

Elementos contextuales, técnicos y de diseño que posibilitan el acceso abierto a colecciones digitales universitarias.

ALMA BEATRIZ RIVERA AGUILERA,
JOSÉ JUAN TÉLLEZ BERTADILLO.
*Biblioteca Francisco Xavier Clavigero
Universidad Iberoamericana Ciudad de México*

En esta ponencia se expone la experiencia del Programa de Biblioteca Digital/Repositorio Institucional de la Biblioteca Francisco Xavier Clavigero (FXC), de la Universidad Iberoamericana, en lo relativo a ofrecer colecciones de acceso abierto. Este ejercicio se aborda desde la perspectiva del compromiso de una biblioteca universitaria contemporánea con el apoyo a la creación, la preservación y la difusión amplia de los conocimientos generados por la comunidad a la que sirve. En el texto se identifican algunos retos, que también pueden considerarse como oportunidades, urgentes para proporcionar acceso abierto en ambientes universitarios: los derechos de autor, la identificación y caracterización de los actores y la calidad de los contenidos. Se señalan también los elementos de diseño técnico que posibilitan la oferta de acceso abierto en nuestra biblioteca: el modelo conceptual, los perfiles de aplicación, la arquitectura de información y la arquitectura técnica. Asimismo, se mencionan tres casos prácticos: las tesis digitales, la colección de publicaciones en la web de los investigadores nacionales de la Universidad y las colecciones históricas. Se concluye señalando la importancia del acceso abierto en la difusión e impacto social de los nuevos conocimientos generados en las universidades, la oportunidad que el diseño y administración de

este tipo de colecciones representa para las bibliotecas universitarias y la necesidad de sentar bases sólidas en relación con los estándares y diseños que respondan a las preocupaciones contemporáneas en la organización de la información, tales como la Web 2.0 y la subsecuente Web Semántica.

1. Las bibliotecas universitarias y el acceso abierto.

En la actualidad, las bibliotecas digitales (BD) y los repositorios institucionales (RI) universitarios vienen a ser elementos necesarios para el florecimiento de las comunidades que generan y consumen conocimiento. Bustos-González, Fernández-Porcel, y Johnson (2007) señalan que procurar que los productos académicos estén disponibles a través del Internet es un estímulo económico y un vector de desarrollo social, que a su vez genera entornos pedagógicos ricos en información.

El acceso abierto responde a la vocación universitaria de creación y difusión del conocimiento y tiene como antecedente conceptual la definición ofrecida por los documentos de la Iniciativa de Acceso Abierto de Budapest (2002), la cual señala: "...el acceso abierto a la literatura que los académicos ofrecen al mundo sin expectativas de pago, significa que dicha producción está libremente disponible para el público en Internet y que permite a los usuarios leer, bajar, copiar, distribuir, imprimir, buscar o ligar textos completos".

Sube (2004/2010) y Zuccala (2009) identifican dos rutas para el acceso abierto: la verde, que implican los repositorios institucionales; y la dorada, representada por las revistas científicas de acceso abierto. Es claro el beneficio que ambas rutas ofrecen a la comunidad académica. Zuccala (Idem) señala que además, ambas son espacios para el desarrollo de la alfabetización científica en las personas comunes, es decir, no científicas.

El directorio de revistas en acceso abierto (DOAJ)¹ reporta en 2011 más de 2580 títulos y el directorio DOAR² más de 1850 repositorios. De los 14 repositorios incluidos en DOAR de México, 13 de ellos per-

1 <http://www.doaj.org/>

2 <http://www.opendoar.org/>

tenecen a universidades. De las revistas incluidas en DOAJ, al hacer una búsqueda con el término México se recuperan 40 títulos, de los cuales sólo uno es un falso hit (referido a New Mexico) y 29 son revistas de instituciones nacionales de educación superior. Desde luego, existen colecciones y revistas de acceso abierto que no están presentes en los directorios, pero consideramos que tanto DOAR como DOAJ dan información sobre tendencias generales sobre volumen y composición de los contenidos de acceso abierto en el mundo y en México.

La importancia de ofrecer ligas a espacios de acceso abierto a los usuarios no se discute en las bibliotecas universitarias, y el reto de aportar a la construcción de este tipo de espacios ha recaído en muchas de estas instituciones que actualmente acogen el reto como una oportunidad, tal como se verá en el apartado siguiente.

2. El programa biblioteca digital/repositorio institucional de la Biblioteca Francisco Xavier Clavigero.

Puede decirse que desde 1998 hasta ahora, el Programa BD/RI en nuestra biblioteca ha generado diversas colecciones y servicios locales que acompañan a las suscripciones comerciales de software y de contenido y a la identificación y organización de recursos de acceso libre para nuestros usuarios. Las colecciones locales han sido un espacio de múltiples retos tecnológicos y operativos, de reflexión sobre nuestras prácticas catalográficas y aprendizaje de nuevas, de exploración sobre políticas de derecho de autor, preservación y de creatividad a nivel de arquitectura de información.

El programa ha tenido avances tales como las tesis digitales a partir de 2004 con la explotación de textos completos locales vía Harvest,³ el desarrollo local del Gestor de imágenes Digitales GesID, el uso del software Greenstone desde 2008 y en 2010 la versión piloto del “recolector” de subida a Greenstone por parte de los usuarios de materiales digitales y sus correspondientes metadatos (Rivera Aguilera, 2009, Rivera Aguilera y Téllez Bertadillo, 2010; Rivera Aguilera,

3 <http://harvest.sourceforge.net/harvest/doc/index.html>

Téllez Bertadillo y O'Dogherty Madrazo, 2010, Téllez Bertadillo, 2010 y Rivera Aguilera, et. al., 2011). Asimismo, el Programa sirve de marco para generar productos cada vez más sofisticados de desarrollo tecnológico, metodologías de procesamiento y evaluación, actividades de investigación y sobre todo, para acercar a la Biblioteca a la comunidad universitaria y difundir a través de acceso abierto los productos de la comunidad universitaria como por ejemplo, información histórica, didáctica y frutos de investigación.

Este trabajo aborda las experiencias del programa arriba mencionado, en sus vertientes de retos y oportunidades y la respuesta de diseño técnico y de organización de la información de forma articulada requerida por las ciencias de la información y la computación.

