

Colecciones digitales universitarias en México*



J. Alfredo Sánchez Huitrón**

Resumen

Se justifica brevemente el desarrollo de colecciones digitales construidas por y para comunidades de usuarios con características distintas a las que atienden predominantemente los progresos actuales. En particular, se motiva la evolución de colecciones en español, por autores (investigadores, educadores, escritores) mexicanos para un público global. Se revisan algunas estrategias para el desarrollo de colecciones digitales, los aspectos más importantes para explotaras e insertarlas en el contexto de las colecciones disponibles globalmente, así como algunos casos de colecciones disponibles o actualmente en proceso.

Palabras clave: desarrollo de colecciones, bibliotecas universitarias, bibliotecas digitales, México. (FRRE)

Abstract

The document provides a review of the strategies for construction and utilization by universities and other scholarly entities of digital collections in Mexico. Special emphasis is given to the continuity in the development by Mexican authors (researchers, professors and writers) of collections in Spanish language for the public in general. Several digital collections currently available in Mexico are also reviewed. (FRRE)

Keywords: collection development, university libraries, digital libraries, Mexico. (FRRE)



* Conferencia magistral presentada en la Conferencia Internacional sobre Bibliotecas Universitarias "La cooperación en el futuro digital", organizada por la Dirección General de Bibliotecas de la UNAM. Ciudad de México, D.F., 26 y 27 de septiembre de 2002. Mesa redonda, "Desarrollo de Colecciones Digitales Universitarias en México", celebrada el 27 de septiembre de 2002.

** Director de Bibliotecas, Universidad de las Américas, Sta. Catarina Mártir, 72820 Cholula, Puebla, México.
alfredo@mail.udlap.mx

I. Introducción

Hasta hace apenas un par de décadas, los avances en tecnología de información, específicamente en el área de bases de datos relacionales, parecían apuntar hacia la construcción de espacios de información bien organizados y con facilidades de acceso cada vez con mayor estructura y orden. Así, se hizo posible que grandes volúmenes de información pudieran subdividirse en pequeñas unidades o átomos, los cuales pueden representarse y actualizarse digitalmente para después manipularlos y extraerlos. Gigabytes (y hasta terabytes) de datos quedaron accesibles a los usuarios, organizada y controladamente a través de *software* para el manejo de bases de datos; aun los documentos menos estructurados tendrían cabida en el mundo relacional. El caos no existiría más.

¿La vuelta al caos?

Con la caída en los costos del almacenamiento en dispositivos digitales y la popularización desde hace menos de una década de la "telaraña mundial" de información, o *World Wide Web*, las facilidades para publicar y multiplicar documentos multimediales han



provocado una explosión de información que alcanza los *peta-*, *exa-*, *zetta-* y hasta *yotta-* bytes. No sólo se ha rebasado la capacidad para organizar y recuperar de los manejadores de bases de datos comerciales, sino que se ha revertido la tendencia al

"orden" que suponía introducir ese tipo de *software*. Nuevamente, encontrar información en el universo digital es un procedimiento incierto, poco estructurado y con resultados frecuentemente poco confiables. ¿El caos ha vuelto?

Colecciones digitales universitarias en México

El oráculo al alcance de todos

En realidad no, el caos no ha vuelto. Las afirmaciones sobre la dificultad para encontrar información en el ciberespacio son, en general, exageradas. Con mayor frecuencia recurrimos al *web* para resolver necesidades de información cada vez más sofisticadas:

¿alguien no sabe qué es un *yotta-byte*?, ¿o desea acceso a investigadores y publicaciones en el área de medicina genómica?,

¿se dice "el *web*" o "la *web*"?, ¿cómo suena el canto de una ballena? Los ejemplos pueden ser innumerables. Un buen "buscador" de *web* está en posición de ofrecer respuestas de muy buena calidad en cuestión de segundos. Internet 2 debe hacer esto aún más rápidamente. El *software* para manejo de bases de datos coexiste ahora con una amplia gama de mecanismos para almacenar y recuperar información, siempre en proceso, diseñados para enfrentar



los nuevos tipos de datos y sus características, como volumen, complejidad, estructura (o ausencia de ella) y dinamismo.

La necesidad de colecciones propias

Pero, si existen colecciones y medios de acceso tan completos y tan ampliamente accesibles, ¿no basta esperar un poco para que todo lo que necesitamos en materia de información se encuentre disponible en línea? ¿No deberían enfocarse los esfuerzos y recursos únicamente a dotar a las comunidades con infraestructura de cómputo y redes para garantizar el acceso a la información necesaria para el progreso?

