulos de publicaciones especializada en informática que se encuentran en la Biblioteca de la DGSCA. Ponemos a su disposición una breve ayuda sobre la manera de emplear este servicio de información

Determine el criterio para localizar los artículos de su interés:

Palabras a localizar en el título del artículo:

Revista

Aquí se muestran los Títulos de las Revistas Disponibles

Fecha

Desea que las revistas cumplan con:

Todos los valores dados Algún valor

Universidad Nacional Autónoma de México

Dirección General de Servicios de Cómputo Académico

Biblioteca

REVISTAS ELECTRONICAS

Este servicio contiene las tablas de contenido de más de 100 títulos de publicaciones especializada en informática que se encuentran en la Biblioteca de la DGSCA. Ponemos a su disposición una breve ayuda sobre la manera de emplear este servicio de información

Determine el criterio para localizar los artículos de su interés:

Palabras a localizar en el título del artículo:

Revista

Aquí se muestran los Títulos de las Revistas Disponibles

Fecha

Desea que las revistas cumplan con:

Todos los valores dados Algún valor

Resultado de la búsqueda - Netscape

File Edit View Go Communicator Help

Bookmarks Netsite: http://calmecac.dgsca.unam.mx/cgi-bin/busqueda.pl What's Related

Resultado de la búsqueda

ARTICULO: New Media Is the Message Apple's latest PowerPC machines deliver a boxful of easy-to-use multimedia features: audio, video, telephony, CDROM, conferencing, and speech recognition. We check out the PowerMac 7200, 7500, and 8500. 113
REVISTA: BYTE
VOLUMEN: 0021
NUMERO: 0001
FECHA: JANUARY 1996
 Artículo seleccionado

ARTICULO: This image comes from Atlantis, a new CDROM adventure game created by Paris-based Cryo Interactive Entertainment. The game is scheduled to be released early in 1997. For more info on software tools for game developers, see page 29
REVISTA: COMPUTER GRAPHICS WORLD
VOLUMEN: 0019
NUMERO: 0012
FECHA: DECEMBER 1996
 Artículo seleccionado

ARTICULO: THE FUTURE OF CD TECHNOLOGY Next-Generation Compact Discs Alan E. Bell The new versions of compact-disc players and CDROM drives debuting in coming months read small, double-sided discs with enough capacity to hold feature films or music catalogues. Similar devices may soon replace tape-based VCRs. A look at how digital versatile discs (DVDs) work. 26
REVISTA: SCIENTIFIC AMERICAN
VOLUMEN: 0275
NUMERO: 0001
FECHA: JULY 1996
 Artículo seleccionado

ARTÍCULO: THE FUTURE OF CD TECHNOLOGY Next-Generation Compact Discs Alan E. Bell The new versions of compact-disc players and CDROM drives debuting in coming months read small, double-sided discs with enough capacity to hold feature films or music catalogues. Similar devices may soon replace tape-based VCRs. A look at how digital versatile discs (DVDs) work. 28

REVISTA: SCIENTIFIC AMERICAN

VOLUMEN: 0275

NUMERO: 0001

FECHA: JULY 1996

Artículo seleccionado

Para mayor información sobre los artículos seleccionados, favor de proporcionar los siguientes datos

Los campos con nombre resaltado son obligatorios, los demás son opcionales

Nombre Completo:

Ernesto Cruz Guerra

Domicilio electrónico (E-mail):

ecruz@dca.unam.mx

Teléfono: 6223640

Dirección Postal

Busqueda de la Universidad Nacional de Mar del Plata

File Edit View Go Communicator Help

Bookmarks Netsite: http://calmecac.dgscs.unam.mx/cgi-bin/busqueda.pl What's Related

Teléfono: 6223640

Dirección Postal

Calle y Num.
Amapola No. 48

Colonia:
Nueva Santa María

Delegación: Miguel Hidalgo

Estado: México D.F.

País: México

Código Postal: 13400

Realizar la solicitud Limpia la forma y selecciones

CAPÍTULO 6

CAPÍTULO 6

CONCLUSIONES

La tecnología es sin duda alguna el germen que habilita a las organizaciones a elaborar productos mejores, más económicos y convenientes para los usuarios. En la era de la Información, que nos ha tocado vivir y con la acelerada transformación de todos los aspectos fundamentales de la vida, el conocimiento será el ingrediente indispensable para la integración social y política de los individuos, por lo que deberemos tecnificar, modernizar y hacer accesibles los procesos educativos y de divulgación que lleven los cambiantes conocimientos a todos los individuos que los requieran y durante el transcurso de sus vidas.

Cada día se aumenta más la distancia entre la capacidad de innovación y producción tecnológica de nuestro país y la de los países desarrollados y será aún mayor esa distancia si los técnicos, profesionistas e investigadores no buscan proponernos la estrategia de largo plazo que nos incorpore al mundo de la creatividad tecnológica y científica. Es indispensable redefinir la forma de abordar esta problemática para lograr esta incorporación a la base de este nuevo mundo tecnológico, pues las estrategias de muchos años hasta ahora no han dado los resultados suficientes para cambiar la situación en este campo.

En la actualidad con los avances tecnológicos la información tendrá que estar al alcance de toda la sociedad; Internet se está constituyendo como el medio natural de comunicación en todos los ámbitos de la sociedad. Las universidades y sus bibliotecas están conscientes de esta situación y han dirigido sus esfuerzos a incorporarse a este esquema.

La biblioteca de DGSCA tiene como objetivo “obtener información y documentación relevantes los programas de docencia e investigación, hacerla asequible a la comunidad y orientarla en su manejo; así mismo proporcionar los servicios bibliotecarios necesarios para transmitir y acrecentar el conocimiento de la comunidad universitaria y del público en general”.¹

La biblioteca de DGSCA siempre ha buscado medios para cumplir estas metas, y está consciente de que Internet es una forma de efectiva de lograr esto; por ello ha incursionado con el sistema de información bibliográfica a través del software comercial ALEPH con interfaz Web. En este trabajo se planteó el siguiente objetivo de la biblioteca, el tener un sistema de recuperación de información hemerográfica a través de una interfaz Web en Internet.

Del presente trabajo se concluye que la utilización de las características y bondades que brinda actualmente el servicio de WWW en cuanto a la mezcla de sencillez para el usuario (al sólo presentar como interfaz el visualizador) y la riqueza de formatos que puede manejar (texto plano, imágenes, video, sonido, etc.), plantea alternativas muy interesantes a las bibliotecas, ya que muchas de ellas no sólo tienen en sus acervos información de tipo documental o impresa, sino que también incluyen acervos fotográficos, sonidos (como es el caso de la fonoteca de RadioUNAM).

¹ Dirección General de Servicios de Cómputo Académico. Guía de usuario de la Biblioteca DGSCA-CU. México DGSCA

El proyecto planteado fácilmente puede adaptarse a cualquiera de este tipo de acervos y en forma sencilla puede incluir cualquier formato de información.

Se logró cubrir con los objetivos planteados inicialmente, ya que en la actualidad se tiene en línea el servicio de recuperación de información de las publicaciones seriales ofreciendo las opciones de recuperación por orden alfabético de las revistas que anota los volúmenes disponibles (en cada año disponible) y obtener de cada revista su ficha bibliográfica y su tabla de contenido. Esto es equivalente al tradicional kardex manual que muestra a su vez las tablas de contenido de acuerdo a la periodicidad de la revista. Con esto los usuarios pueden revisar las tablas de contenido como si estuvieran hojeando físicamente el índice de la revista.

Por otro lado se tiene la opción de búsqueda en el contenido de los títulos de los artículos de las publicaciones seriales. A través de la interfaz en Web en una caja de diálogo los usuarios indican las palabras o términos a localizar. En esta fase los usuarios tienen la posibilidad de marcar (seleccionar) aquellas referencias que satisfagan el perfil de su búsqueda. Al finalizar el despliegue de los registros se anexa una forma de registro para solicitar los documentos primarios, los cuales se les harán llegar sin costo alguno, vía correo electrónico.

Como ya se ha mencionado, el esquema está pensado para ser extendido y poder poner a disposición del público versiones en texto plano o imágenes digitalizadas de los artículos completos. Evidentemente el proceso de digitalización y reconocimiento óptico de caracteres es un proceso que debe ser planteado y analizado con cuidado, ya que si bien la tecnología en este campo ha avanzado en gran medida logrando alcanzar niveles del 98% de exactitud al aplicar la conversión de imagen a texto, es necesario evaluar si se está en posibilidad de aceptar el 2% de errores en el texto que se publica. En general, este índice de error será aceptable, pero habrá material que por su naturaleza y contenido será necesario reducir tal tasa, lo que implica un proceso de revisión y depuración de la información antes de colocarlo a disposición del público. Evidentemente esto requerirá de la inversión de recursos de cómputo, humanos y de tiempo.

Con el creciente auge de la tecnología y de software libre en la red se logró implementar el sistema con software con esta naturaleza, es decir, se emplearon programas que no implicaron un gasto por parte del personal de la Biblioteca. Se puede pensar que por no emplear software comercial el servicio que se está prestando no es de calidad, sin embargo, ante las pruebas del servicio tanto en la recuperación por orden alfabético de las publicaciones seriales que anota los años disponibles, como en la búsqueda de una palabra clave en las tablas de contenido de las revistas, se observó que la respuesta tanto del servidor Web, como del software que realiza las búsquedas (implementación del protocolo Z39-50) es rápida.

Como se había mencionado, una de las premisas de las que se partía era el emplear herramientas de dominio público, a fin de evitar que las bibliotecas realizaran algún gasto extra en cuanto a aplicaciones o equipo de cómputo especializado. Actualmente, el tener un servidor UNIX ya no resulta ser difícil. La aparición y gran auge que ha recibido Linux (versión del sistema operativo UNIX para plataforma Intel y de distribución gratuita) ha venido a constituirse en una opción muy atractiva para evitar la adquisición de equipos en plataforma RISC, en donde normalmente se encontraba el ambiente operativo UNIX. Con las nuevas características de configuración en constelación de varias máquinas Intel, Claro que la calidad de servicio que se puede brindar en un equipo RISC y con aplicaciones comerciales de indexación y búsqueda pueden verse incrementadas en gran medida, sin embargo, la opción planteada puede resultar atractiva y será necesario que la propia biblioteca sea quien tome la decisión en cuanto a los aspectos de costo beneficio.

Un aspecto importante que deja como resultado el presente trabajo es el hecho de que cuando se realizan las búsquedas de información lleva unos cuantos segundos para revisar la información almacenada en el acervo de las revistas y desplegar los registros localizados en un número de revistas dado en ciertos años; el mismo procedimiento de forma manual hubiera consumido varios minutos u horas. El factor tiempo y el esfuerzo invertido por parte del usuario para la consulta representaron ventajas del sistema. Al igual que la facilidad de proveer los documentos primarios, principalmente, por correo electrónico son características de gran valía.

El personal de la Biblioteca necesita llevar a cabo la digitalización o captura manual de la tabla de contenidos (y más adelante de las digitalizaciones y versiones texto de los artículos) sólo una vez y quedan disponibles en el acervo en forma inmediata.

Desde el punto de vista operativo y económico para las bibliotecas o unidades de información el sistema representó ahorros considerables por que una misma colección de publicaciones seriales, físicamente ubicadas en la biblioteca de DGSCA-Ciudad Universitaria, es cotidianamente utilizada en las otras bibliotecas de los Centros Mascarones y Nuevo Leon. Esta dinámica podría ser ampliada a otros integrantes del sistema bibliotecarios de la UNAM dimensionando los beneficios para la institución y para la comunidad en general.

Es evidente que, la tecnología, siendo un campo tan dinámico, irá presentando diversas opciones y alternativas distintas a las presentadas en el presente trabajo. Es necesario estar atentos a lo que sean las tendencias más fuertes y estables en este sentido. Actualmente se está fortaleciendo el empleo de Java como lenguaje de programación para las interfases de WWW con el usuario. Sin embargo, actualmente el conocimiento de tal lenguaje no está tan diseminado como para poder emplear tal tecnología y fácilmente integrar personal calificado para el mantenimiento del sistema. También en las herramientas empleadas para la búsqueda y localización de información se irán dando avances, a los cuales se deberán estar atentos. Sin embargo, es necesario identificar aquellas tendencias que sean estables y fuertes para poder considerarlas candidatas a ser incorporadas en el presente trabajo.

Desafortunadamente no ha habido un trabajo constante y sostenido en cuanto al desarrollo de los estándares y esquemas para el establecimiento de bibliotecas digitales, por lo que es necesario coordinar y encaminar los esfuerzos de las bibliotecas participantes en este sentido, ya que la tecnología está lista como para brindar servicios electrónicos de nivel aceptable. Sólo es necesario homogeneizar los formatos y estandarizar las metodologías y formas de comunicación y organización de la información. Además, el personal deberá de esta consciente de las implicaciones en cuanto al cambio de los servicios brindados.

APÉNDICE A

APÉNDICE A

Common Gateway Interface (CGI)

El Common Gateway Interface (CGI) surgió como la primera manera de presentar información dinámicamente generada en el World Wide Web. El CGI permite a la computadora generar páginas de Web instantáneamente ante la solicitud de un usuario en lugar de tener que tenerlas físicamente por adelantado. En este momento permanece como el único método estable y bien entendido para crear tales páginas. Java presenta problemas que aun no han sido resueltos y otros productos están por anunciarse pero seguramente pasará tiempo antes de considerarlos como tecnología madura.

Esta es la promesa del CGI. Se pueden mostrar figuras de ventas para productos particulares mes por mes, como es solicitado por el grupo, empleando gráficas complejas. Es posible permitir al cliente que proporcione palabras clave para localizar información de los productos. También se pueden proporcionar adiciones como el recolectar los comentarios de los usuarios en un libro de visi tas.

El mecanismo CGI nos permite realizar una serie de trucos que los usuarios disfrutan viendo en las páginas de Web, pero también existe el lado serio, permitiendo a Internet ofrecer el tipo de interactividad, y manejo de aplicaciones que los usuarios de las computadoras actuales esperan. Los CGI's abren una clase totalmente nueva de aplicaciones modernas a través del Web.