3. Algunos retos y oportunidades para el acceso abierto en bibliotecas universitarias.

Entre los retos y oportunidades para una biblioteca universitaria al administrar una colección de acceso abierto, podemos mencionar los siguientes: preservación digital, capacitación, derechos de autor, la multiplicidad de actores involucrados, la calidad de los contenidos, entre otros. Para este trabajo se han considerado los tres últimos aspectos como los más relevantes, pero no los únicos, para la discusión y Falta completar idea.

3.1 Derechos de autor en el acceso al contenido abierto universitario.

El tema de los derechos de autor, sobre todo en el ambiente digital, ha generado en los últimos 20 años muchas dudas y prácticas de sentido común, atribuibles a la poca experiencia que en ese ámbito se tenía durante el siglo XX. Sin embargo, en el siglo XXI las comunidades de creación intelectual y las legislaciones nacionales e internacionales han ido aclarando cada vez más los aspectos de derechos de autor.

En el caso de México, para los derechos de autor el país se rige por la Ley Federal de Derechos de Autor. Vigente 24-12-1996 y reformas al 23 -07-2003 ⁴ y su correspondiente Reglamento Vigente 22-15-1998 y

⁴ <http://www.diputados.gob.mx/LeyesBiblio/pdf/122.pdf>

reformas al 14-09-2005.⁵ Desde luego, el país también se rige de acuerdo con las leyes internacionales de derechos de autor a las que se ha suscrito. Para el ámbito universitario mexicano, Callejas (2006) indica que los conflictos entre universidades y académicos sobre el tema de derechos de autor son mínimos y suelen ser arbitrados según procedimientos claros respecto a las responsabilidades sobre la originalidad por parte de los autores y del reconocimiento de la autoría por parte de los editores. Sin embargo, conviene recordar que los académicos suelen publicar, tanto en formato impreso como digital, no sólo en editoriales universitarias sino también en las de carácter comercial.

En el ámbito de la Web, ha surgido la Creative Commons⁶ (CC) como una corporación sin fines de lucro, dedicada a facilitar el compartir y construir nuevos contenidos digitales a partir del trabajo de otros, todo ello consistente con las leyes internacionales de derechos de autor. CC provee licencias gratuitas y otras herramientas legales que pueden asociarse a los contenidos, con características de uso tan abiertas o limitadas como el autor decida. El uso de las licencias CC se ha convertido en una práctica que resuelve y clarifica la forma legal y correcta del uso de los contenidos de acceso abierto.

Sin embargo, es todavía un reto asegurar la claridad en los registros de derechos de autor y los tipos de licencia que más se adecúen a las colecciones digitales. Por otra parte, cabe señalar que las bibliotecas universitarias han sido impulsoras de considerar el tema de derechos de autor y el uso de licencias CC en las colecciones que producen y/o administran, y han encontrado un nicho de “expertise” en las instituciones universitarias que usualmente recurren a sus bibliotecas como las especialistas en el tema.

3.2 Principales actores del contenido y del acceso abierto.

Jeffery y Asserson (2008) indican que maximizar el acceso a los productos de investigación mejora la riqueza, la calidad de vida, el ritmo de la investigación y la calidad de la investigación. Dado que gran par-

5 http://www.diputados.gob.mx/LeyesBiblio/regley/Reg_LFDA.pdf

6 <http://creativecommons.org/international/mx/>

te de la investigación generada es financiada por instituciones públicas, ésta debería de ser de acceso abierto y libre.

El uso que dan los diferentes tipos de usuarios a los contenidos es muy variado, señalan Jeffery y Asserson (Idem). Los investigadores necesitan tanto acceder a investigaciones y resultados preexistentes, como buscar posibles colaboradores externos. Los administradores de la información deben establecer las reglas y estrategias para comprobar la integridad de la información que se va a exponer sobre la institución, ya sea de índole interna como externa. El organismo que financia la investigación debe asegurar que se cumplan los objetivos de la investigación y principalmente, que los datos y resultados se publiquen en espacios arbitrados. Los grandes comercializadores de la información proceden a la búsqueda de productos de investigación en todos los continentes, países, instituciones y grupos de investigación. El innovador requiere acceder a encontrar nuevas ideas para explotarlas y mejorar la calidad de vida. El profesor o educador requiere el acceso a la información, con el fin de crear material didáctico; y el estudiante por su parte, con el fin de hacer uso y aprender de los materiales. Por último, los medios de comunicación necesitan recuperarla, para difundir o divulgar los resultados que obtengan sobre temas de interés al público.

La actividad bien llevada a cabo por cada uno de estos actores permite que la información sea mejor administrada, que los nuevos proyectos de investigación sean mejor planteados, y que el acceso y la rapidez a la información sean más eficientes. Como resultado, los estudiantes son beneficiados por el auto-aprendizaje, los profesores sacan provecho haciendo uso y reutilizando los resultados de investigación; y, la publicación abierta y la comparación de productos de investigación conlleva a la conformación de nuevos grupos de investigación.

Actualmente, las instituciones educativas presentan grandes oportunidades y retos en torno a la información digital. Thomas, et. al. (2005), mencionan que el crecimiento de la producción académica en formatos tales como: materiales didácticos, resultados de investigación y material académico en general, nace en formato digital. Por otra parte, las relaciones entre los usuarios, autores, editores y bibliotecas se ha vuelto tan compleja, que comienza a ser un espacio

de confrontación entre las diferentes visiones de acceso comercial y abierto. Los altos precios que muestran las estadísticas de la ARL (Association of Research Libraries) respecto a las suscripciones periódicas han dejado sin presupuesto a las bibliotecas (CINDOC-CSIC, 2004). Existe una diversidad de espacios para publicar, retornando estos a las bibliotecas en diferentes publicaciones, dando la impresión a las universidades de que se hace doble pago por la publicación. Una lógica aparentemente absurda es que el autor publica sus artículos sin esperar remuneración ni regalías. Lo hace mas bien, para conservar cierto prestigio dentro de la comunidad científica y para mantener o lograr su posición laboral en la institución (Melero, 2005).