Desafortunadamente, no es así. Las consideraciones sobre el potencial del medio digital para apoyar el aprendizaje, la investigación y la construcción social del conocimiento han tocado apenas la superficie. Sin embargo,

en este momento las afirmaciones sobre la facilidad de acceso a la información son aplicables a un sector muy pequeño de la comunidad global y todavía más pequeño en el caso de nuestro país. A pesar de que un buen número de universidades y otras instituciones en México cuentan con servicios de información y proveen acceso a grandes colecciones digitales disponibles comercialmente, éstas apoyan a un sector de usuarios relativamente pequeño con costos que consumen nuestros presupuestos. Pero aunque estos servicios de información no fueran tan costosos, la situación en general de las colecciones digitales en México en relación con su relevancia y alcance no mejoraría significativamente. Considérese, por ejemplo, el factor del idioma: según un estudio reciente,¹ la proporción del contenido disponible en *español* en Internet es la mayor después del inglés, pero alcanza apenas un 5.68% del total disponible. La traducción automática de contenidos digitales es un área de investigación pujante y con aplicaciones a futuro, pero en el presente es apenas una herramienta de apoyo a nivel léxico

¹ L5 : *The Fifth Study on Languages and the Internet : Networks and Development Foundation* (FUNREDES) [en línea]. FUNREDES, Union Latine, c 1996-2001. <<http://funredes.org/LC/english/L5/>> [Consulta: febrero 2002].

o para dominios de aplicación muy restringidos.

Otros factores esenciales son, indudablemente, la calidad y la veracidad de la información disponible. Los autores han sido liberados de un complicado procedimiento editorial que autorice publicar las opiniones y los resultados de sus proyectos de creación e investigación. Una consecuencia natural de esto es que la selección de materiales de calidad y el procedimiento editorial se transfieren a los lectores o usuarios del material publicado.

Puesto muy simplemente, si nos limitamos a asegurarnos de contar con infraestructura de acceso y a esperar los contenidos digitales desarrollados por los demás, corremos el riesgo de enfrentar a nuestros usuarios a colecciones vastas pero ininteligibles o inutilizables por la imposibilidad de determinar su validez, pertinencia y actualidad.

La evolución de colecciones digitales propias -construidas y actualizadas por nuestra comunidad de autores y usuarios para nuestro consumo y de los demás- es una necesidad que requiere atención pronta y concertada por actores

tales como investigadores, educadores, editoriales, escritores, instituciones de educación superior y entidades de fomento de los sectores público y privado. Estas colecciones deben integrarse de manera natural a las ya existentes, tanto comerciales como de dominio público. Los medios de distribución existen y están mejorando constantemente (Internet, Internet 2, discos compactos, tarjetas de memoria removibles); pero el contenido, el nuestro, debe producirse para ocupar espacios, atender demandas y generar presencia en la comunidad global.

¿Cómo hacerlo?

Es necesario conocer los medios por los cuales es posible generar colecciones digitales, explorarlos y evaluarlos para diseñar estrategias que permitan acelerar su construcción y disponibilidad. A la par, deben desarrollarse servicios para recuperar información e interfaces para individuos y grupos de usuarios que hagan posible el uso y extensión de estas colecciones como apoyo a actividades de aprendizaje e investigación. A los resultantes espacios virtuales que -al integrar colecciones, servicios y ambientes e interfaces de usuario- hacen

posible la investigación y aprendizaje coparticipativos, les denominamos *bibliotecas digitales*. Intentar separar colecciones de servicios e interfaces para efectos de progreso no es de mucha utilidad, al menos en la fase incipiente en que se encuentran las bibliotecas digitales en México. Debemos realizar esfuerzos concertados que impulsen integralmente el desarrollo de bibliotecas digitales.

En el resto de este documento se revisan los medios existentes para construir colecciones digitales, las oportunidades y retos para su explotación, así como algunos casos específicos de avances exitosos en colecciones y servicios digitales.

2. Construcción de colecciones digitales

Pueden distinguirse dos medios generales para construir colecciones digitales: integrar contenidos creados digitalmente y digitalizar materiales existentes en formatos analógicos. En seguida se discuten ambos enfoques. Por simplicidad, se utilizará el término *documento* para denotar unidades genéricas de información de diversa índole, que incluyen artículos o

Colecciones digitales universitarias en México

libros impresos o microfilmados, piezas de audio o video analógicos, además de archivos digitales de texto, audio, imágenes o video.

2.1 Integración de contenidos digitales

La ubicuidad del medio digital

La inmensa mayoría de los documentos producidos en la actualidad se encuentra en formato digital. Ya en el año 2000 se calculaba que el total de documentos "únicos" producidos anualmente en el mundo era de alrededor de dos exabytes (miles de millones de gigabytes, o 10^{18} bytes) en todos los medios existentes.² De estos documentos, apenas un 0.003% corresponde al medio impreso, mientras que el digital se ha adoptado universalmente para almacenar y diseminar información. Aun los documentos que se producen en papel son precedidos regularmente por al menos una versión digital.

Los problemas de integración

Con todas las bondades relacionadas con la consulta y

recuperación de información digital, la construcción de colecciones digitales parecería casi trivial si se aprovecharan todos los documentos producidos en este formato. Desafortunadamente, no es así. Los documentos pueden encontrarse en formato digital, pero es raro que los documentos relacionados se encuentren en un solo acervo, tengan una misma estructura, usen el mismo idioma o cuenten con una calidad homogénea, por mencionar algunas características deseables en una colección digital ideal.