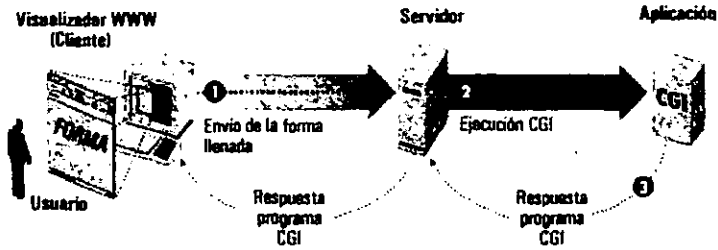
Los usuarios de computadoras esperan actualmente respuestas personalizadas a preguntas particulares. Los días en que el personal quedaba satisfecho cuando el personal del centro de cómputo presentaba un único reporte general a todos los usuarios han terminado. Actualmente, cada persona de ventas, administrador e ingeniero desea formular solicitudes específicas y obtener respuestas adecuadas. Y si una computadora puede hacer esto, ¿Por qué no el Web?.

Esta es la promesa de CGI. Se pueden mostrar figuras distintas para los productos en venta mes con mes, o emplear gráficas de pay o de barras, según lo soliciten. Se puede permitir que el cliente proporcione palabras clave para obtener información de un producto. Además se pueden proporcionar la facilidad de recopilar comentarios de los usuarios, facilitando la búsqueda dentro de los archivos y permitiendo el registro en un libro de huéspedes.

Cuando se navega por Internet, uno puede encontrarse con documentos que resultan espectaculares y nos preguntamos ¿Cómo se hace esto?. Tales documentos pueden consistir, entre otras cosas, de formas que solicitan algún dato o información de registro, mapas o imágenes interactivas que permiten desplazarse a otro documento según la región de la imagen que se seleccionó, contadores que muestran el número de usuarios que han solicitado un documento en particular, utilerías que permiten realizar la búsqueda de información particular dentro de una base de datos. En la mayoría de los casos, se encontrará que tales efectos o servicios son brindados a través de un Common Gateway Interface, conocido más comúnmente como CGI.

El CGI es la parte del servidor de Web que puede comunicarse con otros programas ejecutados por el servidor. Con el CGI, el servidor de Web puede ejecutar un programa, dándole como información de entrada datos del usuario (tales como la máquina desde donde se está conectando, o sí la información de entrada del usuario es proporcionada con la sintaxis correcta definida por HTML). El programa entonces procesa los datos y el servidor envía la respuesta del programa de regreso al visualizador de Web.

El CGI no es mágico; es sólo programación con cierto tipo especial de entrada y reglas estrictas para la presentación de la salida del mismo. Todo lo demás es programación. Por supuesto, hay técnicas especiales que son particulares a los CGI's. El modelo simplificado se muestra en la siguiente figura.



[Diagrama simplificado de CGI]

El CGI convierte el Web de una simple colección de documentos de hipertexto estáticos en un nuevo medio interactivo completo, en donde los usuarios pueden formular preguntas y ejecutar aplicaciones. Revisemos algunas posibles aplicaciones que pueden ser diseñadas empleando CGI's.

Formas

Uno de los usos más socorridos de CGI's es en el proceso de formas. Las formas son un subconjunto de HTML que permiten que el usuario proporcione información. La interfaz de las formas hace que la navegación sea un proceso interactivo para el usuario y el proveedor. La siguiente figura muestra una forma simple.

The image shows a screenshot of a web browser window displaying a simple form. The browser's title bar reads 'Información Organizada de Temas en Mexico - Registro de Vi...'. The address bar shows 'Location: w.unam.mx/Temas/formquestbook.html'. The form is divided into two main sections: 'Datos Personales' and 'Cuestionario'. Under 'Datos Personales', there are three text input fields labeled 'Nombre:', 'B Dirección electronica: [opcional]', and 'URL, si Ud. tiene página personal: [opcional]'. Below these fields is a link that says 'Reservar a las instrucciones'. The 'Cuestionario' section contains a single question: '1. Que piensa sobre la calidad de la presentación de la página de Información Organizada de Temas en México?'. At the bottom of the form, there is a radio button labeled 'Excelente'.

[Forma simple ilustrando diferentes elementos]

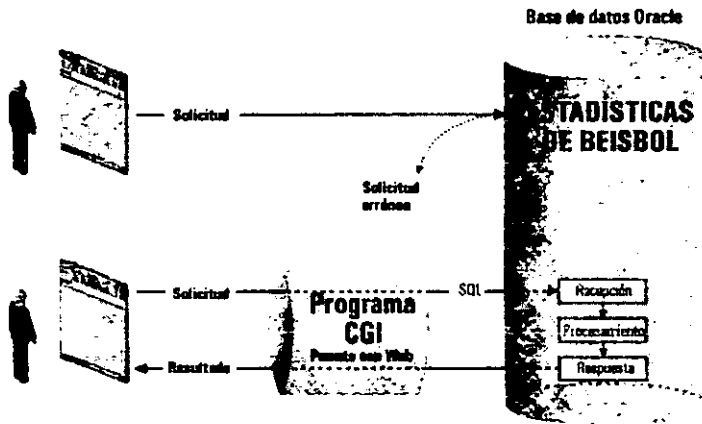
Como podemos observar, diversos elementos gráficos están a disposición para la creación de una forma, tales como botones de selección, campos de recepción de texto, cajas y/o listas de selección. Cuando la forma es completada por el usuario, el botón de envío es empleado para mandar la información al servidor, el cual ejecutará el programa asociado a la forma en particular que interpretará los datos.

Generalmente las formas son empleadas con dos propósitos principales. En la forma más simple, las formas pueden ser empleadas para coleccionar información del usuario. Pero pueden ser utilizadas de una manera más compleja proporcionando interacción en ambos sentidos. Por ejemplo, el usuario puede recibir una lista de documentos disponibles en el servidor, así como una opción para buscar información particular dentro de tales documentos. Un programa CGI puede procesar esta información y recuperar el(los) documento(s) que cumplen con el criterio de selección del usuario.

Puentes (Gateways)

Los puentes (gateways) en Web son programas empleados para tener acceso a información que no está directamente disponible al cliente. Por ejemplo, digamos que se tiene una base de datos en Oracle que contiene estadísticas de jugadores de béisbol del equipo de una compañía y se desea proporcionar tal información a través de Web. ¿Cómo se puede lograr?. Ciertamente no se puede acceder con el cliente directamente al archivo de la base de datos (es decir, abrir un URL asociado a tal archivo) y esperar observar datos con algún significado.

Un CGI representa la solución al problema en la forma de un puente. Se puede emplear un lenguaje como `oraperl` o una extensión de `perl` para formar consultas SQL para leer la información contenida dentro de la base de datos. Una vez que se tiene la información, es posible darle formato y enviarla al cliente. En este caso, el programa CGI sirve como puente para la base de datos Oracle, como se muestra en la siguiente figura.



[Puente con una base de datos]

De manera similar, se puede escribir programas puente para cualquier otro servicio de información en Internet, incluyendo Archie, NNTP (Usenet News), WAIS, etc. Además, se puede amplificar el potencial de los puentes empleando formas que soliciten al usuario una cadena de búsqueda para realizar la recuperación y mostrar información dinámica.

Documentos Dinámicos

La creación de documentos dinámicos son el corazón de los CGI's. Dichos documentos son creados en el momento como respuesta a una solicitud de un usuario. Se pueden crear documentos dinámicos en formato HTML, texto plano, imágenes, y documentos de audio. Un ejemplo simple de un documento dinámico puede ser algo tan trivial como:

Bienvenidos al Servidor de WWW de la Biblioteca de DGSCA

Usted está realizando su conexión desde sky.dgsc.unam.mx. La carga promedio de esta máquina es 1.35.

En este ejemplo, hay dos piezas de información dinámica: el nombre de la máquina desde la que el usuario está realizando la conexión y la carga promedio de la máquina donde reside el servidor.

Por otra parte, documentos dinámicos muy complejos pueden ser creados escribiendo programas que empleen librerías gráficas, puentes y formas. Como un ejemplo más sofisticado, pensemos en que somos los administradores de una galería de arte que se especializa en la venta de réplicas de pinturas renacentistas y estamos interesados en mostrar imágenes de las obras maestras en Web. Se puede empezar creando una forma que solicite al usuario información para propósito de enviar promociones vía correo electrónico, además se presenta un campo de búsqueda para que el usuario especifique el nombre de la pintura, así como una lista de obras selectas populares. Una vez que el usuario envía la forma al servidor, un programa puede enviar por correo electrónico la información del usuario o almacenarla en algún archivo. Dependiendo de la selección del usuario, enviarle un mensaje indicando que la pintura no está en existencia, o enviando una imagen con información histórica localizada en cualquier parte en Internet.

Además de la imagen y la historia, otra forma puede mostrar opciones de procesamiento de imágenes para modificar la brillantez, contraste y/o tamaño de una pintura a ser mostrada. También se pueden desarrollar otros CGI's para modificar otras propiedades de la imagen en el momento empleando librerías gráficas y, enviando la imagen resultante al cliente.

Trabajo interno de un CGI

Pero, ¿qué logra que la interfaz trabaje?. La mayoría de los servidores esperan que los programas CGI's residan en un directorio especial, usualmente llamado cgi-bin, y/o que sean archivos con una determinada extensión. Cuando un usuario define un URL asociado con un programa CGI, el cliente envía una solicitud al servidor solicitando un archivo.

En su mayor parte, la solicitud de un programa CGI es igual a la solicitud de cualquier documento de Web. La diferencia es que el servidor reconoce que la dirección solicitada es un programa CGI, por lo que el servidor no regresa el contenido del archivo, sino que trata de ejecutarlo. A continuación se muestra como es una solicitud del cliente para un CGI:

```
GET /cgi-bin/welcome.pl HTTP/1.0
Accept: www/source
Accept: text/html
Accept: image/gif
User-Agent: Lynx/2.4 libwww/2.14
From: carlost@sky.dgsca.unam.mx
```

Esta solicitud GET identifica el archivo a recuperar como /cgi-bin/welcome.pl. Debido a que el servidor está configurado para reconocer a todos los archivos en el directorio cgi-bin como programas CGI, entiende que debe ejecutar el programa en lugar de enviarlo directamente al visualizador. La cadena HTTP/1.0 identifica el protocolo de comunicación a utilizar.

El cliente además de la solicitud, define los formatos de datos que puede aceptar (www/source, text/html, image/gif), se identifica a si mismo como un cliente Lynx, y envía información del usuario. Toda esta información es puesta a disposición del programa CGI, además de información del propio servidor.

La manera en la que un programa obtiene la información de entrada depende del servidor y del sistema operativo en el que este último esté siendo ejecutado. En los sistemas UNIX, los programas CGI obtienen la información de entrada de la entrada estándar (STDIN) y de las variables de ambiente de UNIX. Estas variables almacenan información tal como la cadena de búsqueda de entrada (en el caso de una forma), el

formato de la entrada, la longitud de la entrada (en bytes), la máquina remota y el usuario que proporciona la entrada, e información propia del cliente. Además almacenan el nombre del servidor, el protocolo de comunicación y el nombre de la aplicación que constituye el servidor.

Una vez que el programa CGI inicia su ejecución, puede crear como respuesta un nuevo documento o proporcionar un URL hacia uno ya existente. En UNIX, los programas envían su respuesta a la salida estándar (STDOUT) como una cadena de datos. Tal cadena consiste de dos partes. La primera es un encabezado parcial o total HTTP, que al menos describe qué formato de datos es regresado (HTML, texto plano, GIF, etc.). Una línea en blanco que indica el final de la sección de encabezado. La segunda parte es el cuerpo, que contiene los datos acordes al formato definido en el encabezado. El cuerpo no es modificado o interpretado por el servidor de ninguna manera.

regresado (HTML, texto plano, GIF, etc.). Una línea en blanco que indica el final de la sección de encabezado. La segunda parte es el cuerpo, que contiene los datos acordes al formato definido en el encabezado. El cuerpo no es modificado o interpretado por el servidor de ninguna manera.

Un programa CGI puede elegir enviar los datos recién generados directamente al cliente o enviarlos indirectamente a través del servidor. Si la salida consiste de un encabezado HTTP completo, los datos son enviados directamente al cliente sin modificación por parte del servidor, o como usualmente se hace, la salida es enviada al servidor como una cadena de datos. El servidor es entonces responsable de agregar la información completa en el encabezado y emplear el protocolo HTTP para transferir los datos al cliente.

A continuación se muestra una salida de un programa generando un documento dinámico HTML, con el encabezado HTTP completo:

```
HTTP/1.0 200 OK
Date: Thursday, 23-July-98 18:00:00 GMT
Server: NCSA/1.4.2
MIME-version: 1.0
Content-type: text/html
Content-length: 2000

<HTML>
<HEAD>
  <TITLE>
  Bienvenidos al Servidor de la Biblioteca de DGSCA
</TITLE>
</HEAD>
<BODY>
  <H1>
  Bienvenidos
</H1>
  ...
</BODY>
</HTML>
```

El encabezado contiene el protocolo de comunicación, la fecha y la hora de respuesta, el nombre del servidor y su versión, y la versión del protocolo MIME). Lo más importante, consiste en la definición Content-type MIME y el número de caracteres (equivalente al número de bytes) de los datos enviados, así como los datos por sí mismos. Ahora, veamos la salida con encabezado parcial:

```
Content-type: text/html
```

```
<HTML>
<HEAD>
<TITLE>
  Bienvenidos al Servidor de la Biblioteca de DGSCA
</TITLE>
</HEAD>
<BODY>
<H1>
  Bienvenidos
</H1>
...
</BODY>
</HTML>
```

En este caso, la única línea de encabezado es la definición Content-type, que describe el formato MIME de la salida. Debido a que la salida está en formato HTML, el tipo declarado es text/html.

La mayoría de los programadores de CGI's prefieren proporcionar sólo un encabezado parcial. Es mucho más simple indicar el formato de los datos y los mismos datos que formular el encabezado completo, cosa que puede ser dejada al servidor. Sin embargo, hay ocasiones en donde se necesita enviar la información directamente al cliente.