En un afán por comprender a los actores de los contenidos de la información: institución educativa, autores, usuarios y biblioteca, Bustos-Gonzalez, Fernández-Porcel y Johnson (2007) proponen considerar las actividades asociadas a cada uno de los roles involucrados en el acceso abierto para la producción intelectual universitaria.

a) Las instituciones a través de sus Bibliotecas.

- Administrar y publicar la producción digital de la propia institución y de otras instituciones académicas
- Maximizar la visibilidad de la institución, investigadores y profesores por medio de publicaciones digitales de la producción científica y académica.
- Retroalimentar científica y académicamente a investigadores y profesores, lo cual puede hacerse entre la misma institución o desde alguna otra.
- Apoyar la educación en línea o a distancia.
- Administrar el conocimiento que se genera alrededor de los procesos de gestión de la educación.
- Preservar la producción digital.
- Compartir la producción digital con toda la comunidad y con la sociedad en general

b) Los autores: Investigadores, profesores y estudiantes.

- Encuentran nuevas formas de publicación.
- Tienen mayor acceso a publicar y consultar pre-publicaciones

- Expresan preocupación por la preservación de sus obras, (algunos encuentran poco confiable el medio electrónico).
- Tienen mayor miedo al plagio.
- Disponen de más recursos para añadir material complementario en sus artículos digitales.
- Pueden conocer inmediatamente por medio de aplicaciones Web, la situación de su artículo durante el proceso de producción.
- Se comienza a creer que Internet permite una difusión directa de sus obras como parte suya o de sus instituciones.

c) Los usuarios potenciales y a la sociedad en general.

- Aprecian la inmediatez de consulta y lectura, y el acceso directo al texto completo.
- Pueden realizar búsquedas retrospectivas y elaborar su propio archivo digital.
- Pueden consultar artículos de diferentes revistas desde su propia computadora, incluso accediendo a un único servidor, o a través de una misma interfaz de consulta, a través de bases de datos o RI.
- Los lectores desempeñan un papel más activo gracias al hipertexto, elementos multimedia, posibilidad de ejecutar programas y a las posibilidades de interactividad (mail, foros...)
- En algunos casos el usuario puede localizar artículos gratuitos que en versión impresa no lo son; también tienen opción al pago individual por artículo sin necesidad de suscribirse a toda la revista (pay per use).

3.3 Calidad de los contenidos de acceso abierto

Como se mencionó en el apartado 1, se establece una diferencia entre dos tipos de rutas de acceso abierto: las revistas arbitradas (oro) y los contenidos en repositorios no necesariamente arbitrados (verde). La diferencia entre ambas rutas es la calidad de la información. Las revistas de OA (oro) son revistas científicas revisadas estrictamente por pares o arbitradas, pero de acceso libre. Los repositorios no necesaria-

mente tienen información científica solamente; pueden incluir también preprints y postprints de artículos de revistas, tesis y disertaciones, materiales de cursos, bases de datos departamentales, archivos de datos, archivos de audio, video, registros institucionales, o las colecciones especiales de las bibliotecas digitalizadas. No son revisadas exclusivamente por pares, pero sí incluyen criterios de revisión por expertos y son de libre acceso (Sube, 2004/2010).

A los repositorios que permiten que el propio autor deposite su trabajo se les llama repositorios de auto-archivo. Por ejemplo, para SHERPA-Romeo, un recolector de ciencia abierta, se pueden revisar las reglas y condiciones para agregar contenido de acceso abierto en: <http://www.recolecta.net/romeo/index.php>. Cada comunidad tiene reglas y condiciones que son establecidas por la institución que alberga sus colecciones.

Otra clasificación de publicación científica o académica de acceso libre, incluye además del RI construido a través del concepto de autoarchivos y revistas de acceso abierto, la de depósito temático externo.

En resumen, los contenidos de la información de acceso abierto están validados según criterios que aseguran la calidad de los mismos. Podríamos señalar una diferencia general de acuerdo con los criterios de selección de los materiales por cada uno de los RI, como sigue: los contenidos de revistas de acceso abierto arbitradas (oro) tienen mayor calidad que los contenidos de los almacenados en RI, los cuales aceptan el autoarchivo de artículos u otros contenidos. Los criterios de aceptación del depósito temático definirán la calidad del mismo, más cercano a oro o a verde.

4. Posibilitando el acceso abierto en la Biblioteca Francisco Xavier Clavigero de la Universidad Iberoamericana.

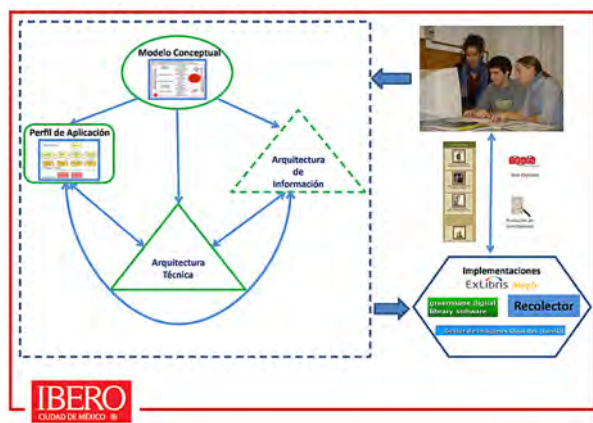
En la sección anterior, se revisaron las condiciones para la administración de colecciones de acceso abierto que impactan a las implementaciones concretas. En este apartado, se plantean dos grandes ámbitos que conjuntamente posibilitan la existencia de colecciones de acceso abierto en bibliotecas universitarias en general y en nuestra biblioteca en particular: los modelos y diseños (4.1). Es a partir de los modelos y

diseños, que se posibilita la creación y administración de colecciones concretas. A su vez, las experiencias concretas van afinando los modelos y diseños, de forma tal que cada fase de nuestro Programa de Biblioteca digital/repositorio institucional va evolucionando.

4.1 Elementos para el diseño y la implementación de colecciones de acceso abierto.

El modelo general de nuestras colecciones, cuya conceptualización inicial data de 1998, ha sido dinámico y en su versión actual incluye una concepción general de los contenidos y servicios pertinentes y el desarrollo de perfiles de aplicación, necesarios para definir los metadatos de colecciones específicas (Rivera Aguilera, et. al 2011) (Ver figuras 2 y 3). Desde el ámbito del cómputo se ha desarrollado una arquitectura técnica de repositorio que puede visualizarse en la *figura 1* (Téllez Bertadillo, 2010) y una arquitectura de información (ver figura 4) a partir de menús y ligas que ofrezcan al usuario el acceso a los contenidos con el menor número posible de clicks. Actualmente, se están conceptualizando modelos, diseños y pruebas piloto acorde a la Web 2.0 y la subsecuente Web Semántica.