Las oportunidades de integración

Sin embargo, el uso tan amplio del medio digital en la actualidad puede capitalizarse para orientar la producción de documentos e integrar colecciones digitales con características deseables de calidad y manejabilidad. Por ejemplo, un área de particular potencial en México, como se ha empezado a hacer en otros países, es la producción de tesis de licenciatura y posgrado. Miles de estos textos se producen regularmente en un medio digital en nuestras instituciones de educación superior

y centros de investigación. Una práctica arraigada en estas comunidades ha sido convertir el documento final a formato impreso, el cual se encuaderna y se almacena en los departamentos académicos o las bibliotecas tradicionales. En muchas ocasiones, ni siquiera se requiere entregar la versión digital de los escritos, de manera que el uso que pueden tener las contadas copias de las tesis encuadernadas es, en el mejor de los casos, sumamente limitado. Diseñar y adoptar políticas institucionales muy simples puede permitir integrar colecciones de tesis digitales con facilidades de acceso y consulta, así como con un gran potencial de difusión y uso como base de la colaboración entre comunidades de investigación.

Las producciones académicas y científicas representan un área especialmente susceptible de integrar en colecciones digitales útiles y trascendentes. Por un lado se encuentran los vastos volúmenes de datos recolectados en campo o experimentalmente. Dado que estos datos se generan digitalmente, su disponibilidad inmediata y su actualización

² LYMAN, Peter and VARIAN, Hal R. *How Much Information?* [en línea]. University of California, c2000. <<http://www.sims.berkeley.edu/research/projects/how-much-info/>> [Consulta: septiembre 2002].



dinámica agilizarían diseminar los resultados, aumentarían la conciencia de grupo en las comunidades de investigación y propiciarían la colaboración. Por otro lado se encuentran las publicaciones digitales formales. El número de revistas de investigación disponibles sólo en formato digital crece continuamente. Nuestras comunidades académicas y de investigación tienen la posibilidad y, de algún modo, la obligación de crear foros propios de publicación y difusión con ayuda de autores, editores y lectores locales y globales.

Otras áreas con buen potencial para integrar documentos digitales son los reportes internos de investigación, las publicaciones institucionales de diversos tipos, las memorias de congresos, los materiales de apoyo a cursos, las conferencias almacenadas digitalmente y los documentos internos de las organizaciones.

2.2 Digitalización de materiales en formatos analógicos

En relación con los milenarios medios analógicos para almacenar y difundir la información, el medio digital acaba de nacer. No obstante, su profundo impacto -actual o potencial- en el manejo de la

información y en nuestras prácticas para diseminar y generar el conocimiento es incuestionable.

Integración de materiales analógicos a colecciones digitales

Para aprovechar mejor el nuevo medio, nuestras colecciones idealmente deberían ofrecer acceso al conocimiento sin distinguir las formas como fue representado y transmitido originalmente. Lo mismo una comunidad puede buscar acceso a documentos muy recientes difundidos vía *web*, que otra a acervos generados hace siglos y almacenados en pergamino o papel, o a materiales reproducidos

originalmente en microfichas, cintas y discos de audio o video analógicos. Digitalizar materiales que se encuentran actualmente en esos medios e integrarlos a bibliotecas digitales en desarrollo tiene implicaciones y complicaciones diversas.

Avances en digitalización

Los aspectos técnicos involucrados en la digitalización de materiales impresos en buen estado y usando tipos uniformes pueden considerarse prácticamente resueltos: el equipo para digitalizar libros u otros documentos en papel o microformatos produce resultados de excelente calidad para textos e



Colecciones digitales universitarias en México

imágenes, los cuales pueden ser alimentados a sistemas de reconocimiento óptico de caracteres y utilizados en mecanismos sofisticados de búsqueda y recuperación de información.

Problemas en digitalización

Este no es el caso de materiales en mal estado, manuscritos o con tipografía poco uniforme. Transformarlos al nuevo medio puede requerir equipo especializado como digitalizadores cenitales, y la recuperación automatizada de información de estos materiales, que se encuentra aún en fase experimental, utiliza técnicas como visión computacional y redes neuronales. Similarmente, digitalizar y explotar medios analógicos temporales como audio y video representa todavía desafíos importantes que se están enfrentando con técnicas de reconocimiento automático de voz, razonamiento basado en casos, reconocimiento de rostros y estudio de escenarios espacio-temporales.

3. Explotación de colecciones digitales

Como se mencionó, el construir colecciones digitales debe ligarse

estrechamente con el proceso de mecanismos que permitan su difusión y empleo por comunidades amplias de usuarios. Sin duda, las colecciones accesibles a través de Internet (o mejor aún, de Internet 2) tienen un potencial de uso mucho mayor que las que se reproducen y se distribuyen por medios como los discos compactos. Sin embargo, no es suficiente colocar documentos digitales en sitios *web*, sino que es necesario desarrollar servicios e interfaces que permitan explotarlos y aprovecharlos para apoyar las tareas de los usuarios.