Introducción de información a un CGI

Cuando un programa CGI es ejecutado, la información de entrada es puesta a su disposición dividida en tres grandes grupos:

- Información sobre el cliente, servidor y usuario
- Datos que el usuario proporcionó a través de la forma
- Información adicional acerca de rutas

La mayor parte de la información sobre el cliente, servidor o usuario es colocada en variables de ambiente. Los datos de la forma pueden ser incorporados en variables de ambiente, o incluidas en el "cuerpo" de la solicitud. La información adicional sobre rutas es colocada en variables de ambiente.

Como se puede observar, las variables de ambiente son cruciales para que los CGI's puedan obtener sus datos de entrada.

Empleando las variables de ambiente

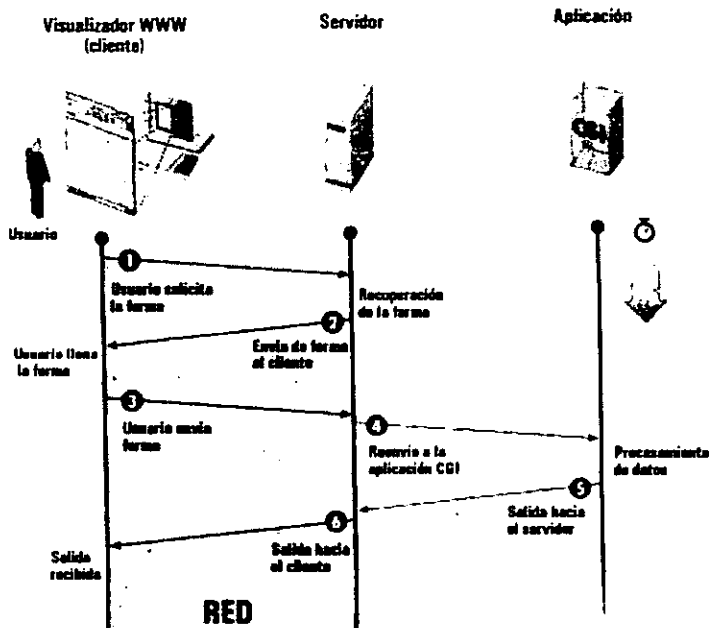
La mayor parte de la información crucial necesaria para un CGI es puesta a disposición a través de variables de ambiente UNIX. Los programas pueden acceder esta información como lo hacen con cualquier otra variable de ambiente (es decir, a través del arreglo asociativo %ENV en el caso de Perl).

A continuación se muestra una lista completa de las variables de ambiente disponibles para un CGI.

Variable de Ambiente	Descripción
GATEWAY_INTERFACE	El número de revisión de Common Gateway Interface que el servidor emplea
SERVER_NAME	El nombre canónico o la dirección IP del servidor
SERVER_SOFTWARE	El nombre y la versión de la aplicación servidora que responde las solicitudes del cliente
SERVER_PROTOCOL	El nombre y la revisión del protocolo de información con la que se realizó la solicitud
SERVER_PORT	El número de puerto al que el servidor está asociado
REQUEST_METHOD	Método con el que la información está siendo proporcionada
PATH_INFO	Información adicional sobre rutas proporcionada al programa CGI
PATH_TRANSLATED	La versión traducida de la ruta proporcionada en la variable PATH_INFO
SCRIPT_NAME	La ruta virtual (es decir, /cgi-bin/programa.pl) del programa que se encuentra en ejecución
DOCUMENT_ROOT	El directorio donde son tomados los documentos de Web proporcionados por el servidor
QUERY_STRING	La información de la forma que se proporciona al programa. Es agregada al URL después de un "?"
REMOTE_HOST	El nombre de la máquina desde donde se está realizando la solicitud
REMOTE_ADDR	La dirección IP de la máquina desde donde se está realizando la solicitud
AUTH_TYPE	Método de validación para autenticar al usuario
REMOTE_USER	Nombre auténtico del usuario
REMOTE_IDENT	Usuario que realiza una solicitud. Esta variable es definida sólo si la bandera IdentityCheck en NCSA está habilitada, y el cliente soporta el esquema de identificación RFC 931 (proceso ident)
CONTENT_TYPE	El tipo MIME de los datos solicitados, como puede ser "text/html"
CONTENT_LENGTH	La longitud de los datos (en bytes o número de caracteres) proporcionados al programa CGI a través de la entrada estándar
HTTP_FROM	El domicilio electrónico del usuario realizando la solicitud. La mayoría de los visualizadores no soportan esta variable
HTTP_ACCEPT	Lista de los tipos MIME que el cliente puede aceptar
HTTP_USER_AGENT	El tipo de visualizador que el usuario está empleando para realizar la solicitud
HTTP_REFERER	El URL del documento que el cliente solicitó antes de acceder al programa CGI

Obteniendo la información de la forma

Como ya habíamos mencionado, las formas proporcionan una manera para obtener información del usuario y proporcionarla a un programa CGI, como se muestra en la figura.



Interacción entre una forma y un CGI

El navegador de Web permite al usuario seleccionar el tipo de información, y la envía al servidor cuando el botón de envío es presionado.

Una de las maneras en las que una forma envía información al programa CGI es agregando la información al URL después de un signo de interrogación. Seguramente alguna vez se ha visto un URL similar a:

<http://calmecac.dgsca.unam.mx/cgi-bin/programa.pl?fortuna>

Antes del signo de interrogación, todo es familiar. Esta es la manera de ejecutar el CGI programa.pl.

Lo que es nuevo es la parte que viene después del signo "?". Esta información es conocida como "cadena de solicitud" o "query string". Cuando al servidor se le envía un URL con una cadena de solicitud, éste ejecuta el programa CGI identificado en la primera parte del URL (antes del caracter "?") y almacena la parte restante en la variable de ambiente QUERY_STRING que puede ser manipulada por el programa CGI.

Veamos un pequeño ejemplo, en donde se presenta una forma al usuario, solicitando proporcione los datos en la siguiente forma:

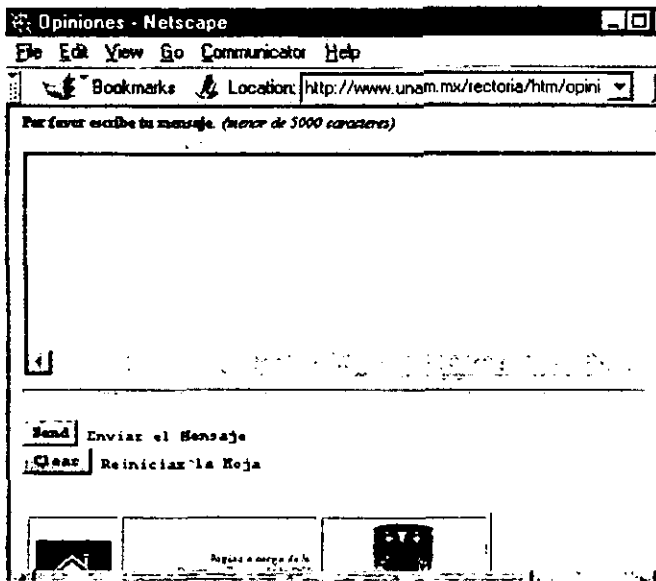
```
<HTML>
<HEAD>
<TITLE>
Ejemplo de forma
</TITLE>
</HEAD>
<BODY>
<H1>
Ejemplo de forma
```

```

</H1>
<HR>
<FORM ACTION="/cgi-bin/ejemplo.pl" METHOD="GET">
Información: <INPUT TYPE="text" NAME="info" SIZE=40>
<P>
<INPUT TYPE="submit" VALUE="Enviar i nformación">
<INPUT TYPE="reset" VALUE="Limpiar forma">
</FORM>
<HR>
</BODY>
</HTML>

```

La apariencia de la forma puede variar entre visualizadores, pero en esencia, la forma tendría este aspecto:



Ejemplo de forma en Netscape

Esta forma consiste de un campo de texto llamado "Información:" y dos botones. El botón con el título "Enviar información" es empleado para enviar los datos proporcionados por el usuario en la forma hacia el programa CGI especificado en el atributo ACTION. El botón con el título "Limpiar forma" elimina la información en el campo de texto.

El atributo METHOD="GET" dentro de la sección <FORM> en parte determina la manera en que los datos son pasados al servidor. Más adelante abordaremos en los métodos existentes para enviar la información al CGI. Por lo pronto tomaremos por omisión el método GET. Ahora, si el usuario escribe la palabra "suerte"

en el campo de texto y luego selecciona el botón de "Enviar información", el visualizador envía la siguiente solicitud al servidor:

```
GET /cgi-bin/ejemplo.pl?info=suerte HTTP/1.0
```

....

El servidor ejecuta el programa llamado ejemplo.pl ubicado dentro del directorio cgi -bin, y coloca la cadena "info=suerte" en la variable de ambiente QUERY_STRING. Podemos pensarlo como que se asigna la información proporcionada por el usuario a la variable "info" (nombre definido en el atributo NAME de la etiqueta <INPUT>). Ahora podemos desarrollar un programa que reciba y maneje estos datos.

```
#!/usr/bin/perl
```

```
print "Content-type: text/plain \n\n";
```

```
$query_string = $ENV{'QUERY_STRING'};  
($campo,$sinformacion) = split (/=/,$query_string);
```

Después de imprimir el tipo de contenido (text/plain) y obtener la cadena de solicitud del arreglo %ENV, se emplea la función split para separar la cadena en dos partes a partir del carácter "=", colocando la primera parte en la variable \$campo y la segunda parte en la variable \$sinformacion (para este caso en particular, \$campo contiene "info" y \$sinformacion contiene "suerte").

Debido a que se emplea el método GET, todos los datos proporcionados a la forma son incluidos en el URL, por lo que podemos tener acceso directo al programa sin necesidad de emplear la forma, empleando el URL:

```
http://calmecac.dgsca.unam.mx/cgi-bin/ejemplo.pl?info=suerte
```

Esto trabaja en forma idéntica a haber llenado la forma y enviarla.

Métodos GET y POST

En el ejemplo anterior, se empleó el método GET para procesar la forma. Sin embargo, existe otro método llamado POST. Al emplear este método, el servidor envía los datos al programa como un paquete de entrada. Esto es, si en el ejemplo anterior, la etiqueta <FORM> fuera:

```
<FORM ACTION="/cgi-bin/ejemplo.pl" METHOD="POST">
```

la siguiente solicitud hubiera sido enviada al servidor:

```
POST /cgi-bin/ejemplo.pl?info=suerte HTTP/1.0
```

....

```
Content-length: 11
```

```
info=suerte
```

El programa que interpreta este método varía, ya que como la información es proporcionada como un paquete, se define la variable de ambiente CONTENT_LENGTH con el tamaño de los datos en número de bytes (o caracteres). De esta manera, es posible leer exactamente la cantidad de datos que provienen de la entrada estándar.

```
#!/usr/bin/perl
```

```
$size_info = $ENV{'CONTENT_LENGTH'};
```

Después se lee el número de bytes especificados en la variable \$size_info de la entrada estándar en la variable \$datos.

```
read(STDIN,$datos,$size_info);
```

Ahora, podemos dividir el contenido de la variable \$datos en \$campo y \$informacion como lo hicimos en el ejemplo anterior con el método GET, con el mismo resultado.

Debido a que es la forma la que determina qué método es empleado (GET o POST), el programador de CGI's no puede controlar con qué método el programa será ejecutado, por lo que en general son diseñados para soportar ambos métodos. Para lograrlo, se hace uso de la variable de ambiente REQUEST_METHOD que almacena el método definido en la forma.

Codificación de los datos

Hasta ahora sólo hemos visto ejemplos con manejo de información muy simple obtenida de la forma. Sin embargo, esta información puede ser mucho más complicada. Debido a que con el método GET la información de la forma es enviada como parte del URL, no se permite que haya espacios o caracteres especiales que no son permitidos en los URL's. Por lo tanto, una codificación especial es empleada. Veamos el siguiente ejemplo, considerando la siguiente forma:

```
<HTML>
<HEAD>
  <TITLE>
    Cuando es tu cumpleaños?
  </TITLE>
</HEAD>
<BODY>
  <H1>
    Cuando es tu cumpleaños?
  </H1>
  <HR>
  <FORM ACTION="/cgi-bin/cumpleanos.pl" METHOD="POST">
    Cumpleaños (con formario dd/mm/aa): <INPUT TYPE="text" NAME="fecha" SIZE=40>
    <P>
    <INPUT TYPE="submit" VALUE="Enviar información">
    <INPUT TUPE="reset" VALUE="Limpiar la forma">
  </FORM>
  <HR>
</BODY>
</HTML>
```

Cuando el usuario envía la información, el cliente genera la siguiente solicitud para el servidor (asumiendo que el usuario proporcionó la fecha 05/11/73):

```
POST /cgi-bin/cumpleanos.pl HTTP/1.0
```

```
...
```

```
Content-length: 18
```

```
fecha=05%2F11%2F73
```

En forma codificada, ciertos caracteres, como los espacios y otros símbolos son reemplazados por sus equivalentes hexadecimales. En este ejemplo, el CGI necesita decodificar estos datos convirtiendo el "%2F" a "/".

La siguiente expresión regular es empleada para decodificar los datos:

```
$datos =~ s/%([\dA-Fa-f][\dA-fa-f])/pack ("C",hex($1))/eg;
```

En este caso, convierte "%2F" en "/".

Salida de los CGI's

Como hemos visto, los programas CGI's son solicitados como cualquier otro documento. La diferencia es que en lugar de responder un documento estático, el servidor ejecuta un programa y responde su salida. En la parte que corresponde al cliente, éste espera una respuesta de la misma manera en la que solicitó el documento, es decir, espera una respuesta que sea capaz de interpretar.

En su forma más simple, la salida de un programa CGI es un documento en texto plano o HTML, que el visualizador muestra como cualquier otro documento de Web. Sin embargo hay otras cosas que se pueden realizar como:

- Enviar gráficas y otros datos binarios
- Indicarle al navegador cuándo almacenar en caché el documento dinámico
- Enviar códigos de estado especiales al navegador
- Indicarle al servidor que se envíe un documento ya existente

Cada una de estas técnicas involucra el manejo de encabezados adicionales desde el programa CGI.