Figura 1. Elementos de diseño e implementación de la Biblioteca Digital/Repositorio Institucional. Biblioteca Francisco Xavier Clavigero Universidad Iberoamericana Ciudad de México. (Elaboración propia).

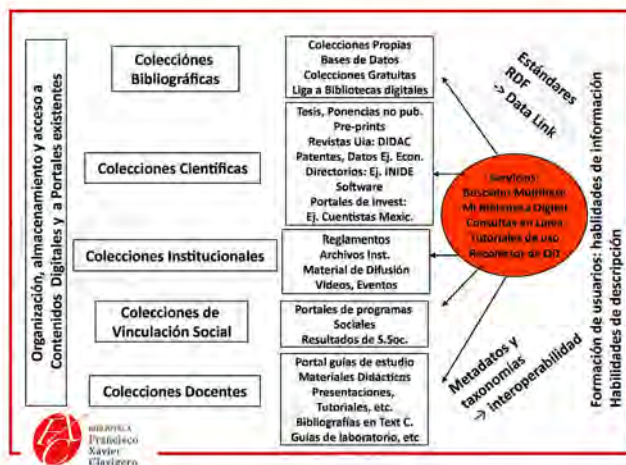


4.1.1 Diseño conceptual de la Biblioteca Digital/ Repositorio Institucional

El programa Biblioteca Digital/Repositorio Institucional ha pasado por diversas etapas en nuestra institución: Etapa I. 1998-2003 Concepto, primeras colecciones y servicios digitales. Etapa II. 2004-2007. Búsqueda en texto completo y definición de políticas y procedimientos de operación. Etapa III. 2008-2011. Colecciones dinámicas e integración del concepto de repositorio institucional al programa, uso de estándares de metadatos específicos a las colecciones y estrategias de la Web 2.0.

A partir del próximo año, se prevé una cuarta etapa (2012-2015), que integrará la consolidación de las fases anteriores a nivel de diseño y proceso, sin perder de vista los retos de la Web 3.0 (también llamada semántica). Este objetivo marca el modelo conceptual, el modelo de datos y la arquitectura técnica y de información de las aplicaciones. El concepto eje en todas las etapas es la implementación de colecciones y servicios que atiendan con calidad las necesidades de los usuarios. En la *figura 2*, puede verse el concepto que ha definido las acciones de la etapa III y que sienta las bases para la implementación de la etapa IV.

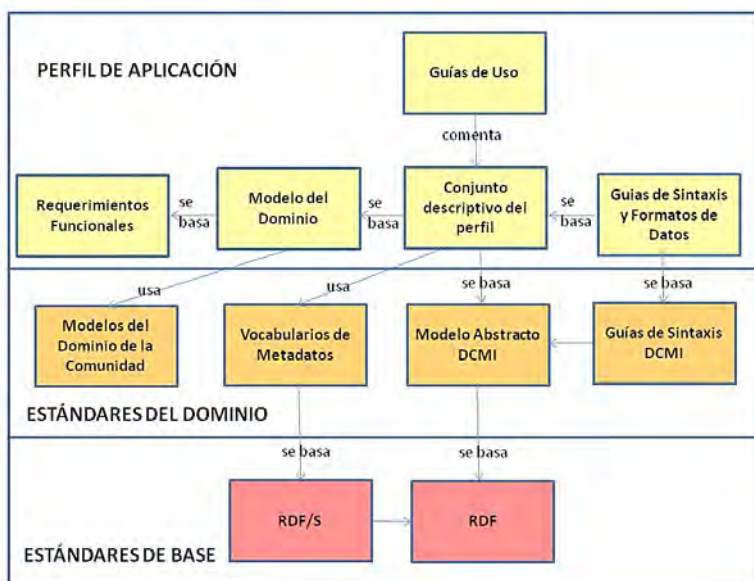
Figura 2. Modelo conceptual. Biblioteca Digital/Repositorio Institucional. Biblioteca Francisco Xavier Clavigero Universidad Iberoamericana Ciudad de México. (Elaboración Alma Rivera).



4.1.2 Perfiles de aplicación: un método para diseñar e implementar esquemas de metadatos descriptivos para colecciones digitales.

El papel de los metadatos en una biblioteca digital o un repositorio de información es clave para concretar el acceso a los contenidos que ahí se albergan. Sin perder de vista que existen muchos estándares de metadatos, Dublin Core es un referente fundamental para las bibliotecas universitarias comprometidas con el acceso abierto. Es por ello, que para definir los perfiles de aplicación que describen nuestras colecciones digitales, se han tomado como referente el Marco de Singapur y algunas experiencias locales

Figura 3. Marco de Singapur para perfiles de aplicación Dublin Core (Nilsson, Baker y Johnston, 2008). Traducción Alma Rivera.



En 2008 se publicó como documento oficial de Dublin Core el llamado Marco de Singapur o Singapore Framework (Nilsson, Baker y Johnston), el cual propone una estructura para el diseño de aplicaciones de metadatos para colecciones específicas fundamentada en los estándares del dominio a describir y como estándar genérico el modelo de datos Resource Description Framework y la semántica asociada al mismo. Dicha propuesta establece como obligatorias para un perfil de aplicación la documentación correspondiente a los requerimientos funcionales de la colección, el modelo del dominio y el conjunto de metadatos descriptivos del perfil (Ver *Figura 3*). Las guías de uso y de sintaxis y formato de datos fueron consideradas opcionales. Este mismo marco sirvió de base para el desarrollo del documento llamado *Guías para perfiles de aplicación Dublin Core* (Coyle y Baker, 2009). En ese mismo año, Rivera Aguilera (2009) presentó un trabajo sobre perfiles de aplicación, en el que indicaba la importancia de analizar la colección a enriquecer con metadatos definidos por un perfil de aplicación, la revisión detallada de los estándares pertinentes y la importancia de proponer nuevas etiquetas cuando se considerara necesario.