Estructura y semiestructura de colecciones

Puede afirmarse que la gama de servicios posibles así como de metodologías y técnicas aplicables para explotar colecciones digitales depende, en gran medida, del modo como éstas se organizan. Desde este punto de vista, pueden distinguirse dos grandes tipos de colecciones: las altamente estructuradas y las semiestructuradas. En el caso de las primeras, los contenidos digitales se organizan uniformemente, especificando de antemano los atributos de interés y sus tipos de datos. Las segundas,

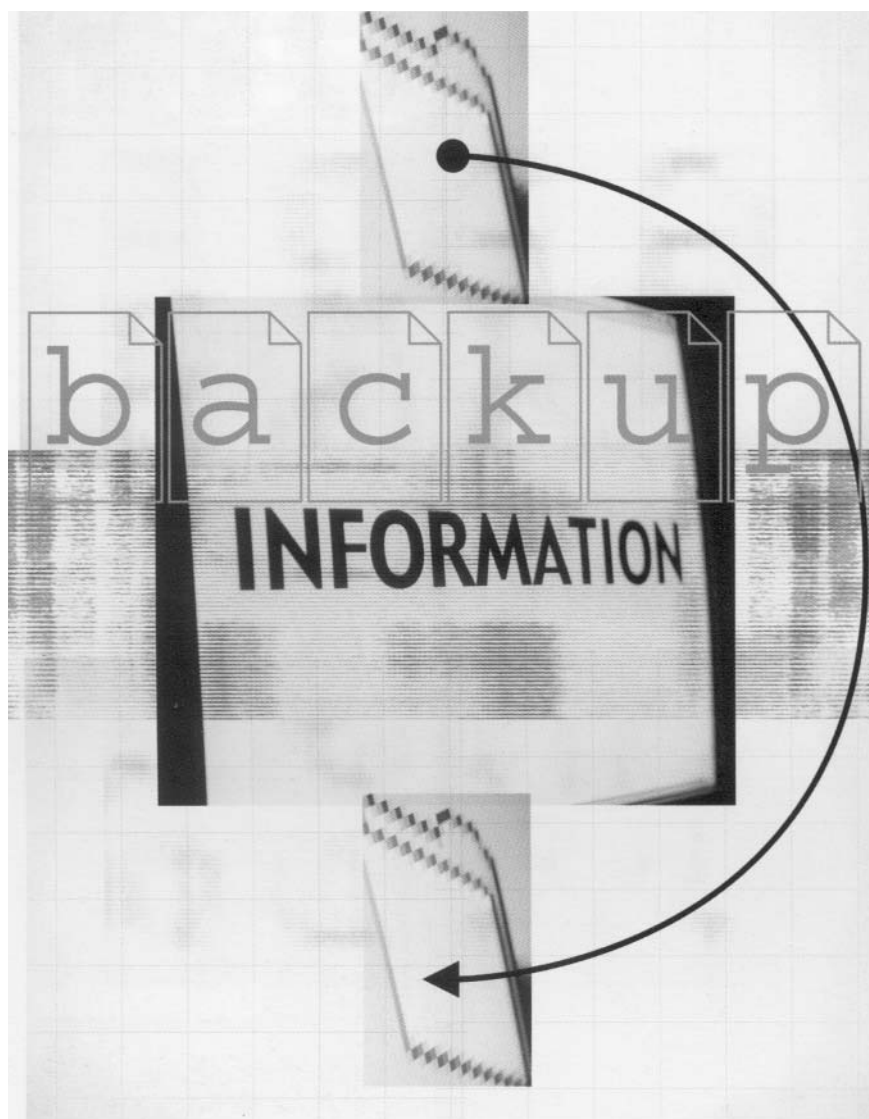
por su parte, comprenden contenidos de tipos y extensión variables y son descritos por algunos metadatos que auxilian en su explotación. La mayoría de los documentos disponibles vía *web* son semiestructurados (los metadatos que los describen, a veces mínimos, son normalmente su título, autor y algunas palabras clave). Entre los esfuerzos enfocados a describir documentos semiestructurados mejor y más útilmente, destaca el lenguaje XML (Extensible Markup Language).³

Recuperación de datos y de información

Los sistemas comerciales para manejo de bases de datos permiten construir y explotar las colecciones altamente estructuradas a través de lenguajes de consulta y aplicaciones accesibles vía *web*. Las colecciones semiestructuradas plantean retos significativos para la disciplina conocida como Recuperación de Información (RI). A diferencia de la recuperación de datos típica del *software* comercial de bases de datos, la RI no se enfoca a encontrar documentos que contengan exactamente los términos de interés para el usuario, sino los que pueden ser relevantes

³ Extensible Markup Language, XML [en línea]. World Wide Web Consortium, c1996-2002. <<http://www.w3.org/XML/>> [Consulta: 2002].





técnicas para manejo de datos comprimidos y diversos modelos de RI.

Las nuevas colecciones, que idealmente incluyan datos multimediales y multilingües, tenderán a ser semiestructuradas. Afortunadamente, la variedad de técnicas y métodos desarrollados en el área de RI pueden utilizarse para asegurar que esas nuevas colecciones puedan explotarse. Aunque la implementación de los diversos modelos de RI puede no resultar trivial, varios proyectos en curso apuntan a generar componentes de *software* disponibles para el dominio público y que pueden adaptarse a las necesidades de cada biblioteca digital.