Hasta este punto, hemos tomado la línea de salida "Content-type" como indispensable. Pero éste es sólo un tipo de encabezado que un programa CGI puede emplear. "Content-type" es un encabezado HTTP que indica el tipo de contenido MIME describiendo el formato de los datos que enseguida se envían. Otros encabezados indican:

- El tamaño de los datos
- Otro documento que el servidor debe enviar (esto es, en lugar de regresar un documento dinámico por el programa)
- Códigos de estado HTTP

En la siguiente tabla se muestran los encabezados que son útiles

Encabezado	Descripción
Content-length	La longitud (en bytes) del paquete de salida. Implica datos binarios
Content-type	El tipo de contenido MIME del paquete de salida
Expires	Fecha y hora en la que el documento ya no es válido y debe ser obtenido nuevamente por el navegador
Location	Redirección de servidor (no puede ser enviado como parte de un encabezado completo)
Pragma	Evita o permite que el documento sea almacenado en caché
Status	Estado de la solicitud (no puede ser enviado como parte de un encabezado completo)

Los siguientes encabezados son interpretados sólo por los visualizadores compatibles con Netscape:

Encabezado	Descripción
Refresh	El cliente recarga un documento específico
Set-Cookie	El cliente almacena datos específicos. Normalmente para mantener un registro de datos entre solicitudes

La lista completa de encabezados HTTP se puede consultar en:

<http://www.w3.org/Protocols/HTTP/Object-Headers.html>

Además, hay un par de cosas que se debe saber sobre la sintaxis de un encabezado:

- Las líneas del encabezado no van en un orden especial. En general, los encabezados que se generan a partir de un programa CGI pueden ser colocados en cualquier orden.
- El bloque de encabezados terminan con una línea en blanco. HTTP es un protocolo muy simple. La manera en la que el servidor sabe que se ha terminado un encabezado es buscando una línea en blanco. Todo lo que vaya antes de la línea en blanco se considera como información del encabezado; todo aquello que va después de la línea en blanco se toma como datos. En Perl, la línea en blanco es generada por dos caracteres nueva línea (`\n\n`) que son impresos después de la última línea del encabezado. Si no se incluye la línea en blanco después del encabezado, el servidor asumirá en forma incorrecta que todo el paquete de información es encabezado HTTP y generará un estado de error.

Tipos aceptados y tipos de contenido

Las aplicaciones CGI pueden regresar prácticamente cualquier tipo de documento dinámico, siempre y cuando el cliente pueda manejarlo correctamente. Puede regresar un archivo de texto plano, un archivo HTML o puede enviar documentos PostScript, PDF, SGML, etc.

Esta es la razón por la que el cliente envía una lista de "tipos aceptados" "accept types" que soporta, directa o indirectamente a través de aplicaciones de apoyo, al servidor cuando envía una solicitud. El servidor almacena esta información en la variable de ambiente `HTTP_ACCEPT`, y el programa CGI puede verificar esta variable para asegurar que el archivo enviado presenta un formato que el visualizador puede manejar.

También es él por qué cuando se regresa un documento, el programa CGI necesita emplean el encabezado "Content-type" para notificar al cliente el tipo de datos que se están enviando, de tal forma que el visualizador pueda dar formato y mostrar el documento apropiadamente.

El encabezado Content-length

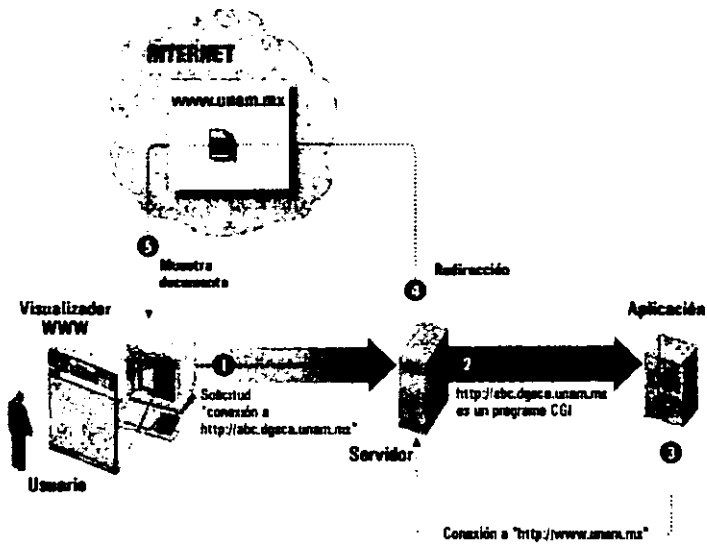
Como se ha observado, no estamos limitados a trabajar sólo con texto HTML (definido por el tipo MIME `text/html`), sino que podemos responder documentos formateados de diversas maneras, como texto plano, imágenes GIF o JPG, etc. Cuando se va a enviar una imagen, lo primero que se debe hacer es que el programa CGI indique (con el encabezado `Content-type`) que el elemento a regresar es una imagen. Esto le indicará al visualizador cómo mostrar el elemento. El siguiente elemento a definir es el encabezado `Content-`

length, que notifica la cantidad de datos que serán enviados. Esto permite evitar errores de fin de datos no esperados, principalmente cuando se manejan datos binarios.

Redireccionamiento

Hasta ahora hemos visto que los CGI's regresan documentos dinámicos creados en el momento. Sin embargo, otra de las cosas que puede hacer un CGI es indicarle al servidor que es necesario recuperar un documento estático. Esto es conocido como redireccionamiento.

Para realizar esta operación, es necesario enviar el encabezado Location para indicarle al servidor qué documento enviar. El servidor recuperará el documento especificado, dando la apariencia que el cliente no ha solicitado un CGI sino el documento (Ver la siguiente figura).



Redireccionamiento

Un uso común de esta característica es regresar un documento genérico que contiene información estática. Por ejemplo, digamos que se tiene una forma que se desea sea llenada por el usuario, y se quiere mostrar un mensaje de agradecimiento después de que alguien complete la forma. Se puede tener el programa CGI que cree y muestre el mensaje cada vez que es ejecutado. Pero una manera más eficiente puede ser diseñar el programa de tal forma que envíe instrucciones al servidor y le indique regresar un archivo que contiene el mensaje genérico, a través de la instrucción:

```
print "Location: /archivo.html \n\n";
```

donde archivo.html es el nombre del archivo que se desea mostrar.

Encabezados Expires y Pragma

La mayoría de los visualizadores almacenan internamente (caché) los documentos que se obtienen. Esto es una característica muy útil que ahorra recursos; el visualizador no tiene que recuperar el documento cada vez que se desea verlo. Sin embargo, puede ser un problema cuando se está trabajando con documentos dinámicos creados por programas CGI. Una vez que el visualizador muestra un documento dinámico producido por un CGI, lo almacenará. La siguiente vez que se trate de tener acceso al mismo archivo, el visualizador no hará una solicitud al servidor, sino que mostrará el documento que mantiene en su caché. Sólo hasta que se indique explícitamente que es necesario recargar el documento es cuando se ejecutará el programa nuevamente.

Afortunadamente, hay una solución a este problema. Si no se desea que un documento dinámico sea almacenado, se puede emplear el encabezado Expires o el encabezado Pragma para indicarle al cliente que no guarde el documento. Sin embargo, algunos visualizadores no manejan correctamente estos encabezados, por lo que es necesario tener cuidado.

Códigos de estado

Los códigos de estado son empleados por el protocolo HTTP para informar el estado de una solicitud. Por ejemplo, si un documento no existe, el servidor regresa un código de estado 404 al visualizador. Si un documento ha sido movido, un código de estado 301 es presentado.

Los programas CGI pueden enviar información de estado como parte del documento dinámico.

Un encabezado Status consiste de un código numérico de tres dígitos, seguido por una cadena reportando el código. Un valor dentro de 200 indica éxito, mientras que los valores dentro de 400 constituyen una solicitud incorrecta. En adición a estos dos rangos, hay otros códigos de estado que se pueden emplear en diversas situaciones, que van desde una solicitud no autorizada o prohibida hasta errores internos. La siguiente tabla muestra una lista de algunos códigos de estado comúnmente utilizados.

Código de estado	Mensaje
200	Success
204	No Response
301	Document Moved
401	Unauthorized
403	Forbidden
404	Not Found
500	Internal Server Error
501	Not Implemented

Para una relación completa de códigos de estado, referirse a:

<http://www.w3.org/Protocols/HTTP/HTRESP.html>

Desafortunadamente no todos los visualizadores soportan todo el conjunto de códigos.

Un código que merece especial atención es el 204, que produce un "no response". En otras palabras, el visualizador no mostrará una nueva página si el programa CGI regresa un código de estado 204.

Este código puede ser empleado cuando se trabaja con formas o mapas de imágenes. Por ejemplo, si el usuario proporciona un valor no válido en uno de los campos o selecciona una sección no definida en un

mapa de imágenes, se puede contestar un código de estado 204 para indicar al cliente que no muestre una nueva página.

Formas y CGI

Como hemos visto, Common Gateway Interface (CGI) es empleado principalmente para dos propósitos: la obtención de datos y la comunicación interactiva. Se pueden presentar registros o solicitudes en línea a través del uso de formas. También son empleadas para crear un medio interactivo entre el usuario y el servidor de Web. Por ejemplo, una forma puede solicitar al usuario que seleccione un documento de un menú, para que entonces el servidor regrese el documento solicitado.

La principal ventaja de las formas es que se pueden emplear para crear una vista para numerosos puentes (con bases de datos y otros servicios de información) que pueden ser accedidos por cualquier cliente sin necesidad de preocuparse de la dependencia de plataformas. En otro sentido, hay limitaciones con las actuales implementaciones:

La interfaz no soporta cualquier tipo de datos además del tipo genérico "text". La siguiente especificación de HTML puede contemplar otro tipo de datos, como es "int", "date", "float" y "url".

La información proporcionada por el usuario no puede ser validada por el cliente; el usuario tiene que seleccionar el botón de envío y el programa CGI en el servidor verifica que la entrada sea válida.

Etiquetas HTML

Una forma consiste de dos partes distintas: el código HTML y el programa CGI. Las etiquetas HTML crean la representación visual de la forma, mientras que el programa CGI interpreta (o procesa) la información contenida dentro de la forma.

Etiqueta FORM

El inicio de una forma simple es:

```
<FORM ACTION="/cgi-bin/programa.pl" METHOD="POST">
```

La etiqueta <FORM> inicia la forma. Un documento puede consistir de varias formas, pero no pueden estar anidadas, es decir, una forma no puede ser colocada dentro de otra forma.

Los dos atributos dentro de la etiqueta <FORM> (ACTION y METHOD) son muy importantes. El atributo ACTION especifica el URL del programa CGI que procesará la información de la forma. Evidentemente no se está limitado a emplear programas CGI en el propio servidor para interpretar la información; se puede especificar un URL hacia un servidor remoto si el programa deseado se encuentra a disposición en alguna parte.

El atributo METHOD especifica la manera en la que el servidor enviará la información al programa. POST envía los datos a través de la entrada estándar, mientras que GET pasa la información a través de variables de ambiente. Si ningún método es especificado, el servidor emplea por omisión el método GET.

Campos de texto y de contraseña

La mayor parte de los elementos de una forma son implementados empleando la etiqueta <INPUT>. El atributo TYPE en <INPUT> determina qué tipo de entrada está siendo solicitada. Diversos tipos de elementos están definidos: texto, contraseña, botones de selección y menús. Las siguientes líneas son ejemplos de entradas de texto simple.

```
Nombre: <INPUT TYPE="text" NAME="usuario" SIZE=40><BR>
Edad: <INPUT TYPE="text" NAME="edad" SIZE=3 MAXLENGTH=3><BR>
Contraseña: <INPUT TYPE="password" NAME="contraseña" SIZE=10><BR>
```

En este caso, dos campos de texto y un campo de contraseña son creados empleando los argumentos "text" y "password", respectivamente. El campo de contraseña es básicamente el mismo que un campo de texto excepto por el hecho de que los caracteres introducidos serán mostrados como asteriscos o puntos. Si no se define un atributo TYPE, un campo de texto es creado por omisión.

El atributo NAME define el nombre del elemento de entrada en particular. No es mostrada por el visualizador, pero es empleada como etiqueta del dato cuando es transferido al programa CGI. Por ejemplo, el primer campo de entrada tiene el atributo NAME="usuario". Si alguien escribe "lalo" en el primer campo de entrada, entonces una parte de los datos enviados por el visualizador será:

```
usuario=lalo
```

El programa CGI puede obtener esta información y procesarla como sea necesario.

El atributo opcional VALUE puede ser empleado para definir un valor por omisión para el campo. Esta cadena puede ser redefinida por el usuario. Otros atributos opcionales son SIZE y MAXLENGTH. SIZE es el tamaño físico del elemento de entrada; el campo puede desplazarse si la entrada supera el tamaño definido. El tamaño por omisión es 20 caracteres. MAXLENGTH define el número máximo de caracteres que serán aceptados por el visualizador; por omisión no hay límite.

Existe otro tipo de campo de texto. Es llamado el campo oculto (hidden) y permite almacenar información en la forma. El cliente no muestra el campo. Estos campos son muy útiles para transferir información de un CGI a otro.

Botones de envío e inicialización

Dos de los más importantes tipos de la etiqueta <INPUT> son envío (Submit) e inicialización (Reset).

```
<INPUT TYPE="submit" VALUE="Enviar la información">
<INPUT TYPE="reset" VALUE="Limpiar la forma">
```

Casi todas las formas presentan estos dos botones. El botón de envío permite pasar toda la información de la forma al programa CGI especificado por el atributo ACTION. Sin este botón la forma sería inútil debido a que nunca podría ejecutar el programa CGI.

Los visualizadores proporcionan por omisión las etiquetas Submit y Reset para cada uno de los botones respectivamente. Sin embargo, pueden ser redefinidas con el atributo VALUE.

Además se pueden tener múltiples botones de envío:

```
<INPUT TYPE="submit" VALUE="opcion" VALUE="Opcion 1">
<INPUT TYPE="submit" VALUE="opcion" VALUE="Opcion 2">
```

Si el usuario selecciona en "Opcion 1", el programa CGI recibirá:

```
opcion=Opcion 1
```

Además se pueden tener imágenes como botones:

```
<INPUT TYPE="image" SRC="/icons/button.gif" NAME="instalar" VALUE="Instalacion"">
```

Cuando se selecciona la imagen, el visualizador envía las coordenadas en donde se seleccionó la imagen.

```
instalar.x=250&instalar.y=20
```

Nótese que la información de cada campo es delimitada por el caracter "&". Si se emplea un visualizador de sólo texto, lo que se envía es:

```
instalar=Instalacion
```

El botón Reset limpia la información proporcionada por el usuario. Esto se puede hacer cuando se quiere iniciar con una forma limpia.