Actualmente, se establecen los requerimientos funcionales y los conjuntos descriptivos metadatos que describan mejor los contenidos y la colección, junto con los metadatos administrativos correspondientes. Se tiene claridad en los modelos de dominio y las guías de sintaxis y formatos de datos. (Rivera Aguilera, 2010 y Rivera Aguilera, et. al. 2010). Se espera, a partir de 2011, integrar a nuestros diseños toda la documentación requerida por el Marco de Singapur y las taxonomías locales requeridas.

4.1.3 Arquitectura de información

Morville y Rosenfeld (2006), indican que la arquitectura de información se basa en encontrar el equilibrio entre: las necesidades y el comportamiento del usuario, el contexto y por supuesto, el contenido. El resultado de articular estos conceptos es el producto que la institución expone o vende a sus usuarios agregando el valor que espera el usuario.

La arquitectura de la información de una colección de acceso abierto responde al modelo conceptual y a los perfiles de aplicación

Acceso Abierto a la Información en las Bibliotecas...

correspondientes a dicho modelo. De ahí, que el producto resultante será un sistema Web complejo que armonice el diseño, la funcionalidad, los cómo y las formas de recuperar la información para satisfacer las necesidades primarias del usuario.

Los autores antes mencionados identifican cuatro tipos de necesidades de la información: 1. Exploración. 2. Investigación exhaustiva, 3. Búsqueda de un tema y 4. Rehacer una búsqueda anterior. En el mundo real, puede considerarse que existe la combinación de todas ellas. Por otro lado, tenemos el comportamiento de los usuarios en relación con sus habilidades informáticas y sus expectativas de: tiempos de respuesta mínimos, localización de sugerencias que le ayuden a comprender mejor lo que busca, posibilidad de replantear su búsqueda dependiendo de resultados anteriores. En algunos casos, los usuarios comienzan haciendo una búsqueda de acuerdo a su necesidad primaria y terminan consumiendo información que nada tiene que ver con la búsqueda inicial.

Figura 4. Portal Biblioteca Digital Biblioteca Francisco Xavier Clavigero, diseño 2009.



Nuestro modelo incipiente de arquitectura de información está plasmado en los interfaces web a colecciones y servicios digitales. El diseño y navegación actual están siendo revisados según criterios de usabilidad acordes a los usuarios de la nueva década, en los cuales deben quedar bien identificados los cuatro componentes de la arquitectura de la información: sistema de organización, sistema de etiquetado, sistema de navegación y sistema de búsqueda. (Morville y Rosenfeld, 2006).

4.1.4 Arquitectura técnica: un acercamiento a través de MOCIRI (Modelo Conceptual de un Repositorio Institucional).

La arquitectura técnica de la información es la responsable de captar todos los requerimientos definidos en la arquitectura de la información y en el perfil de aplicación.

La implementación de la arquitectura técnica tiene como resultado un sistema Web que responda a la arquitectura de la información. La entrada al sistema son los contenidos y la información administrativa relacionada, los cuales se procesan y organizan agregando valor de acuerdo a las necesidades del usuario, para finalmente ponerla a disposición del usuario final y de la sociedad en general.

En la Biblioteca de la Universidad Iberoamericana Ciudad de México, se ha implementado parcialmente la arquitectura técnica basada en el modelo MOCIRI (Modelo Conceptual de un Repositorio Institucional) para el sistema Web GesID. Dicho modelo también tiene elementos que coinciden con la aplicación Greenstone. (Ver figura 5 y apartado 4.2).

El MOCIRI es un modelo técnico que sienta las bases para el desarrollo de una arquitectura técnica que considere el modelo conceptual, la arquitectura de la información y los perfiles de aplicación que en nuestro programa de Biblioteca Digital/Repositorio Institucional han ido desarrollándose y madurando de forma paralela y articulada. En este apartado, se describe MOCIRI como un antecedente de lo que será la arquitectura técnica de la fase IV de nuestro programa. El objetivo de este diseño técnico es articular una serie de herramientas

que permitan: coleccionar, administrar, organizar, procesar, preservar, compartir y publicar la producción de la institución, logrando satisfacer con esto las necesidades de la institución, de los autores y por supuesto, del usuario final de la información, así como la interacción entre estos. Para lograr este objetivo, se han definido seis capas, las cuales se describen brevemente a continuación:

Capa 1. Autores - Contenido

Identifica la información que ya exista o se esté generando en los departamentos de la institución. Después de una revisión rápida de los contenidos junto con los autores o responsables de la información sobre reglas y condiciones de publicación abierta, se toma la decisión sobre la viabilidad de publicación o no. Esta capa está articulada con el modelo conceptual y la definición de requerimientos funcionales del perfil de aplicación.

Los autores de los contenidos generalmente son los profesores, investigadores, estudiantes y administrativos. Cabe mencionar que cualquier autor que desee subir información al sistema, lo puede hacer libremente en tiempo real bajo las reglas y condiciones establecidas.

Capa 2. Colecciones digitales

En esta capa se definen la o las colecciones necesarias, con el fin de agrupar los contenidos o materiales según su temática, autoría, disciplina o alguna otra. A través de los metadatos descriptivos, debe quedar estructuralmente bien definida la colección. Estos metadatos deben quedar representados en un sistema de captura. La semántica de una colección está definida de acuerdo con las taxonomías, ontologías y axiomas; de esta manera, quedan los materiales completamente agrupados, identificados y relacionados. Este nivel corresponde integralmente al perfil de aplicación.

Capa 3. Recolección de contenido

Una vez definida la colección junto con el conjunto de metadatos establecidos en la capa 2, es necesario generar el ambiente o la estructura en la cual los autores puedan subir libremente sus contenidos o bien, en la que estos puedan ser transferidos desde otros sistemas.

Consiste básicamente en definir un interface sencillo con una plantilla de captura de metadatos, en la que los autores puedan sentirse con libertad para subir sus contenidos. Para lograr esto, es necesario ocultar los aspectos engorrosos de los metadatos y generarles una máscara simple y clara, para lograr una buena comunicación entre el sistema y los creadores. En nuestra experiencia, los metadatos junto con su respectiva máscara, valores por default y elementos de personalización deben de estar almacenados en una base de datos relacional. Esta capa debe responder a los estándares de almacenamiento y a su vez, a la arquitectura de información en su aspecto de captura directa por los creadores.