Interoperabilidad e interfaces de usuario

Considerando que cada colección digital se desarrollará con una estructura propia, deben tomarse en cuenta las dificultades inherentes al acceso federado durante el diseño y desarrollo de nuevas colecciones digitales, buscando que sean interoperables. Idealmente, el acceso a múltiples

para una consulta dada, acompañados por una medida de su relevancia esperada.⁴

Los avances en el área de RI, motivados por el auge de la *web*,

son de enorme importancia. Gracias a ellos, los llamados *motores de búsqueda* de Internet nos permiten explorar eficientemente acervos inmensos y heterogéneos. Entre estos avances, destacan las

⁴ RIJSBERGEN, C.J. VAN. Information Retrieval [en línea]. 2ª edición, c1999. <<http://www.dcs.gla.ac.uk/~iain/keith>> [Consulta: 1999].

Colecciones digitales universitarias en México

colecciones digitales debería ser transparente para el usuario. En este sentido, un enfoque que está cobrando importancia es la adopción de un protocolo de intercambio como el que propone la Iniciativa de Archivos Abiertos.⁵

Por último, para el mejor aprovechamiento de las colecciones digitales, los usuarios deben contar con interfaces de acceso apropiadas y ambientes virtuales que les permitan desarrollar actividades de colaboración relacionadas con los contenidos. El desarrollo de aplicaciones e interfaces descansa sobre los servicios de acceso e intercambio descritos anteriormente.

Bibliotecarios digitales

En general, las habilidades y prácticas desarrolladas durante siglos para el manejo y recuperación de información en el medio impreso y otros formatos analógicos no son aplicables al medio digital. Al incorporar acervos utilizando el nuevo medio a las colecciones de nuestras bibliotecas es fundamental considerar las nuevas habilidades y los cambios culturales necesarios,

no solamente entre nuestros usuarios, sino también entre los bibliotecarios. Las nuevas generaciones de bibliotecarios digitales requieren formarse rompiendo paradigmas fuertemente arraigados y siguiendo programas académicos que difieren en gran medida de los utilizados en la formación de profesionales en biblioteconomía y bibliotecología tradicionales. La capacitación y actualización continuas de nuestros bibliotecarios son indispensables para garantizar que pueden apoyar a los usuarios y atender sus demandas de servicios de información.

4. Casos de colecciones digitales en México

En México, algunas instituciones han promovido la construcción de colecciones digitales propias y ampliamente accesibles. Aunque los servicios desarrollados alrededor de ellas con frecuencia se limitan a búsquedas convencionales, es posible explorar medios que permitan aumentar su difusión y su potencial como apoyo a actividades de aprendizaje, enseñanza e investigación.

Colecciones de la Universidad de Colima

Sin duda, uno de los pioneros en México en crear y difundir colecciones en formato digital, la Universidad de Colima por medio de su Centro Nacional de Edición de Discos Compactos,⁶ ha

producido más de 200 títulos en áreas tan diversas como arte, historia, economía, ciencias de la salud, bibliotecas y tecnología. El disco compacto hace posible distribuir las colecciones en comunidades que no cuentan aún con acceso a Internet.



⁵ *Open Archives Initiative* [en línea]. 1999. <<http://www.openarchives.org>> [Consulta: septiembre 2002].

⁶ *Centro Nacional de Edición de Discos Compactos de la Universidad de Colima* [en línea]. Universidad de Colima, c2002. <<http://www.ucol.mx>> [Consulta: septiembre 2002].



Colecciones del ILCE

A través del Instituto Latinoamericano de Comunicación Educativa (ILCE),⁷ puede accederse a colecciones que incluyen la versión digital de cientos de libros de áreas como ciencias sociales, literatura, arte, ciencia y tecnología. Los documentos, disponibles en HTML con algunos medios de búsqueda básica, han sido editados por entidades como el Fondo de Cultura Económica y el Consejo Nacional para la Cultura y las Artes.



que pueden utilizarse para construir y administrar algunas colecciones digitales- es un proyecto de la Red de Desarrollo e Investigación en Informática (REDII) del CONACYT, realizado por el Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM),⁸ que ha integrado algunas colecciones usando Phronesis, principalmente del área de computación. Otras organizaciones mexicanas, como el Centro de Investigación en Computación (CIC) del Instituto Politécnico Nacional (IPN) y el Instituto Nacional de Astrofísica, Óptica y Electrónica, han iniciado la construcción de colecciones digitales empleando también Phronesis. Sin embargo, el volumen de documentos disponible actualmente en estas colecciones es todavía relativamente pequeño.

Colecciones digitales en IPN y UNAM

La integración de colecciones de tesis digitales se ha iniciado en varias instituciones del país. Por ejemplo, el CIC del IPN publica en este formato las tesis de su

programa doctoral y de sus dos programas de maestría, que incluyen documentos producidos desde 1999. A través de su sitio web,⁹ el usuario puede navegar en la colección, examinar las sinopsis de las tesis y descargar los documentos completos comprimidos (un total de 40 a la fecha del presente escrito). Similarmente, la colección de tesis digitales de la Universidad Nacional Autónoma de México (UNAM) ofrece, por medio de una interfaz de búsqueda, algunos documentos (31 al consultar su sitio web) en formato HTML, correspondientes a tesis de diversas áreas y niveles de estudio.¹⁰ Otras colecciones disponibles mediante el sitio web de la UNAM incluyen la Biblioteca Jurídica Virtual (actualmente con más de 360 libros, varias publicaciones periódicas y más de 2,200 artículos), revistas y monografías del Instituto de Biología, así como un volumen considerable de libros y revistas digitalizadas por la Coordinación de Publicaciones Digitales.