Botones de selección y de relación

Los botones de selección y de relación son empleados típicamente cuando se le muestran al usuario varias opciones.

Un botón de selección crea botones cuadrados (o cajas) que pueden ser "encendidas" o "apagadas". A continuación se muestra un ejemplo de cómo crear cuatro botones de selección:

```
<FORM ACTION="/cgi-bin/programa.pl" METHOD="POST">
¿Qué película desea ordenar?: <BR>
Amadeus <INPUT TYPE="checkbox" NAME="amadeus">
El Ultimo Emperador <INPUT TYPE="checkbox" NAME="emperador">
Gandhi <INPUT TYPE="checkbox" NAME="gandhi">
La Lista de Schindler <INPUT TYPE="checkbox" NAME="schindler">
<BR>
```

Si el usuario selecciona o "enciende" uno de los botones de selección y envía la forma, el visualizador asigna el valor "on" al nombre de variable. Por ejemplo, si alguien selecciona el botón de Gandhi en el ejemplo anterior, el visualizador enviará:

```
gandhi=on
```

Este valor puede ser alterado empleando el atributo VALUE:

```
Gandhi <INPUT TYPE="checkbox" NAME="gandhi" VALUE="yes">
```

Ahora, cuando el botón de Gandhi es seleccionado, el visualizador enviará:

gandhi=yes

Un botón de selección no tiene relación con los demás. Cualquier número de ellos pueden ser seleccionados al mismo tiempo. Los botones de relación difieren de los botones de selección porque en el primero sólo un valor puede ser habilitado a la vez. Por ejemplo:

```
¿Cómo será la forma de pago? : <BR>
Master Card: <INPUT TYPE="radio" NAME="pago" VALUE="MC" CHECKED><BR>
Visa: <INPUT TYPE="radio" NAME="pago" VALUE="Visa"><BR>
American Express: <INPUT TYPE="radio" NAME="pago" VALUE="AE"><BR>
Discover: <INPUT TYPE="radio" NAME="pago" VALUE="Discover"><BR>
</FORM>
```

Algunos lineamientos para lograr que los botones de relación funcionen apropiadamente:

Todas las opciones deben llevar el mismo atributo NAME (en el ejemplo "pago"). Esto debido a que el visualizador sabe que deben estar agrupados y de esta manera asegurar que sólo una opción esté empleando el NAME asignado en un momento dado.

Mientras que con los botones de selección el proporcionar diferentes atributos VALUE es sólo cuestión de gustos, con los botones de relación el tener diferentes atributos VALUE es crucial para tener resultados con sentido. Sin un atributo VALUE específico, no importa qué elemento es seleccionado, el visualizador asignará la cadena "on" a la variable con NAME pago. El CGI no tiene forma de saber qué elemento fue seleccionado. Por lo tanto cada elemento en los botones de relación necesita tener asignado un atributo VALUE distinto para asegurar que el programa conozca cuál fue seleccionado.

Para ambos casos, el atributo CHECKED determina qué elemento debe estar habilitado por omisión. En el ejemplo de los botones de relación, la opción "Master Card" tiene el atributo CHECKED, haciendo que efectivamente esta sea la opción seleccionada por omisión.

Menús y listas deslizables

Los menús y las listas deslizables son empleadas generalmente para presentar un gran número de opciones al usuario. El siguiente es un ejemplo de un menú:

```
<FORM ACTION="/cgi-bin/programa.pl" METHOD="POST">
Seleccione una forma de pago:
<SELECT NAME="tarjeta" SIZE=1>
<OPTION SELECTED> Master Card
<OPTION>Visa
<OPTION>American Express
<OPTION>Discover
</SELECT>
```

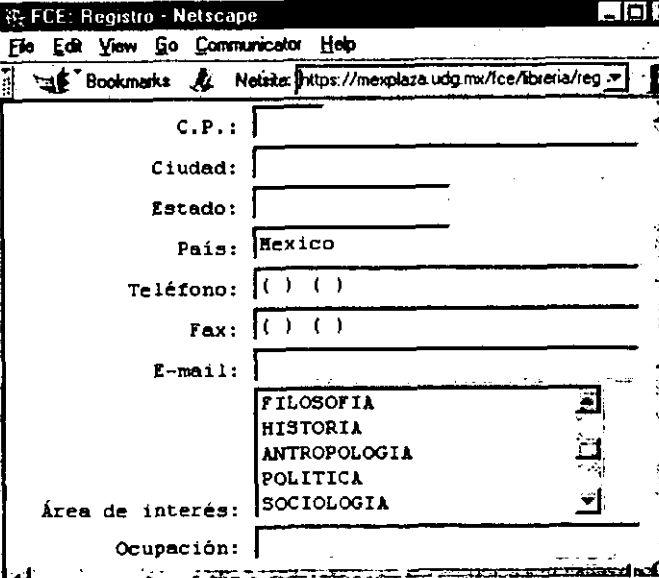
Los menús de opciones y las listas deslizables son creadas empleando la etiqueta SELECT, de la cual existe una de apertura y otra de finalización. El atributo SIZE determina si un menú o una lista se muestra. Un valor de 1 produce un menú, y un valor mayor o igual a 2 produce una lista deslizable, en cuyo caso el número representa la cantidad de elementos que serán visibles a un mismo tiempo.

Una selección de un menú o una lista es agregada empleando la etiqueta OPTION. El atributo SELECTED de la etiqueta OPTION nos permite determinar una selección por omisión.

Ahora, tenemos un ejemplo con lista deslizable

```
<SELECT NAME="libros" SIZE=3 MULTIPLE>  
<OPTION SELECTED>TCP/IP Network Administration  
<OPTION>Linux Network Administrators Guide  
<OPTION>DNS and BIND  
<OPTION>Computer Security Basics  
<OPTION>System Performance Tuning  
</SELECT>  
</FORM>
```

Este ejemplo crea una lista deslizable con tres elementos visibles y la capacidad de seleccionar múltiples opciones. (El atributo MULTIPLE especifica que más de un elemento puede ser seleccionado). La forma resultante tiene el siguiente aspecto:



The screenshot shows a Netscape browser window titled "FCE: Registro - Netscape". The address bar displays "https://mexplaza.udg.mx/fce/libreria/reg". The form contains the following fields:

- C.P.:
- Ciudad:
- Estado:
- País: Mexico
- Teléfono: () ()
- Fax: () ()
- E-mail:
- Área de interés: A dropdown menu with the following options: FILOSOFIA, HISTORIA, ANTROPOLOGIA, POLITICA, and SOCIOLOGIA.
- Ocupación:

Forma con menús y lista deslizable

Enviando los datos al servidor

Los visualizadores emplean el tipo MIME `application/x-www-form-urlencoded` para codificar los datos de la forma.

Primero, el nombre de cada elemento de la forma - especificado por el atributo `NAME` - se iguala con el valor introducido por el usuario para crear un par variable-valor. Por ejemplo, si el usuario proporciona "30" cuando se le pregunta por su edad, el par variable-valor será `edad=30`. Cada par variable-valor es separado por el carácter "&".

Segundo, debido a que los nombres de variables de los elementos de la forma y los datos obtenidos de ella están en texto estándar, es posible que contengan caracteres que puedan confundir al visualizador. Para prevenir posibles errores, el esquema de codificación traduce todos los caracteres especiales a su correspondiente valor hexadecimal. Estos caracteres especiales incluyen caracteres de control y ciertos símbolos alfanuméricos. Por ejemplo, la cadena "Gracias por su ayuda" será convertida a "Gracias%20por%20su%20ayuda". Este proceso será repetido para cada par variable-valor para crear la cadena de solicitud.

Cada par variable-valor lleva los campos de texto y de contraseña, la información proporcionada por el usuario constituirán el valor. Si no se proporciona información, el par variable-valor será enviado de todas formas, pero seguramente será ignorado.

El programa CGI tendrá entonces que decodificar la información antes de procesar la información de la forma. El esquema de codificación es el mismo para GET y POST.

GET vs. POST

Como lo mencionamos, hay dos métodos para enviar los datos de una forma: GET y POST. La principal diferencia entre ambos es la manera en que los datos son pasados al programa CGI. Si el método GET es empleado, la cadena de solicitud es agregada al URL del programa cuando el cliente envía la solicitud al servidor. Esta cadena de solicitud puede ser accedida empleando la variable de ambiente `QUERY_STRING`. Aquí se muestra un ejemplo de solicitud GET por parte del cliente:

```
GET /cgi-bin/programa.pl?usuario=Rafael%20Perez&edad=25&contrasenia=prueba HTTP/1.0
Accept: www/source
Accept: text/html
Accept: text/plain
User-Agent: Lynx/2.4
```

Este método presenta ventajas y desventajas. La principal ventaja es que se puede acceder al programa con una solicitud sin necesidad de emplear la forma. Básicamente se pasan los parámetros al programa. Por ejemplo, si se desea enviar la solicitud anterior al programa directamente, se puede hacer:

```
<HREF="/cgibin/programa.pl?usuario=Rafael%20Perez&edad=25&contrasenia=prueba">Programa
CGI</A>
```

Si se emplea este método, es necesario ser cuidadoso, ya que el visualizador o el servidor pueden truncar aquellos datos que excedan una determinada cantidad de caracteres.

Ahora, una solicitud POST simple:

```
POST /cgi-bin/programa.pl HTTP/1.0
Accept: www/source
```

Accept: text/html
Accept: text/plain
User-Agent: Lynx/2.4
Content-type: application/x-www-form-urlencoded
Content-length: 47

usuario=Rafael%20Perez&edad=25&contrasenia=prueba

La principal ventaja del método POST es que la longitud de la solicitud es ilimitada y por lo tanto no es necesario preocuparse sobre el truncamiento por parte del servidor o el cliente. Para obtener los datos enviados con el método POST, el programa CGI lee de la entrada estándar.

Proceso de Decodificación

Para obtener la información contenida en la forma, un protocolo de decodificación debe ser aplicado a los datos. Primero, el programa debe determinar la manera en la que el cliente envió los datos. Esto se hace examinando el valor de la variable de ambiente REQUEST_METHOD. Si el valor indica una solicitud GET, la cadena de solicitud debe ser obtenida de la variable de ambiente QUERY_STRING. De otra manera, si es una solicitud POST, el número de bytes especificados en la variable de ambiente CONTENT_LENGTH debe ser leído de la entrada estándar. A continuación el algoritmo para la decodificación de los datos:

1. Determinar el protocolo de solicitud (GET o POST) verificando la variable de ambiente REQUEST_METHOD
2. Si es el protocolo GET, leer la cadena de solicitud de la variable de ambiente de QUERY_STRING
3. Si es el protocolo es POST, determinar el tamaño de la solicitud empleando la variable de ambiente CONTENT_LENGTH y leer esa cantidad de datos desde la entrada estándar
4. Dividir la cadena de solicitud a partir del carácter "&", que separa los pares variable-valor (recordar que el formato es variable1=valor&variable2=valor...)
5. Decodificar los caracteres hexadecimales y "+" en cada par variable-valor
6. Crear una tabla de variable-valor con la variable como índice

La razón por la que el programa necesita verificar el protocolo de solicitud es para poder diseñarlo correctamente. Sin embargo, en ocasiones es recomendable desarrollar el programa de tal forma que sea capaz de manejar ambos protocolos de solicitud.

APÉNDICE B

APÉNDICE B

PROTOCOLO Z39.50

El Protocolo ANSI/NISO Z39.50 es empleado para la Recuperación de Información en la Infraestructura y Servicios de Información.

Las computadoras y las redes de telecomunicaciones ofrecen un potencial enorme para compartir datos en formato electrónico. El reto ha sido desarrollar mecanismos que puedan estandarizar la comunicación entre los sistemas de cómputo existentes. NISO responde a este reto estableciendo el Comité de Estándares en 1979 para trabajar en un protocolo de recuperación de información. Este trabajo culminó con Z39.50- 1988. También, a finales de los 70's se funda el Consejo en Recursos Bibliotecarios (CLR por sus siglas en inglés) llevando a cabo el desarrollo de un protocolo experimental como parte del proyecto Linked System Project (LSP) para búsquedas en bases de datos bibliográficas y transferir registros entre la Biblioteca del Congreso, the Online Computer Library Center (OCLC), el Grupo de Desarrollo en Bibliotecas (the Research Libraries Group RLG), y la Red de Bibliotecas Washington (ahora Red de Bibliotecas Western (Western Library Network WLN). Este Protocolo de Sistemas Ligados proporcionó las bases para el protocolo Z39.50.

¿Qué hace Z39.50?

Z39.50 reconoce que la recuperación de información consiste de dos componentes principales: selección de información basados en algún criterio y la recuperación de la información; Z39.50 proporciona un lenguaje para ambas actividades. Estandariza la manera en la que el cliente y el servidor se comunican y operan entre sí, aún cuando existen diferencias entre los sistemas de cómputo, aplicaciones de búsqueda y bases de datos.

Una serie de mensajes son intercambiados entre el cliente y el servidor (definido por lo que el estándar nombra Initialization Facility), se establece la conexión, se inicia una sesión Z39.50, se negocian expectativas y límites para todas las actividades que van a ocurrir (es decir, tamaño máximo de los registros que serán transferidos del servidor al cliente, la versión de protocolo aceptada, etc.) Después de que estos arreglos son realizados, el cliente envía su solicitud. El cliente Z39.50 traduce la solicitud con una representación estándar y la envía al servidor Z39.50 (definida por la Search Facility). El servidor ejecuta la búsqueda sobre la(s) base(s) de datos y un conjunto de resultado es generado. El cliente puede solicitar entonces los registros del conjunto resultado o solicitar al servidor un proceso adicional del conjunto resultado (definido en la Retrieval Facility). Una vez que el cliente recibe los registros, éste puede procesar los registros y mostrarlos al usuario. Las extensiones de que un cliente pueda realizar proceso adicional en los registros obtenidos (es decir, combinar registros desde varias búsquedas separadas) dependerá de la interfaz del usuario que es totalmente independiente de la aplicación Z39.50.