Capa 4. Preservación

La estructura definida en esta capa se basa en la conformación de archivos planos en formato XML, donde queden asentados todos los metadatos que están almacenados en la base de datos definida en la capa anterior; también debe de contemplar todos los elementos necesarios para poder compartir las colecciones con otros repositorios. El GesID tiene una herramienta para importar-exportar colecciones completas a archivos planos en formato XML; por su parte, Greenstone posee almacenamiento de metadatos en XML. Con respecto a los objetos digitales, esta capa implica tanto almacenamientos en formatos vectoriales independientes del administrador de la colección y los editores originales, como la asignación de metadatos propios de la preservación.

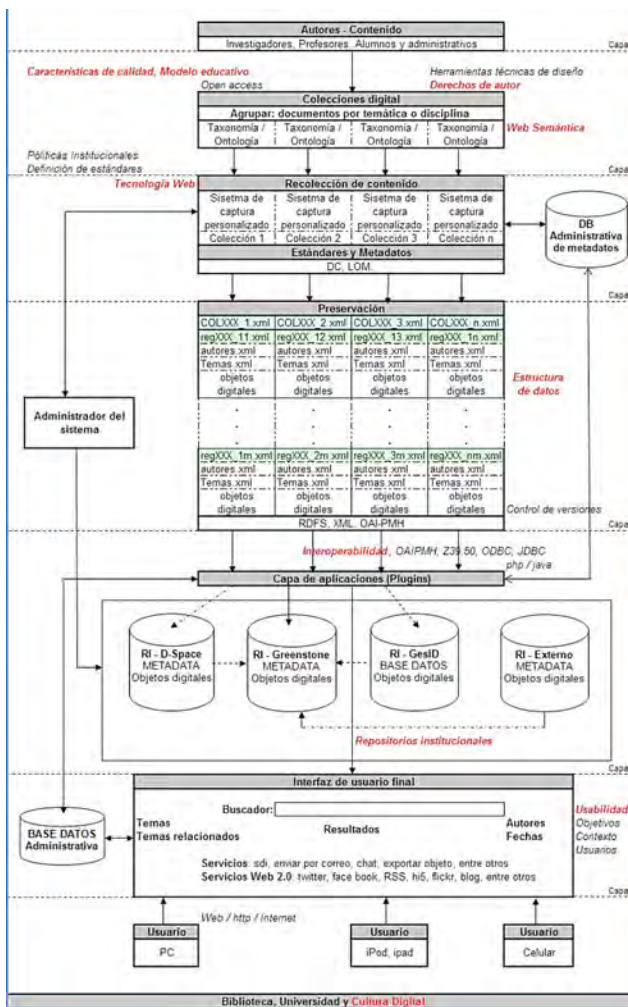
Capa 5: Aplicaciones (plugins)

Esta capa describe los programas (llamados plugins) que sirven para importar contenidos en diferentes formatos e importar y exportar registros de otros RI, como pueden ser Greenstone, DSpace, GesID, e inclusive entre bases de datos relacionales. Esta capa proporciona la fortaleza de indexar, por ejemplo, colecciones en texto completo; es decir, un RI indexa los metadatos y los contenidos de los objetos digitales, teniendo así un potencial amplio en cuanto a la recuperación de información, que abarca los metadatos capturados, los metadatos incrustados en los objetos y los textos completos.

Capa 6: Interfaz de usuario final

Es la capa en la que las necesidades y comportamientos de los usuarios, los autores y la institución deben de quedar representados a través de servicios como: el sistema de búsqueda, el despliegue de resultados, la promoción de otros materiales y colecciones, promoción

Figura 5. MOCIRI: Modelo Conceptual de un Repositorio Institucional. (Elaboración José Juan Téllez).



de otros autores, o cualquier otra información a través de una relación definida con anterioridad, permite la interacción entre autores, institución y usuarios y finalmente, la vinculación con las redes sociales. Esta capa asegura los interfaces con los diferentes dispositivos con los que cuenta el usuario final autor/consumidor de información. Esta capa responde integralmente a la arquitectura de información.

Los cuatro elementos ya señalados (modelo conceptual, perfiles de aplicación, arquitectura técnica y de información) han sido parcialmente implementados en algunas colecciones concretas cuyas características responden a las necesidades y características de nuestros usuarios y se exponen en el apartado siguiente.

4.2 Ejemplos de colecciones de acceso abierto en la Biblioteca Francisco Xavier Clavigero.

Las colecciones digitales de acceso abierto implementadas durante la fase III Biblioteca Digital/Repositorio Institucional (Ver figura 1) son las tesis digitales, publicaciones libres en la Web de los investigadores nacionales de la institución y algunas colecciones de acervos históricos.

4.2.1 Tesis digitales: Aleph, Harvest y Greenstone.

En el año 2004, se inauguró la primera colección digital en texto completo que daba forma sistemática e integral (metadatos y texto completo) y explotaba la información digital. Esta colección fue la de tesis digitales y estuvo basada en estándares y en el uso de una combinación de software comercial (Aleph)⁷ para las búsquedas vía metadatos y software libre (Harvest)⁸ para aquéllas de texto completo. La colección de tesis digitales de la UIACM se conformó originalmente en un sitio en HTML, que ligaba en una estructura jerárquica organizada por departamentos académicos a los archivos pdf de las tesis, almacenadas en un servidor Linux, las cuales podían recuperarse también a través del catálogo en línea Aleph 500 vía la etiqueta 856 del formato

7 <http://www.exlibrisgroup.com/category/Aleph>

8 <http://harvest.sourceforge.net/harvest/doc/index.html>

MARC. Una vez disponible el servicio, la aplicación fue evaluada y encontrada satisfactoria por los usuarios (Rivera Aguilera, Ramírez Moya y Pozo Marrero, 2005).

En 2008 se probó el administrador de colecciones digitales Greenstone (Witten, 2009), el cual es libre y de forma integral almacena y explota los metadatos y el texto completo de los objetos digitales textuales, administra las imágenes digitales con eficiencia y genera estructuras jerárquicas de acceso vía un html dinámico. Este modelo implica para las tesis digitales una transferencia desde MARC en Aleph a Dublin Core en Greenstone.