Colecciones Phronesis

Phronesis -nombre con el que se identifica a un conjunto de herramientas de dominio público

⁷ ILCE : Instituto Latinoamericano de la Comunicación Educativa [en línea]. <<http://www.ilce.edu.mx>> [Consulta: septiembre 2002].
⁸ Proyecto Phronesis [en línea]. <<http://copernico.mty.itesm.mx/phronesis/proyecto>> [Consulta: septiembre 2002].
⁹ Unidad de Tecnologías Educativas, Instituto Politécnico Nacional, Centro de Investigación en Computación [en línea]. CIC-IPN, 2002. <<http://www.cic.ipn.mx/ute/>> [Consulta: septiembre 2002].
¹⁰ Coordinación de Bibliotecas : Cómputo Académico UNAM [en línea]. <<http://biblioteca.dgsca.unam.mx/acervo.html>> [Consulta: septiembre 2002].

Colecciones digitales universitarias en México

Colecciones digitales en la Universidad de las Américas, Puebla

Bajo una iniciativa denominada "Bibliotecas Digitales Universitarias para Todos" (U-DL-A, por sus siglas en inglés), la Universidad de las Américas, Puebla (UDLA), ha dedicado esfuerzos y recursos de investigación y desarrollo a la construcción de colecciones y servicios digitales que apoyen, por un lado, las actividades académicas de profesores y estudiantes universitarios y, por otro, que actúen como una plataforma para explorar cuestiones abiertas en el área de bibliotecas digitales.

Entre los servicios e interfaces desarrollados, se incluyen recomendaciones de materiales bibliográficos a usuarios, ambientes de referencia virtual para bibliotecarios y usuarios en general, además de herramientas para revisar y anotar documentos digitales. Los detalles de estos componentes de bibliotecas digitales se encuentran fuera del alcance del presente documento, pero pueden hallarse en

publicaciones del grupo de investigación y desarrollo.¹¹

Respecto a la construcción de colecciones digitales, énfasis de este trabajo, U-DL-A ha dedicado esfuerzos a los dos enfoques generales referidos en el apartado 2: integración de contenidos digitales y digitalización de materiales analógicos. En el primer rubro se ubica principalmente nuestro proyecto de tesis digitales. En el segundo, la digitalización de nuestros documentos de reserva semestral en la biblioteca y de secciones clave de nuestras colecciones especiales. Las características más importantes de estos desarrollos, disponibles a través de las páginas *web* de las bibliotecas de la UDLA,¹² se presentan a continuación.

Tesis digitales de la UDLA

Nuestra colección incluye tesis digitales producidas desde 1998. A partir de 1999, la entrega de tesis en formato digital es un requisito de graduación para los estudiantes de licenciatura y posgrado de ingeniería en sistemas

computacionales y ciencias de la computación. La entrega de tesis digital ha sido opcional para todas las demás disciplinas. A la fecha, se han integrado 133 tesis: 67 pertenecientes a ingeniería en sistemas computacionales, 36 al posgrado en computación y el resto a voluntarios de otras áreas. Se han registrado alrededor de 45,000 accesos desde noviembre de 1999, de los cuales cerca del 50% se han originado en otros países. Actualmente, la Junta Académica de la Universidad está considerando la implantación de la entrega de tesis digital como requisito de graduación para todos los programas vigentes. Esta aprobación implicaría el crecimiento de la colección en un promedio de 800 tesis anualmente.

Entre las características importantes del proyecto de tesis digitales, cabe destacar las siguientes:

- La colección está disponible vía *web* con facilidades para buscar y navegar por ella según diferentes criterios (título, autor, asesor, etc.).
- Pueden efectuarse búsquedas

¹¹ *Publicaciones del Laboratorio de Tecnologías Interactivas y Cooperativas* [en línea]. <<http://ict.udlap.mx/pubs>> [Consulta: septiembre 2002].

¹² *Bibliotecas Digitales de la UDLA* [en línea]. <<http://biblio.udlap.mx>> [Consulta: septiembre 2002].