Cada base de datos que reside en un servidor de información puede tener características únicas. Por ejemplo, las bases de datos pueden diferir en la manera en que los datos son almacenados y en los puntos de acceso disponibles para la búsqueda. Los registros en cada base de datos pueden tener distintas estructuras y consistir de diversos elementos. El objetivo de Z39.50 es proporcionar comunicación computadora a computadora en términos estándares y mutuamente entendibles, así como soporte a la transferencia de datos entre sistemas independientemente de su estructura, contenido o formato de datos. Sin embargo, en implementaciones individuales, los servidores pueden estar limitados por formatos específicos de datos que pueden ser exportados y los puntos de acceso que son proporcionados para la búsqueda.

Buscando en una base de datos

Cuando se realiza una búsqueda en una base de datos, el cliente envía la solicitud al servidor. La solicitud contiene los términos buscados (es decir, aquellos términos que el usuario ha identificado para ser buscados en los puntos de acceso de la base de datos) y atributos de tales términos (es decir, especificar el término como "autor", "título", etc.). Las solicitudes pueden incluir diversos tipos de atributos. Por ejemplo, si un usuario desea buscar por el nombre del autor, un atributo "uso" especifica el término a buscar como "autor". Si el usuario desea buscar por todos aquellos libros publicados después de cierta fecha, un atributo "use" especifica al término "fecha de publicación" y un atributo "relation" especifica que el usuario desea todas las fechas de publicación "mayores a" una fecha particular. El estándar Z39.50 ANSI/NISO enumera todos los tipos de atributos y sus valores en el conjunto de atributos registrados. El conjunto de atributos estándares y mutuamente reconocidos permite a los desarrolladores tener una base común para la comunicación entre sistemas.

Después de que el servidor realiza una búsqueda en una base de datos, crea un conjunto resultado consistente en aquellos registros que cumplen con el criterio de la solicitud. Los clientes pueden solicitar que los servidores regresen aquellos registros contenidos en el conjunto resultado o empleen el conjunto resultado como argumento en búsquedas subsecuentes.

Recuperando registros desde la Base de Datos

Cuando un usuario quiere que se muestren los registros del conjunto resultado, Z39.50 proporciona opciones sobre qué elementos (es decir, conjunto de elementos) de los registros en la base de datos un usuario puede hacer una solicitud. Además proporciona opciones sobre el formato de transferencia del registro (es decir, una sintaxis de registro) del servidor al cliente. Z39.50 registra conjuntos de nombres estándares de elementos y sintaxis de registros para soportar la comunicación cliente/servidor en este aspecto de la recuperación de información.

Otras características del Estándar que soporta la Recuperación de Información

Además de estandarizar los acuerdos iniciales entre los dos sistemas al iniciar la sesión, el envío de la solicitud, y la recuperación de los registros resultantes, Z39.50 también incluye características que permiten:

- Que el cliente solicite al servidor eliminar uno o más conjuntos de resultados que han sido creados a partir de la búsqueda en la base de datos (Result-set-delete Facility)
- Que el servidor solicite al cliente proporcionar información tal como una contraseña para autenticación (Access Control Facility)
- Que el cliente y el servidor soliciten y proporcionen información relacionada con los recursos que serán o están siendo empleados para llevar a cabo una consulta específica, tal como el costo de una búsqueda en particular (Accounting/Resource Control Facility).
- Maneras de que cualquiera (cliente o servidor) finalice una sesión

Mejorando el Estándar

Aunque Z39.50 ha sido empleado inicialmente para sistemas que manejan datos bibliográficos (es decir, catálogos de acceso público en línea), el protocolo es actualmente muy general y extensible. Manteniéndose en el ambiente de información en red dinámica en la que fue desarrollado, Z39.50 crece y se expande en funcionalidad y riqueza. Las revisiones a los estándares, sin embargo, reconocen y soportan a la base instalada de versiones anteriores. Lo que proporciona estabilidad al estándar, compatibilidad entre versiones y protege la inversión en productos Z39.50 mientras se extiende la utilidad del estándar.

Basado en los requerimientos de los desarrolladores y usuarios de Z39.50, las nuevas versiones contienen nuevas características y mejoras. No todas las implementaciones deciden incluir todas las características. Las siguientes indicaciones muestran mejoras que estarán disponibles para los desarrolladores de Z39.50.

- **Ordenamiento:** Permite al cliente solicitar al servidor que ordene el conjunto de resultados de acuerdo a un criterio proporcionado por el usuario
- **Revisión:** Permite al cliente revisar la lista de términos (es decir, los valores de los puntos de acceso) disponibles en una base de datos o grupo de ellas
- **Servicios extendidos:** Define un conjunto de tareas u operaciones que el cliente puede solicitar que el servidor ejecute, como es salvar un conjunto de resultados para su uso posterior, ejecutar búsquedas en un esquema periódico, exportar los registros en un conjunto resultado, ordenar los documentos y solicitud de impresión
- **Explicación:** Permite al cliente obtener información sobre la implementación de un servidor, incluyendo la base de datos disponible para búsqueda, restricciones en el uso del servidor, horas de operación y disponibilidad y, una amplia gama de información importante que el cliente puede utilizar para facilitar la recuperación de información efectiva con un servidor en particular
- **Segmentación:** Proporciona una transferencia efectiva de partes de un registro cuando el registro entero excede el tamaño de transferencia máximo negociado entre el cliente y el servidor; esto es especialmente crítico para bases de datos de imágenes y servicios multimedia
- **Búsqueda aproximada:** Tipo de búsqueda que permite al cliente realizar búsquedas de proximidad específicas
- **Nueva Sintaxis de Registros:** Define diversas sintaxis de registro que el objetivo puede emplear para empaquetar los registros de la base de datos para transferir al origen incluyendo:
 - Sintaxis Simple sin Estructura de Registro de Texto (SUTRS) empleado para datos textuales y permite que el cliente muestre los datos recuperados con poco o nulo proceso
 - Sintaxis de registro OPAC, incluye datos de circulación y préstamo
 - Sintaxis de Registro Genérico un formato de propósito general para empaquetar registros de complejidad variada con datos potencialmente arbitrarios en campos individuales

Se anticipa que como Z39.50 está en pleno desarrollo soportará servicios de recuperación de información relacionados con servicios como entrega de documentos y ofrecerá una nueva flexibilidad en el rango de información que puede ser manejada por el estándar.

APÉNDICE C

APÉNDICE C

Introducción al SSL y certificación utilizando SSLeay

Introducción

El servicio de Web transfiere información del servidor al cliente de forma transparente, es decir, lo que viaja en la red es la información que se está enviando tal cual del navegador al servidor y viceversa, si se desea que los datos viajen de una forma no transparente se tiene que recurrir al cifrado de información.

Para trabajar con información cifrada se requiere trabajar con un protocolo especial llamado SSL (Secure Sockets Layer).

Certificados

Aunque la persona pudo haber enviado un mensaje privado, firmado y asegurando la integridad del mensaje, ella todavía necesita asegurarse de que ella esta realmente comunicándose con el banco. Esto significa que necesita estar segura de que la llave pública que esta usando corresponde a la llave privada del banco. Y viceversa, el banco también necesita verificar que el mensaje firmado realmente corresponde a la firma de la persona. Si cada parte tiene un certificado el cual valida la identidad del otro, confirma la llave pública y es firmado por una institución que da fe de la autenticidad, entonces ellos estarán seguros que se están comunicando con quien cada uno piensa que es. La institución que da fe de la veracidad de sus identidades se denomina Autoridad certificadora, y los certificados son usados para la autenticación.

Contenido de un certificado

Un certificado asocia una llave pública con la identidad real de un individuo, un servidor o alguna otra entidad, conocido como el individuo. Como se muestra en la Tabla 1, la información relativa al individuo incluye información sobre su identificación (nombre distintivo), y la llave pública. También incluye la identificación y firma de la Autoridad certificadora que emite el certificado (emisor) y el periodo durante el cual el certificado es válido. Puede tener información adicional (o extra) también como información administrativa del uso de la Autoridad certificadora, tal como el número de serie.

Individuo	Nombre distintivo, Llave pública
Emisor	Nombre distintivo, Firma
Periodo de valides	No antes de..., No después de...
Información administrativa	Versión, número de serie
Información extra	

Tabla 1 - Información del Certificado

Un nombre distintivo es usado para proveer una identidad en un contexto específico, por ejemplo, un individuo pudiera tener un certificado personal al igual que otro para su identificarse como un empleado. Los nombres distintivos están definidos por el estándar X.509, el cual define los campos, nombre de campos y abreviaciones utilizadas para referirse a los campos (ver Tabla 2).

Campo	Abreviación	Descripción	Ejemplo
Common Name	CN	Nombre que se certifica	CN = Miguel Estrada
Organization or Company	O	Nombre está asociado con dicha organización.	O = U.N.A.M.
Organization Unit	OU	Nombre está asociado con dicha unidad organizacional, tal como un departamento.	OU = Coordinación de Servicios de Red
City/Locality	L	Nombre esta localizado en esta ciudad.	L = Ciudad de México
State/Province	SP	Nombre esta Localizado en este Estado/Provincia	SP = D.F.
Country	C	Nombre esta localizado en este país (código ISO)	C = MX

Tabla 2 - Información de un nombre distintivo

Una Autoridad Certificadora puede definir una política específica la cual distinguirá qué campos son opcionales y cuales son necesarios. Puede también establecer requerimientos sobre el contenido de los campos, para que los usuarios puedan certificarse. Como un ejemplo, un navegador Netscape requiere que el campo Common Name para un certificado representando un servidor tenga un nombre el cual corresponda a una expresión regular para el nombre de dominio del ese servidor, tal como *.unam.mx.

El formato binario de un certificado esta de finido usando la notación ASN.1. Esta notación dice el cómo especificar los contenidos y las reglas de codificación para definir cómo esta información se traduce a un formato binario. El binario cifrado del certificado está definido usando las "Reglas de Cifrado Distintivo" (DER - Distinguished Encoding Rules), las cuales están basadas en las "Reglas Básicas de Cifrado" (BER - Basic Encoding Rules). Para aquellas transmisiones en las cuales no se pueden utilizar binarios, la forma binaria puede ser traducida en formato ASCII usando el cifrado base 64. Esta versión de cifrado es llamada cifrado PEM, cuando es colocado entre las siguientes líneas:

```
----- BEGIN CERTIFICATE -----
----- END CERTIFICATE -----
```

Autoridades certificadoras

Primero se verifica la información en una solicitud de certificado antes de garantizar el mismo, la Autoridad certificadora asegura la identidad del dueño de la llave privada. Es decir, si la persona solicita un certificado personal, la Autoridad certificadora debe asegurarse primero que esta persona es realmente la persona que la solicitud afirma.

Certificados encadenados

Una Autoridad certificadora puede también emitir un certificado para otra Autoridad certificadora. Cuando examinamos un certificado, esta persona puede necesitar examinar el certificado del que lo creo, de cada Autoridad certificadora padre, hasta encontrar alguno que sea de su confianza. Ella puede decidir confiar únicamente en una cadena limitada de generadores de certificados, para reducir su riesgo de un "mal" certificado en la cadena.

Creando un Nivel Raíz de Autoridad Certificadora

Como se apreció rápidamente, todo certificado requiere de alguien que lo emita para verificar la validez de la identidad del individuo certificado, hasta el nivel más alto: la Autoridad certificadora. Esto presenta un problema: Ya que ésta es quién verifica la emisión de los certificados para el nivel más alto de autoridad, ¿no existe quién la certifique? En este único caso, el certificado es "auto-firmado", es decir, quién emite el certificado es el mismo que el individuo. Como resultado de lo anterior, uno debe proceder con mucho cuidado en confiarse a sí mismo un certificado auto-firmado. La amplia publicación de una llave pública por la autoridad raíz reduce el riesgo en confiar dicha llave-- puede ser obvio cuando alguien publica una llave anunciando ser la autoridad. Los navegadores están prediseñados a confiar en autoridades certificadoras muy reconocidas.

Un número de compañías, tales como VeriSign, se tienen establecidas así mismas como Autoridades certificadoras. Estas compañías proveen los siguientes servicios:

- * Verificación de solicitudes de certificados.
- * Procesamiento de las solicitudes de certificados.
- * Emisión y administración de certificados.

Es posible crear su propia Autoridad certificadora. Si bien el riesgo en el ambiente de Internet existe, puede ser útil en una Intranet donde la organización puede fácilmente verificar las identidades de los individuos y los servidores.

Administración de certificados

Establecer una Autoridad certificadora es una responsabilidad la cual requiere una sólida estructura tanto administrativa, técnica y directiva. Las Autoridades certificadoras no solo emiten certificados, ellas pueden también administrarlos -- esto es, pueden determinar cuanto tiempo los certificados son válidos, cuando los renuevan, y mantienen listas con los certificados que han emitido pero que no son ampliamente válidos (listas de revocación de certificados, o CRL - Certificate Revocation Lists). Se requiere que esta persona tenga un certificado como empleada, pero cuando deje la compañía su certificado deberá ser

revocado. Ya que los certificados son objetos que se transmiten, es imposible decir solamente que han sido revocados. Cuando se examina certificados para validar, sin embargo, es necesario contactar a la Autoridad Certificadora que los emitió para verificar sus CRL -- esto no es una parte automatizada del proceso.

NOTA

Si utiliza una Autoridad certificadora que no esta configurada en su navegador por omisión es necesario cargar el certificado de la Autoridad Certificadora al navegador, permitiéndole validar al servidor certificado firmado por dicha autoridad. Esto último resulta ser muy peligroso, ya que una vez cargado, el navegador aceptará todo certificado firmado por esa Autoridad certificadora.

SSL

El SSL (Secure Sockets Layer) es una capa de protocolo la cual puede ser colocada entre una capa de red confiable orientada a la conexión (por ejemplo TCP/IP) y la capa de aplicación del protocolo (por ejemplo HTTP). SSL provee comunicación segura entre el cliente y el servidor al permitir mutua autenticación, el uso de firmas digitales para la integridad y el cifrado para la privacidad.