En el caso de las tesis digitales, diversos elementos vaciados en el registro MARC no tenían cabida en DC y eran necesarios para mantener el acceso por diversos puntos a la colección de tesis. Para poderlos incluir, se integraron campos del esquema de metadatos DC, Electronic Thesis and Dissertation Metadata Standard (ETD-MS)⁹ y fue necesaria también una ampliación local del mismo. (Rivera Aguilera, 2009).

En esta colección, el tema de la cesión de derechos de publicación digital se resolvió a partir de una carta que firmaron voluntariamente los posgraduados. Los tesistas entregan su material en disco compacto; un 85% de ellos firman el permiso de publicación digital de acceso abierto en la red y un 15% prefiere que su material se consulte en un CD, directamente en las instalaciones de la Biblioteca.

En lo referente a la calidad de las tesis, ésta es avalada por un comité departamental; y el hecho de abrir las tesis al acceso abierto ha hecho que tanto los tesistas como los comités, pongan especial cuidado en los detalles de forma y fondo, tal como lo han expresado los autores en el proceso de entrega.

4.2.2 Productos de los Investigadores Nacionales de la Universidad Iberoamericana Ciudad de México. Greenstone.

En el año 2008, se inició también una colección que describe materiales digitales de autores SNI de la UIA. Esta colección está instalada en Greenstone, utiliza metadatos de Dublin Core y emplea algunas eti-

9 <http://www.ndltd.org/standards/metadata/etd-ms-v1.00-rev2.html>

quetas locales para asociar el material a departamentos académicos y fundamentalmente describe los materiales (artículos, libros, ponencias, etc.) disponibles por acceso abierto en la red.

Esta modalidad de contenidos referenciados permite describir y mantener juntos los productos disponibles de nuestros investigadores y dado que describe ligas de materiales publicados por otras instituciones, no representa limitaciones de derechos de autor. Inicialmente, se consideró solicitar directamente a los investigadores el material, pero sólo 10% del total se mostró dispuesto participar en la creación de la base en esa modalidad. Por ello, se decidió hacer directamente las búsquedas del material de los investigadores disponible en la Web desde la Biblioteca. No se descarta en el futuro alimentar esta colección con información no publicada como son las presentaciones, reportes de investigación, datos de investigación, etc., con el permiso de los creadores.

4.2.3 Colecciones históricas: GesID

“Como parte de sus tareas fundamentales, la Biblioteca Francisco Xavier Clavigero de la Universidad Iberoamericana, Ciudad de México se ha dado a la tarea de digitalizar algunas de sus colecciones históricas. Producto de este esfuerzo es el sitio electrónico en el que se muestra parte de la riqueza de los acervos con los que cuenta la Biblioteca. Las colecciones que lo integran, a saber, la Colección Digital de la Independencia de México, la Colección de Postales de la Decena Trágica y el Archivo Fotográfico Alberto Salinas Carranza constituyen una importante fuente para la investigación sobre la historia de nuestro país y nos invitan a conocer más estos acervos. A través de la versión digital de los materiales que integran las colecciones, la Biblioteca logra dos de sus tareas fundamentales. La primera es la conservación de las fuentes primarias que constituyen documentos únicos e irremplazables; y la segunda, igual en importancia, es la difusión de piezas a las que difícilmente tendrían acceso muchos investigadores“. (Harari Betancourt, V., 2009).

El software utilizado para la administración de estas colecciones es un de carácter local y se denomina GesID (Gestor de Imágenes Digitales), utiliza las etiquetas de metadatos Dublin Core y está desarrollado con software libre usando como base de datos PostgreSQL,¹⁰ la cual es a su vez explotada con el lenguaje de programación interpretado PHP.¹¹ Este sistema incluye la administración de vocabularios controlados; y se espera desarrollar a través de dicho administrador taxonomías y eventualmente, ontologías. (Téllez Bertadillo, 2010).

El tema de los derechos de autor no representa dificultades en este tipo de colecciones históricas, dado que los materiales pertenecen a los acervos físicos propios y los derechos patrimoniales de los autores han prescrito, debido a la antigüedad de los mismos.

4.2.4 Colecciones y servicios a nivel de pilotaje: materiales didácticos en formato digital y recolector de objetos

Existe una pequeña colección de videos con mensajes del rector José Morales Orozco que puede recuperarse a través del catálogo en línea en Aleph 500, vía la etiqueta 856 del formato MARC.

Otra colección de acceso abierto que se pondrá en operación para consulta y captura directa por los autores es la de material didáctico en formato digital. (Rivera Aguilera, 2009). Dicha colección ha servido para el diseño y pruebas del aspecto de los autores universitarios, lo cual ha enriquecido la perspectiva del mero nombre del creador en relación con sus datos complementarios, utilizando etiquetas Vcard y algunas locales. (Rivera Aguilera, et.al., 2010).

Actualmente, se encuentra en fase de pruebas un recolector de objetos digitales que facilitará la subida y etiquetado descriptivo hecho directamente por los autores. Dicha aplicación también está desarrollada en un ambiente de software libre MySQL¹² y PHP.

10 <http://www.postgresql.org/>

11 <http://www.php.net/>

12 <http://www.mysql.com/>

Conclusiones

Desde la perspectiva del contenido, el acceso abierto es fundamental en la difusión e impacto social de los nuevos conocimientos generados en las universidades.

Entre los diferentes retos que enfrentan las bibliotecas respecto de las colecciones de acceso abierto en instituciones de educación superior, se discuten en este trabajo tres aspectos: la complejidad social de las comunidades universitarias y sus actores, los derechos de autor y la calidad en los contenidos. Dichos aspectos son básicos, puesto que representan para las bibliotecas tanto un reto como una oportunidad de posicionarse en sus instituciones con los servicios y el “expertise” requeridos.