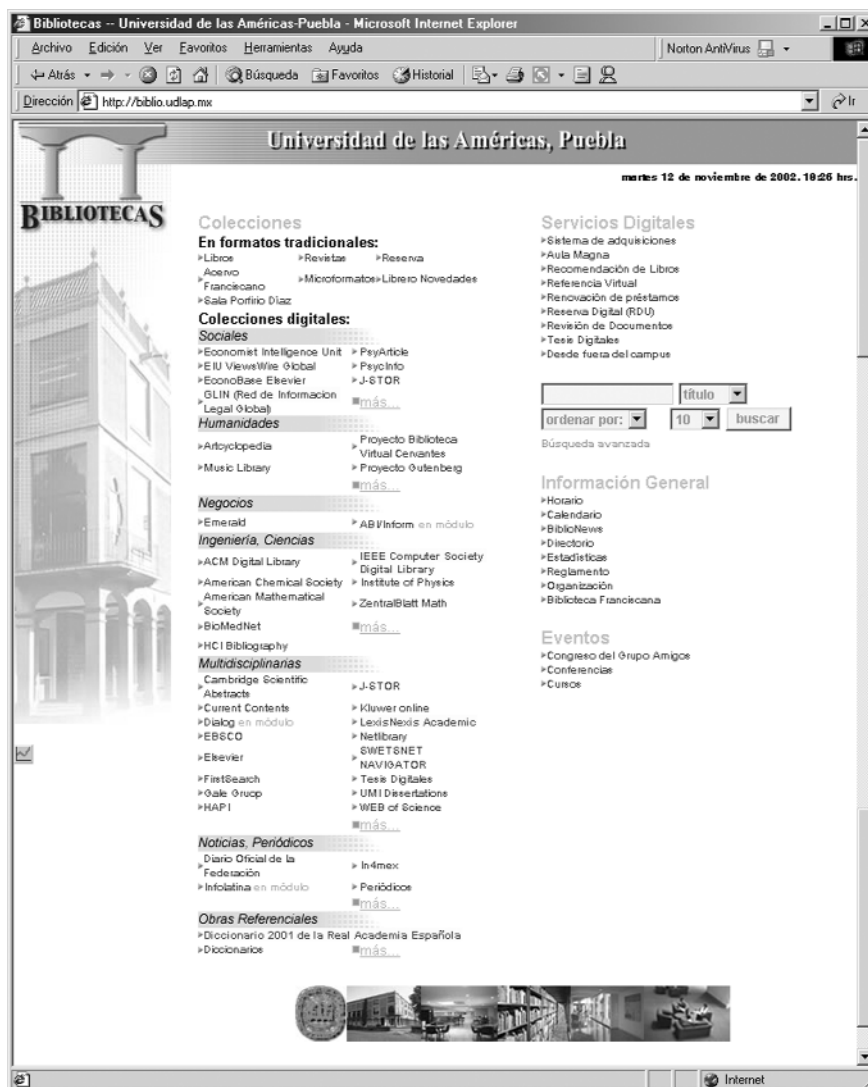


usando los metadatos que describen la colección, así como términos incluidos en cualquier componente de los documentos (texto, fórmulas, tablas, etc.).

• Los departamentos académicos, los estudiantes y la biblioteca se encargan de administrar la colección.

• Nuestro proyecto participa, con más de 125 universidades del mundo, en la iniciativa internacional *Networked Digital Library of Theses and Dissertations* (NDLTD).¹³ La colección es la única participante mexicana y una de cinco latinoamericanas (una chilena y tres brasileñas). NDLTD busca, entre otros objetivos, aumentar la visibilidad y disponibilidad de la investigación realizada a través de proyectos de tesis, reducir los costos de producción y manejo de este tipo de documentos, y promover las bondades del medio digital.

• La colección es la única en Latinoamérica registrada entre los más de cien proveedores de datos de la organización internacional denominada Iniciativa de Archivos Abiertos (OAI), que define estándares para facilitar el intercambio y la



diseminación eficientes de contenidos digitales. En términos prácticos, la participación en OAI ha implicado el desarrollo de un servidor que puede atender peticiones de datos utilizando el

protocolo de intercambio de dicha organización.

• La colección de tesis digitales ha sido la base para desarrollar un proyecto de recuperación federada de información

¹³ *Networked Digital Library of Theses and Dissertations* [en línea]. NDLTD, c2002. <<http://www.ndltd.org>> [Consulta: septiembre 2002].

Colecciones digitales universitarias en México

empleando el concepto de agentes móviles.¹⁴ Este proyecto, apoyado por el programa CONACYT-NSF, es una colaboración con el ITESM (*campus* Monterrey) y el Instituto Tecnológico de Virginia (Virginia Tech) en los Estados Unidos.

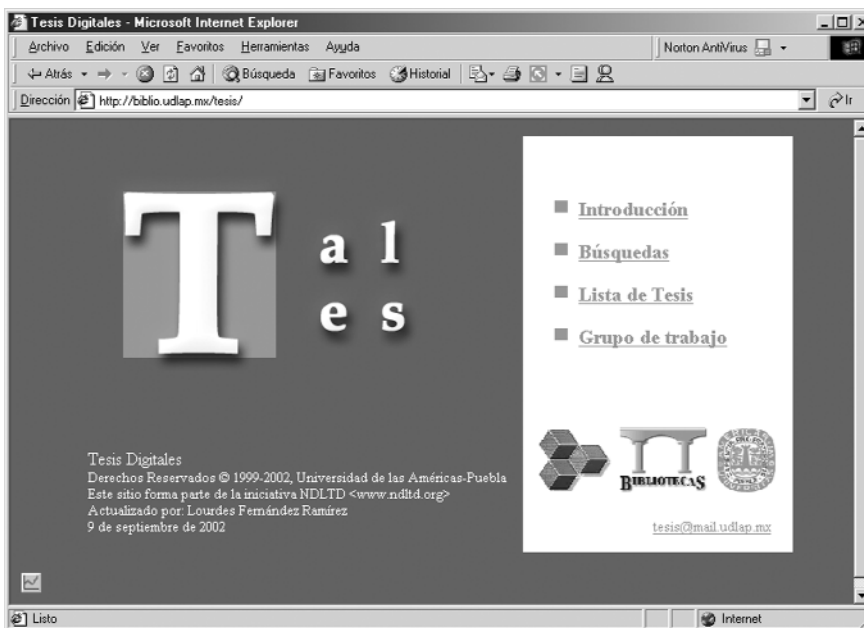
cual los profesores e investigadores pueden poner materiales de apoyo para cursos y proyectos a disposición de sus estudiantes y colegas. Este servicio sustituye al que tradicionalmente se ofrecía mediante el préstamo de materiales para fotocopiado. Con RDU, los

utilizan RDU y se han digitalizado más de 110,000 páginas.