El protocolo esta diseñado para soportar un rango de opciones para especificar algoritmos usados para la criptografía, la codificación de los mensajes y las firmas. Esto permite la selección de algoritmos para permitir a servidores específicos estar basados en aspectos legales, de exportación u otros conceptos, y también permitir al protocolo tomar ventaja de nuevos algoritmos. Las selecciones son negociadas entre el cliente y el servidor al iniciar a establecer una sesión por el protocolo.

Hay un número de versiones del protocolo SSL, como se muestra en la Tabla 3.

Versión	Fuente	Descripción	Navegador que lo incorpora
SSL 2.0	Publicado por Netscape.	Protocolo Original.	Netscape 3.0 Internet Explorer3.0
SSL 3.0	Borrado de Internet expirado.	Revisiones para prevenir ataques específicos de seguridad.	Netscape 3.0 Internet Explorer 3.0
TLS 2.0	Borrador de IETF.	Revisión de SSL 3.0	Ninguno

Tabla 3 - Versiones del protocolo SSL

Como nota de la Tabla 3, uno de los beneficios en SSL 3.0 es que agrega soporte para de carga de cadenas de certificados. Esta característica permite a un servidor pasar un certificado de servidor junto con los certificados de quienes los expiden al navegador. La carga de cadenas también permite al navegador validar el certificado del servidor, si es que la Autoridad certificadora que certificó no este instalada para el emisor intermedio, ya que son incluidos en la cadena de certificados. SSL 3.0 es la base para el protocolo estándar TLS (Transaction Layer Security), actualmente en desarrollo por Internet Engineering Task Force (IETF).

Estableciendo una sesión

Una sesión SSL se establece al seguir una secuencia de reconocimiento entre el cliente y el servidor, como se muestra en la Figura 1. Esta secuencia puede variar, dependiendo si el servidor está configurado para *proveer un certificado de servidor* o *solicitar un certificado de cliente*. No obstante existen los casos, donde pasos adicionales de comunicación se requieren para administrar el cifrado de la información, este artículo resume un escenario común: vea la especificación de SSL para revisar el rango completo de posibilidades.

NOTA

Una vez que una sesión de SSL ha sido establecida, ésta puede ser rechazada, evitando así la baja en el desempeño al repetir los pasos necesarios para comenzar una sesión.

Los elementos de una secuencia de reconocimiento, como la utilizan el cliente y el servidor, se listan a continuación:

1. Negocian el esquema de cifrado para ser utilizado durante la transferencia de datos.
2. Establecen y comparten una llave para la sesión entre el cliente y el servidor.
3. Opcionalmente el servidor autentica al cliente.
4. Opcionalmente el cliente autentica al servidor.

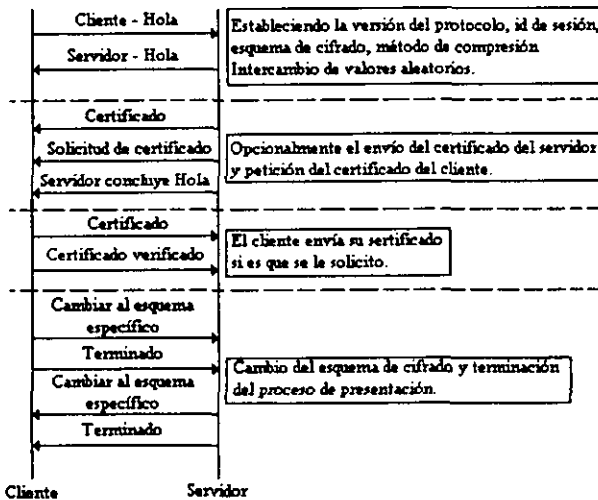


Figura 1

El primer paso, la negociación del esquema de cifrado, permite al cliente y al servidor escoger un esquema soportado por ambos. La especificación del protocolo SSL 3.0 define 31 esquemas. Un esquema está definido por los siguientes componentes:

- * Método para intercambiar la llave.
- * Cifrado de los datos transmitidos.
- * Codificación de los mensajes para crear el Código de Autenticación de Mensajes (MAC - Message Authentication Code).

Estos tres elementos se describen más adelante.

Método de intercambio de llave

El método de intercambio de llave define cómo compartir la llave oculta "criptográfica simétrica". usada por la aplicación para transferir, datos será convenida entre el cliente y el servidor. SSL 2.0 usa la llave de intercambio RSA, mientras que SSL 3.0 soporta una selección de los algoritmos para el intercambio de llave incluidos el RSA cuando los certificados son utilizados, y la de Diffie -Hellman para intercambiar llaves sin certificados y sin comunicación previa entre el cliente y el servidor.

Una variable en la selección de los métodos de intercambio de llaves son las firmas digitales. Firmar con una llave privada, provee seguridad contra un ataque de un tercero durante el intercambio de información usado en la generación de una llave compartida.

Cifrado para transferir datos

SSL usa el algoritmo convencional de la criptografía (criptografía simétrica) descrita inicialmente para cifrar mensajes en una sesión. Hay nueve selecciones posibles incluyendo la opción para el funcionamiento sin cifrado.

- * Sin cifrado.
- * Flujo cifrado.
 - o RC4 con llaves de 40 -bit.
 - o RC4 con llaves de 128-bit.
- * Cifrado de bloque CBC
 - o RC2 con llave de 40 -bit.
 - o DES40, DES, 3DES_EDE.
 - o Idea.
 - o Fortezza.

"CBC" se refiere a Cipher Block Chaining, lo cual significa que una porción de texto (previamente cifrado y codificado), se utiliza en el cifrado del bloque actual. "DES" se refiere al estándar DES (Data Encryption Standard), el cual tiene un número de variantes (incluyendo DES40 y 3DES_EDE). "Idea" es uno de los mejores y más robustos algoritmos criptográficos disponibles, y "RC2" es un algoritmo propiedad de RSA.

Función de codificación

La elección de una función de codificación determina cómo dicho proceso se crea a partir de una unidad de registro. SSL soporta a los siguientes:

- No-codificación (selección nula).
- MD5, un hash de 128 bits.
- Secure Hash Algorithm (SHA), un hash de 160 bits diseñado para usarse con el estándar Digital Signature Standard (DSS).

El mensaje codificado se usa para crear un Message Authentication Code (MAC) el cual es cifrado con el mensaje para proveer integridad, prevenir y repeler ataques.

Protocolo de secuencia de reconocimiento

La secuencia de reconocimiento utiliza tres protocolos:

- El "SSL Handshake Protocol" para permitir el establecimiento de una sesión entre un cliente y un servidor SSL.
- El "SSL Change Cipher Spec Protocol" para establecer realmente la forma conjunta del esquema de cifrado para la sesión.
- El "SSL Alert Protocol" para transmitir mensajes de error SSL entre el cliente y el servidor.

Estos protocolos, también como los datos de protocolo de aplicación, son encapsulados en el "SSL Record Protocol" (Protocolo de Registro SSL), como se muestra en la Figura 2. Un protocolo de encapsulamiento es transferido como datos por la capa más baja del protocolo, el cual no examina los datos. El protocolo encapsulado no tiene conocimiento del protocolo subyacente.

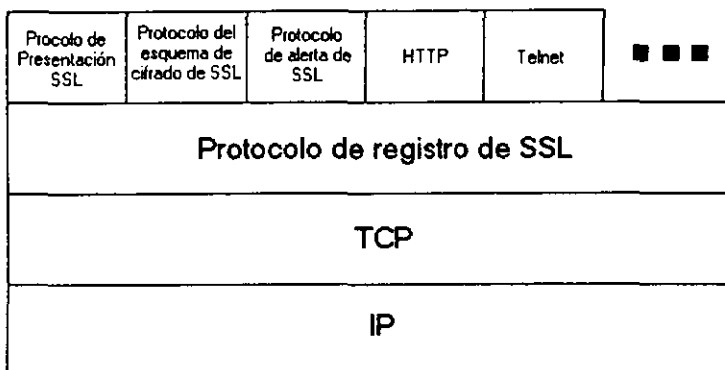


Figura 2

El encapsulamiento de los protocolos de control de SSL por el protocolo de registro significa que si una sesión activa es renegociada los protocolos de control serán transmitidos cuidadosamente. Si no hubiera sesión previa, entonces el esquema de cifrado NULL es usado, lo cual significa que no hay cifrado y los mensajes no tienen capacidades para verificar desde su integridad hasta que la sesión ha sido establecida.

Transferencia de datos

El protocolo SSL Record Protocol (Protocolo de registro de SSL), mostrado en la Figura 3, es utilizado para transferir aplicaciones y datos del SSL Control entre el cliente y el servidor, posiblemente fragmentando dicho dato en unidades más pequeñas, o combinando múltiples protocolos de alto nivel en unidades simples. Puede comprimir, añadir codificaciones firmadas y cifrar esas unidades antes de transmitir las usando protocolos básicos de transporte confiables.

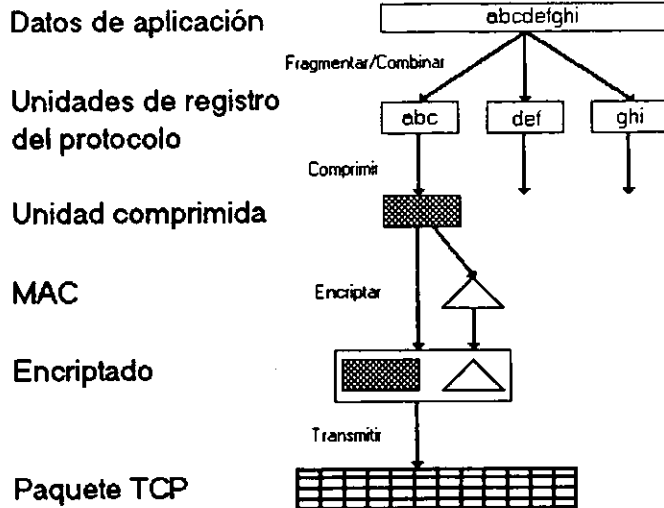


Figura 3

Asegurando la comunicación del Web HTTP

Una utilización común del SSL es asegurar la comunicación con el Web HTTP entre un navegador (cliente) y un servidor Web. Este caso no excluye el uso de HTTP no seguros. La versión segura utiliza URL's que vienen con https más que http, y un diferente puerto del servidor (por omisión 443). El navegador mantendrá las llaves privadas certificadas del cliente cuando son garantizados y activa un indicador si una conexión segura esta siendo utilizada.

Implementando SSL

Aunque uno puede escribir una implementación de SSL de la nada siguiendo la especificación, es mucho más fácil usar uno de los juegos de bibliotecas ya existentes de SSL. Además, debido a las patentes, es usualmente necesario contar con la licencia de alguna biblioteca criptográfica, al menos en los E.U. Los juegos de bibliotecas SSL incluyen rutinas de cifrado, de mensajes y para el manejo de certificados. Cada uno requiere el uso de la licencia de algún paquete de llave pública en los E.U., de Security Dynamics (Originalmente disponible de Public Key Partners [RFC1170] y entonces del RSA, la cual fue adquirida por Security Dynamics).

Hay dos prominentes paquetes de llave pública de RAS disponibles:

- **RSARef**
La implementación de RSA Reference, un juego no provisto de código fuente de RSA, puede ser utilizada para aplicaciones Freeware y no comerciales. Consensus Development Corp utilizó una licencia para uso comercial pero no tiene mucho haciéndolo.

- BSAFE3.0

Implementación comercial de RSAREf.

Estas dos implementaciones de llaves públicas incluyen un juego completo de algoritmos de llave pública (incluyendo cifrado RSA y llave de intercambio Diffie-Hellman), algoritmos de cifrado simétricos y funciones de codificación. Ellos pueden ser utilizados con otras herramientas SSL, los más prominentes de los cuales se incluyen los siguientes:

- SSLRef

Ejemplo de la implementación SSL 3.0 de Netscape Communications Corporation.

- SSLPlus

Código fuente comercial del paquete disponible de Consec Development Corp, una extensión de SSLRef3.0. Requiere BSAFE3.0 (de RSA) para utilizarse.

- SSLava

Paquete compatible de SSL3.0 escrito en Java de Phaos Technology.

- SSLeay

SSLeay-0.8.1es una implementación gratuita no comercial de SSL 2.0 y 3.0. Incluye una implementación de llave pública la cual puede ser utilizada fuera de los E.U. En los E.U., RSAREf o BSAFE3.0 deben ser utilizados pagando las cuotas de las patentes requeridas. SSLeay ofrece una forma barata para poner en funcionamiento SSL.

GLOSARIO

GLOSARIO

Acceso Directo (Archivo de Computadoras)	El uso de archivos de computadoras vía elementos portadores (como discos, cassettes, cartuchos) diseñados para ser insertados en una computadora o a su equipo auxiliar por el usuario. Ver Accesos Remoto (Archivos de Computadora).
Acceso Remoto (Archivos de Computadoras)	El empleo de archivos de computadoras vía dispositivos de entrada/salida conectados electrónicamente a una computadora. Ver Acceso Directo (Archivos de Computadoras).
Archivo de computadora	Un archivo (datos y/o programas) codificados para su manipulación por una computadora.
Archivo de Datos en Formato Máquina	Ver Archivo de Computadora.
Archivo Multiparte	Entidad bibliográfica que consiste de más de un archivo de computadora.
Archivo, Computadora	Ver Archivo de Computadora.
Area	Una sección mayor de una descripción bibliográfica, involucrando datos de una categoría en particular o de un conjunto de categorías.
Autor	Ver Autor Personal.
Autor conjunto	Persona que colaboró con una o más personas para producir un trabajo en relación a una función específica. Ver Responsabilidad Compartida.
Autor Personal	La persona principalmente responsable de la creación del contenido intelectual o artístico de un trabajo.
Campo 856 MARC	El campo 856 en registros bibliográficos permite conectar a recursos electrónicos. Esta característica proporciona el mismo control bibliográfico y acceso tanto para recursos electrónicos como para los materiales de la biblioteca tradicional.
Catálogo	<ol style="list-style-type: none"> 1. Una lista de materiales bibliográficos contenidos en una colección, una biblioteca, o un grupo de bibliotecas, organizados de acuerdo a un plan bien definido. 2. En un sentido más amplio, una lista de materiales preparados para un propósito particular (es decir, una exhibición, catálogo).
Compilador	<ol style="list-style-type: none"> 1. Aquel que produce una colección seleccionando y conjuntando ejemplares con relación a trabajos de varias personas o entidades. 2. Aquel que selecciona y conjunta en una publicación, ejemplares con relación a trabajos de varias personas o entidades. Ver también Editor.
Cuerpo relacionado	Elemento corporativo que tiene relación con otro elemento distinta a la relación jerárquica (es decir, que está fundada pero no controlada por el otro elemento; elemento que sólo recibe apoyo financiero y/o otro tipo de asistencia de otros elementos).