Asimismo, se ha identificado la necesidad de sentar bases sólidas para las aplicaciones, en relación con los estándares y diseños que respondan a las preocupaciones contemporáneas en la organización de la información, tales como la Web 2.0 y la subsecuente Web Semántica. El modelo conceptual y los perfiles de aplicación que requiere cada colección, así como la arquitectura técnica y la arquitectura de información, son los documentos fundamentales para asegurar que las implementaciones respondan a los requerimientos y entornos de los usuarios actuales. Dichos usuarios integran simultáneamente el rol de consumidores y productores de información de acceso abierto en las universidades.

Finalmente, pero no menos importante, señalaremos que las implementaciones de acceso abierto están soportadas coherentemente en un software libre y presentan un fuerte potencial para llevar a cabo investigaciones interdisciplinarias desde la biblioteca universitaria.

Bibliografía

Budapest Open Access Initiative. (2002). *Budapest Open Access Initiative*. Recuperada el 15 de enero de 2011 de: <http://www.soros.org/openaccess/read.shtml>.

Bustos-González, A., Fernández-Porcel, A. y Johnson, Ian. (2007). *Directrices para la creación de repositorios institucionales en universidades y organizaciones de educación superior*. Babel Library. Alfa Network Babel Library. Recuperado el 15 de enero de 2011 de http://eprints.rclis.org/bitstream/10760/13512/1/Directrices_RI_Espa_ol.pdf

CINDOC-CSIC. (2004). *Revistas científicas y electrónicas: estado del arte*. Recuperado el 31 de enero de 2011 de: http://www.tecnociencia.es/e-revistas/especiales/revistas/pdf/e-revistas_informe.pdf

Callejas, C. B (2006). Los derechos de autor. *Revista Ciencia de la Academia Mexicana de Ciencias*. 57(2). Recuperado el 10 de agosto de 2010 de http://revistaciencia.amc.edu.mx/images/revista/57_2/derechosdeautor.pdf

Harari Betancourt, V. (2009). Colecciones documentales mexicanas: presentación. Portal Biblioteca Francisco Xavier Clavigero, Universidad Iberoamericana Ciudad de México. Recuperado el 30 de enero de 2011 de <http://www.bib.uia.mx/centenarios/galeria/indexGaleria.php>

Jeffery, K. y Asserson, A. (2008). Institutional repositories and current research information systems. *New Review of Information Networking*, 14:71-83.

Melero, R. (2005). Acceso abierto a las publicaciones científicas: definición, recursos, copyright e impacto. *El profesional de la información*. 14(4):255-266. Recuperado el 30 de enero de 2011 de: <http://elprofesionaldelainformacion.metapress.com/index/FYAA5WFK2KEFT6AL.pdf>

- Morville, P. y Rosenfeld, L. (2006). *Information Architecture for the World Wide Web*. 3a. Ed. Sebastopol, CA: O'Reilly Media. (Versión electrónica)
- Nilsson, M., Baker, T. y Johnston, P. (2008). *The Singapore framework for Dublin Core application profiles*. Recuperado 15 de agosto de 2010 de <http://dublincore.org/documents/2008/01/14/singapore-framework/>
- Rivera Aguilera, A. B. (2010) Perfiles de aplicación en Greenshine: integración de diversos esquemas de metadatos para la catalogación de las colecciones digitales de la Biblioteca Francisco Xavier Clavigero. Martínez Arellano, F. F. Comp. *4º Encuentro de Catalogación y Metadatos UNAM/Centro de Investigaciones Bibliotecológicas*. Ciudad de México. 215-232 pp. Disponible en: http://132.248.242.3/~publica/archivos/libros/iv_encuentro_catalogacion.pdf
- Rivera Aguilera, A. B., Ramírez Moya, A. M. y Pozo Marrero, A. (2005). *Evaluación de las colecciones digitales*, Foro Internacional sobre Tecnologías de Información Interfases 2005. Universidad de Colima, Colima, México
- Rivera Aguilera, A. B. et. al (2011) Conjunto descriptivo de un perfil de aplicación en lo referente a creadores de material didáctico digital en ambientes de biblioteca 2.0. *5º Encuentro de Catalogación y Metadatos UNAM/Centro de Investigaciones Bibliotecológicas*. Ciudad de México, 29/09/10 -01/10/10. En Prensa.
- Rivera Aguilera, A. B. y Téllez Bertadillo, J. J. (2010) La biblioteca digital universitaria: centro de recursos para el aprendizaje, la investigación y la vinculación social. *Mesa Bibliotecas Digitales. Reunión de Primavera 2010*. Corporación Universitaria para el Desarrollo de Internet (CUDI)/Red Abierta de Bibliotecas Digitales (RABiD). Morelia, 22/04/10. Videoconferencia.

Rivera Aguilera, A. B., O'Dogherty, R. y Téllez Bertadillo, J.J. (2010). "Visión del patrimonio cultural digital desde la Biblioteca Francisco Xavier Clavigero". *Foro Políticas de Digitalización del Patrimonio Cultural*. UIA 21/10/2010.

Rivera Aguilera, A. B. et al. Conjunto descriptivo de un perfil de aplicación en lo referente a creadores de material didáctico digital en ambientes de biblioteca 2.0. *5º Encuentro de Catalogación y Metadatos UNAM/Centro de Investigaciones Bibliotecológicas*. Ciudad de México, 29/09/10 -01/10/10

Téllez Bertadillo, J. J. (2010). Arquitectura de un repositorio: GesID un primer acercamiento. *II Jornada de Investigación en la Biblioteca Francisco Xavier Clavigero: Biblioteca Universitaria e Investigación*. Universidad Iberoamericana Ciudad de México.

Thomas, C., et. al. (2005). The new frontier of institutional repositories: a common destination with different paths. *New Review of Information Networking*, 11(1): 65-82.

Sube, P. (2004/2010). *Open Access Overview*. Recuperado el 30 de enero de 2011 de: <http://www.earlham.edu/~peters/fos/overview.htm>.

Witten, I. H. (2009). The development and usage of the Greenstone digital library software. *Bulletin of the American Society for Information Science and Technology*. 35 (2): 31-38. Disponible en: http://www.asis.org/Bulletin/Dec-08/DecJan09_Witten.html. Consultado el 9 de septiembre de 2009.

Zuccala, A. (2009). The lay person and open Access En B. Cronin Ed. *Annual Review of Information Science and Technology*. v. 43. (359-396). Medford, New Jersey: Information Today/American Society for Information Science and Technology.