Colecciones especiales de la UDLA

Las colecciones especiales constituyen el sello distintivo de una biblioteca. Están integradas por libros y documentos sumamente valorados en función de propiedades tales como su originalidad, antigüedad, rareza o contenido altamente especializado. Naturalmente, ésta es una área que puede beneficiarse al ponerse en formato digital al alcance de usuarios en todo el mundo. Entre las colecciones especiales de la UDLA que se han integrado a U-DL-A, se encuentran el archivo microfilmado de telegramas de Porfirio Díaz y el acervo franciscano.

Los telegramas generados durante el gobierno de Porfirio Díaz (1876-1911) integran una importante fuente de información para estudiantes e investigadores de diversas disciplinas. Como parte del proyecto de digitalización de estos materiales, se encuentran ya disponibles vía *web* los más de 4,000 telegramas (de un total de



Reserva Digital Universitaria (RDU)

Desde el semestre de otoño 2001, la Dirección de Bibliotecas de la UDLA ofrece el servicio de Reserva Digital Universitaria, por medio del

materiales recibidos se digitalizan y los usuarios pueden visualizarlos, almacenarlos o imprimirlos según sus necesidades. En el diseño de RDU, se ha puesto especial atención en proteger los derechos de autor. A la fecha, más de 600 cursos

¹⁴ Distributed information retrieval from web-accessible digital libraries using mobile agents. [en línea]. J. A. Sánchez [et al.]. *Upgrade: The Magazine of the Council of European Professional Informatics Societies*, 3, 2, <<http://www.upgrade-cepis.org>> [Consulta: septiembre 2002].



70,000) correspondientes al crítico año de 1910.

Por otro lado, la *Biblioteca Franciscana* -administrada desde el 2001 por la Dirección de Bibliotecas de la UDLA- comprende archivos históricos y un acervo de casi 25,000 volúmenes de los siglos XVI al XIX. Se ha comenzado la selección de materiales para digitalizarlos y se han desarrollado servicios que permitirán consultar directamente el contenido de las páginas digitalizadas.¹⁵ Finalmente, y también en la línea de explotación de colecciones digitales, se ha desarrollado un servidor y una biblioteca de modelos de recuperación de información, mediante los cuales se facilita el uso, desde diferentes aplicaciones, de una gama de herramientas para explorar colecciones digitales, incluyendo modelos ya clásicos como espacios vectoriales, *booleano* extendido e indexado semántico latente.¹⁶

5. Conclusiones


El formato digital representa un

desafío y una enorme oportunidad como medio para expresar, difundir y generar conocimiento. Aunque los avances en el área de bibliotecas digitales hacen posibles nuevas formas de trabajo para todas las organizaciones que generan, diseminan o usan conocimientos, se requiere un impulso decidido a proyectos nuevos y existentes dedicados a construir colecciones y servicios digitales propios.

El cambio cultural necesario para adoptar el medio digital se está dando aceleradamente entre los propios académicos, de modo que puede esperarse también una mayor demanda de mecanismos para construir y explotar colecciones digitales. Los casos de colecciones digitales universitarias aquí comentados no representan una lista exhaustiva. Desafortunadamente, no hay muchos más. La colaboración entre instituciones interesadas y entidades de fomento es vital para satisfacer significativamente las necesidades de información y para incrementar la presencia de nuestras comunidades de docencia e

investigación en el medio digital global.

Agradecimientos

El desarrollo de proyectos de bibliotecas digitales en la UDLA no sería posible sin la participación de los estudiantes talentosos y entusiastas del Laboratorio de Tecnologías Interactivas y Cooperativas, así como de investigadores del Centro de Investigación en Tecnologías de Información y Automatización, y de nuestro propio Departamento de Bibliotecas Digitales. Además del apoyo que hemos recibido de la UDLA, nuestros proyectos han sido financiados por el Consejo Nacional de Ciencia y Tecnología (proyectos No. 35804-A y G33009-A) y el Programa Iberoamericano de Ciencia y Tecnología para el Desarrollo (proyecto CYTED VII.19 RIBIDI). Agradecemos a todos ellos sus contribuciones. 



¹⁵ GARCÍA JIMÉNEZ, Patricia. *Consulta a textos digitalizados: implementación y análisis en el contexto de las colecciones especiales de la UDLA* [en línea]. Tesis de Ingeniería en Sistemas Computacionales, Universidad de las Américas, Puebla. <<http://biblio.udlap.mx/tesis>> [Consulta: agosto 2002].

¹⁶ MALDONADO, Fernanda. *Servidor y biblioteca de modelos de recuperación de información*. Tesis de Ingeniería en Sistemas Computacionales, Universidad de las Américas, Puebla. <<http://biblio.udlap.mx/tesis>> [Consulta: septiembre 2002].