Designación del material específico	Un término indicando la clase especial de material (usualmente la clase de objeto físico) al que un ejemplar pertenece (es decir, un disco sonoro). Ver además Designación General de Material.
Designación General de Material	Término indicando la clase de material genérico de material al que un ejemplar pertenece (es decir, grabación sonora). Ver Designación Específica de Material.
Distribuidor	Un agente o agencia que tiene derechos exclusivos o compartidos sobre la venta de un ejemplar.
Edición (Archivos de Computadora)	Todas las copias que esencialmente tienen el mismo contenido y ejemplar de una misma entidad.
Editor	Aquel que prepara un ejemplar para su publicación externa. El trabajo editorial puede ser limitado por la preparación del ejemplar por el fabricante, o puede incluir supervisión del fabricante, revisión (restitución), o elucidación del contenido del ejemplar, y la adición de una introducción, notas y otros asuntos críticos. En algunos casos, puede involucrar la dirección técnica del personal encargado de crear o compilar el contenido del ejemplar. Ver Compilador.
Ejemplar	Documento o conjunto de documentos en cualquier forma física, publicada, entregada o tratada como una entidad, y que forma la base de una descripción bibliográfica individual.
Ejemplar Multimedia	Ejemplar conteniendo dos o más categorías de material, ninguno de los cuales se puede identificar como elemento constituyente predominante de un ejemplar; también designado como Kit.
Elemento	Una palabra, frase, o grupo de caracteres representando una unidad distintiva de información bibliográfica y formando parte de un área de la descripción
Elemento Corporativo	Una organización o grupo de personas que son identificadas por un nombre particular y que actúa como una entidad. Ejemplos típicos de elementos corporativos son las asociaciones, instituciones, firmas comerciales, empresas no lucrativas, gobiernos, agencias gubernamentales, instituciones religiosas, conferencias.
E-MAIL	Correo electrónico
Encabezado	Nombre, palabra o frase al inicio de un registro catalográfico para proporcionar un punto de acceso. Ver Punto de Acceso.
Encabezado principal	La primera parte del encabezado incluyendo un subencabezado.
Extensión del ejemplar	El primer elemento del área de la descripción física. Proporciona el número y la designación específica del material de las unidades del ejemplar siendo descrita y, en algunos casos, otras indicaciones de extensión (por ejemplo, duración). Ver Designación Específica de Material.
Formato	En un sentido más extenso, la presentación física particular de un ejemplar.
Fuente de información principal	La fuente de datos bibliográficos con preferencia como fuente de datos para la preparación de descripciones bibliográficas (o parte de ellas).

Información diferente al título	Título que surge de un ejemplar distinto al título propio o a los títulos paralelo o de la serie; además cualquier frase apareciendo en conjunción con el título propio, etc., indicativa del carácter, contenido, etc., del ejemplar, del motivo del mismo o de la ocasión de su producción o publicación. El término incluye subtítulo, pero no incluye variantes del título propio.
Juego	Un ejemplar o grupo de materiales diseñados para jugar reglas implícitas o prescritas de recreación o instrucción. Ver Kit.
Kit	Ejemplar conteniendo dos o más categorías de material, ninguno de los cuales es identificable como elemento constituyente predominante del ejemplar; también designado como "ejemplar multimedia". Ver Juego.
Nombre de archivo (Archivos de Computadoras)	Designación empleada en un sistema de cómputo para identificar un archivo. En algunas ocasiones, un archivo es llamado "nombre de un conjunto de datos". Para designación externa de un archivo de computadora, ver Título Propio.
Nombre de Conjunto de Datos	Ver nombre de archivos (Archivos de Computadoras).
Nombre predominante	Nombre o forma del nombre de una persona o corporación que aparece con más frecuencia (1) en el trabajo de una persona referenciado por la corporación, o (2) en fuentes de referencias en orden de predilección.
Palabra de Registro	Palabra a través de la cual un registro es ordenado en el catálogo, usualmente la primer palabra (distinta a un artículo) del encabezado.
Pantalla de Título (Archivo de computadoras)	En el caso de un archivo de computadoras, el despliegue de datos que incluye el título propio y usualmente, pero no necesariamente, el grado de responsabilidad y los datos relacionados con la publicación.
Portador físico	Medio físico de almacenamiento de datos, sonidos e imágenes. Para ciertas categorías de material, el portador físico consiste de un medio de almacenamiento como cartuchos, cintas, etc.
Punto de Acceso	Un nombre, término, código, etc., bajo el que un registro bibliográfico puede ser buscado y localizado.
Referencia	Una dirección de un encabezado o registro a otro.
Referencia cruzada	Ver Referencias.
Registro	Entrada de un ejemplar en un catálogo.
Registro Adicional	Un registro, adicional al registro principal, por medio del cual que un ejemplar es representado en un catálogo.
Registro Principal	Registro catalográfico completo de un ejemplar, presentado en forma tal que la entidad es identificada y citada en forma única. El registro principal puede incluir trazados.
Responsabilidad compartida	Colaboración entre dos o más personas o entidades desarrollando el mismo tipo de actividades en la creación del contenido de un ejemplar. La contribución de cada persona o entidad puede ser separada de la de otros. Ver además Autor Asociado, Responsabilidad Mezclada.

Responsabilidad mezclada	Un trabajo de responsabilidad mezclada es en el que diferentes personas o entidades contribuyen al contenido intelectual o artístico desarrollado en distintos tipos de actividades (adaptación o ilustración de un trabajo escrito por otras personas). Ver Autor Conjunto, Responsabilidad Compartida.
Serial	Publicación en cualquier medio publicado en partes sucesivas ostentando designaciones numéricas o cronológicas y se entiende que continuará de manera indefinida. Las publicaciones seriales incluyen publicaciones periódicas, periódicos, anuarios (reportes, libros del año, etc.), revistas, memorias, procedimientos, etc., de sociedades; y series monográficas numeradas. Ver además Series I.
Series	<ol style="list-style-type: none"> 1. Un grupo de ejemplares separados relacionados entre sí por el hecho de que cada elemento muestra, además de su propio título, un título colectivo aplicado al grupo en su totalidad. Los ejemplares individuales pueden o no estar numerados. 2. Cada dos o más volúmenes dentro de una secuencia de publicaciones seriales (es decir, Notas y solicitudes, series primera y segunda, etc.).
Sine loco (s.l.)	Sin lugar (es decir, el nombre del lugar de la publicación, distribución, etc., es desconocido).
Sine nomine (s.n.)	Sin nombre (es decir, el nombre de la editorial, distribuidor, etc., es desconocido).
Suplemento	Un ejemplar, usualmente elaborado separadamente, que complementa algún otro ya editado proporcionando una actualización o de otra manera, continuando con el original o conteniendo características especiales no incluidas en el original. El suplemento tiene una relación formal con el original como autoría en común, un título o subtítulo en común, y/o una intención estática de continuar o complementar el original.
Título	Una palabra, frase, carácter o grupo de caracteres, normalmente aparecidos en un ejemplar, que nombran el ejemplar o el trabajo contenido en él. Ver además Título Alternativo, Título Proporcionado, Título Propio, Título Uniforme.
Título Alternativo	La segunda parte de un título propio que consiste de dos partes, cada una de las cuales es un título por sí mismo; las partes son unidas por "or", o su equivalente en otro lenguaje (como La tempestad, o, La Isla Encantada).
Título Propio	El nombre principal de un ejemplar, incluyendo cualquier título alternativo pero excluyendo títulos paralelos y otra información del título.
Título Proporcionado	Un título asignado por el catalogador para un ejemplar que no tiene título propio en la fuente principal de información o en su subtítulo. Puede ser tomado de cualquier parte en el propio ejemplar de la fuente de referencia, o puede ser elaborado por el catalogador.

<i>Título Uniforme</i>	<ol style="list-style-type: none"> 1. El título particular por el que un trabajo puede ser identificado para propósitos de catalogación. 2. El título particular es empleado para distinguir el encabezado de un trabajo del encabezado de un trabajo distinto. 3. Título convencional colectivo empleado para colocar publicaciones de un autor, compositor, o elemento corporativo conteniendo diversos trabajos o extractos, etc. de diversos trabajos (es decir, trabajos completos y diversos en una forma literaria o musical en particular).
<i>URL</i>	Uniform Resource Locator
<i>URN</i>	Uniform Resource Name
<i>USMARC</i>	Los formatos USMARC son estándares en la representación y comunicación de información bibliográfica y relativa en formato de máquina. Es un conjunto de códigos y designaciones de contenido definidos para la identificación de tipos particulares de registros en formato de máquina. Los formatos USMARC están definidos por los siguientes tipos de datos: bibliográficos, de pertenencia y de autoridad.
<i>Vocabularios controlados</i>	Vocabulario de los lenguajes controlados que permiten menor esfuerzo por parte de los usuarios para recuperar la información deseada; se obtiene menor información que cuando se usa el lenguaje natural pero es información más precisa.

REFERENCIAS

REFERENCIAS

Software Engineering : A Practitioner's Approach

Roger S. Pressman
McGraw Hill College
4th edition (August 1996)

Introducing software engineering

Neville J. Ford and Mark Woodroffe
Prentice Hall
New York : 1994

Ingeniería de Software Explicada

Marck Norris y Peter Rigby
Megabyte
México 1994

Biblioteca del Congreso de los Estados Unidos de América

<http://www.loc.gov>

International Standar Serial Number

<http://www.issn.org>

National Serials Data Program (NSDP)

Library of Congress
Washington, D.C. 20540-4160
Tel.: (202) 707-6452
Fax: (202) 707-6333
Email: issn@loc.gov

Universal Library Labels

<http://www.ul.cs.cmu.edu>

ANSI/NISO Z39.56 standard (SICI)

http://www.lib.virginia.edu/speccol/scdc/articles/alcts_brief.html

Finding Aids for Archival Collections

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

SGML concepts

Bradley, Neal. 1992
ASLIB proceedings 44, no. 7/8
(July/August): 271-274.

SGML: an authors' guide to the standard generalized markup language

Bryan, Martin. 1988
Wokingham, Eng.: Addison- Wesley.

Del Catálogo a la ayuda de selección

Buckland, Michael. 1994.

Briefings from the CFCC, no. 2, 1994 in ALCTS Newsletter 5, 1994.

Bibliografía de SGML (Standard Generalized Markup Language) y temas afines

Cover, Robin. 1991.

Kingston, Ontario : Dept. of Computing and Information Science

Queen's University at Kingston.

SGML Web Page. 1994

<http://www.sil.org/sgml/sgml.html>

Colección de Imágenes Digitales: Modelo de Catalogación de Datos y Acceso por Red

Davis, Stephen Paul. 1995

<http://www.cc.columbia.edu/cu/libraries/inside/projects/diap/paper.html>

Hypertext Markup Language

<http://www.w3.org/hypertext/WWW/MarkUp/html-spec>

Panorama—SGML on the Web

<http://www.oclc.org:5047/oclc/research/panorama/>

The Berkeley Finding Aid Project, Standards in Navigation

<ftp://library.berkeley.edu/pub/sgml/findaid/arlpap.txt>

Guidelines for Electronic Text Encoding and Interchange

Sperberg-McQueen, C.M. and Burnard, Lou. 1994.

Chicago, Ill. and Oxford, England: Text

Encoding Initiative., 1994.

<http://etext.virginia.edu/TEI.html>

TEI Home Page, 1994

<http://www.uic.edu/orgs/tei/>

USMARC DTD

<ftp://library.berkeley.edu/pub/sgml/marcdtd/>

Practical SGML

Van Herwijnen, Eric. 1990.

Dordrecht: Kluwer Academic Publishers.

The World Wide Web Consortium

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

PURL (Persistent Uniform Resource Locator)

<http://purl.oclc.org/>

URN (Uniform Resource Name)

URL:<http://www.dlib.org/>

Sistema Handle

URL:<http://www.handle.net/>

Berkeley Digital Library SunSITE

URL:<http://sunsite.berkeley.edu/FindingAids>

EAD Website

URL:<http://lcweb2.loc.gov/ammem/ead>

The Application of Z39.50 in Your Library

Scott Nickerson

Information Technology column - June 1998

Dalhousie University, School of Library and Information Studies of Canada

<http://www.mgmt.dal.ca/slis/etig/>

<http://magi.com/~mmelick/it98jun.htm>

Electronic Document Management in the Digital Library

Sharon Vegh

June 1996 Information Technology

Dalhousie University, School of Library and Information Studies of Canada

<http://www.mgmt.dal.ca/slis/et>

<http://magi.com/~mmelick/it96june.htm>

Implementation of Z39.50 in Canadian Libraries

Carrol D. Lunau

Information Technology column - October 1997

Dalhousie University, School of Library and Information Studies of Canada

<http://www.mgmt.dal.ca/slis/et>

<http://magi.com/~mmelick/it97oct.htm>

Interoperability, Scaling, and the Digital Library Research

A Report on the May 1995

IITA Digital Libraries Workshop

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

Report of the Publishing Perspective Working Group

William Arms

IITA Digital Libraries Workshop

<http://www-diglib.stanford.edu/diglib/pub/reports/iita-dlw/part1.html>

Digital Library: Gross Structure and Requirements:

Report from a March 1994 Workshop

IEEE CAIA'94 Workshop on Intelligent Access to On-Line Digital Libraries

eLib, Electronic Libraries Programme

UK

<http://www.ukoln.ac.uk/services/elib>

Hacia la Biblioteca Electrónica

Heberto Reynel Iglesias

Realidad Virtual

Año 3, Número 4, 1993