



Temas de Biblioteconomía

La biblioteca digital en el ámbito universitario

Integración e interconexión de RE El movimiento OA y los repositorios de investigación

Autor: César Martín Gavilán

Fecha: 20/04/09

La biblioteca digital en el ámbito universitario

Definición de Biblioteca Digital

Si bien existe cierto consenso en la literatura especializada sobre el concepto y definición de la biblioteca tradicional, no sucede lo mismo cuando hablamos de la Biblioteca Digital. El concepto se ha usado con distintas acepciones y matices, siempre relacionadas con el acceso a documentos en formato electrónico por medio de redes de comunicación y asociados al modelo de “biblioteca del futuro”. Para mayor confusión, en el pasado se han utilizado varios términos (biblioteca electrónica, biblioteca virtual, etc.) para referirse a realidades semejantes a lo que ahora se identifica como Biblioteca Digital.

Dejando a un lado el debate puramente terminológico, entre las múltiples definiciones se puede destacar la de la Digital Library Federation (DLF) que define en 1998 las Bibliotecas Digitales como “organizaciones que proporcionan recursos, incluyendo personal especializado, para la selección, acceso intelectual, interpretación, distribución y preservación de la integridad, garantía de persistencia en el tiempo de colecciones de trabajos digitales, de manera que estén disponibles y sea posible económicamente, su uso por una comunidad o conjunto de comunidades.”

Debemos resaltar de esta definición la concepción de la Biblioteca Digital como:

- organización o institución, en lugar de simple conjunto de recursos electrónicos a los que se le aplica una solución tecnológica que permita crear, recuperar y usar la información
- la implicación de personal especializado, en lugar de comunidades de usuarios que de forma autónoma gestionan y cubren sus necesidades de información
- continuidad de las funciones de la biblioteca, destacando la preservación de los documentos.

La Association of Research Libraries (ARL), por su parte, incide en la dimensión globalizadora de la Biblioteca Digital, caracterizándola como

- Ente no aislado, integrado por diversos conjuntos de colecciones creadas y administradas por diferentes organizaciones: colaborar, reutilizar
- Requiere tecnologías específicas para compartir y enlazar recursos dispersos.
- Los vínculos entre las diversas colecciones y servicios de información deben ser transparentes para el usuario.

La Biblioteca Digital, en cualquier caso, entendida como la biblioteca del futuro, deberá basarse en lo aprendido y hecho en el pasado por la biblioteca

tradicional; tendrá que integrar formatos y contenidos analógicos con colecciones digitales, externas y propias, fruto de procesos de digitalización; podrá poseer o simplemente dar acceso a la información; necesitará gestionar los nuevos espacios reales y también los virtuales, etc. De esta idea surge la noción de “biblioteca híbrida”, que es la que mejor define en estos momentos a la mayor parte de las Bibliotecas Digitales. Hay que señalar que las bibliotecas híbridas no parecen ser un estado de transición temporal en la evolución hacia una Biblioteca Digital “pura”, sino que se están convirtiendo de hecho en un tipo específico de Biblioteca Digital.

La Biblioteca Digital en el ámbito universitario

Las bibliotecas orientadas a dar servicio a la comunidad universitaria son quizás las que más se han desarrollado en el terreno de las Bibliotecas Digitales, muchas veces adoptando el esquema de biblioteca híbrida. Los proyectos innovadores relacionados con las bibliotecas digitales de las universidades se han ido consolidando progresivamente durante esta década, llegando a convertirse en uno de los recursos más importantes para el aprendizaje y la investigación.

Es difícil encontrar en la actualidad una biblioteca universitaria en España sin un proyecto digital. Entre los muchos existentes, podemos destacar algunos pioneros como la Biblioteca Digital Dioscórides (Universidad Complutense de Madrid), Biblioteca Virtual Miguel de Cervantes (Universitat d’Alacant y Banco de Santander), la Biblioteca Digital de Catalunya (CBUC) o la Biblioteca Digital de la Universitat de València (precedente de Somni), todos ellos dirigidos a ofrecer acceso público a patrimonio bibliográfico. Por su parte, REBIUN lleva organizando anualmente desde 2001 un Workshop sobre Proyectos Digitales que ha constituido un foro de estudio y de debate relevante sobre la cuestión para todas las bibliotecas universitarias de España.

Desde el punto de vista de los contenidos, la Biblioteca Digital universitaria se ha ampliado y ahora aparecen dos grandes grupos de información claramente diferenciados:

- por una parte la información electrónica de pago que se adquiere por medio de suscripción y compra a los grandes editores y productores;
- por otra, la producción propia de documentos electrónicos y colecciones (por ejemplo, procesos de digitalización realizados con el objetivo de compatibilizar la difusión del patrimonio bibliográfico con su conservación) pero cada vez más la aportación de la documentación digital que se crea y distribuye en las múltiples plataformas educativas por profesores y estudiantes, tanto a nivel de docencia como de investigación.

Aunque toda esta documentación se ofrezca a texto completo al usuario y en formato compatible y fácilmente exportable, cada una de estas dos grandes tipologías tiene mecanismos y procesos claramente diferenciados en su gestión.

La introducción de las TIC en los procesos docentes, en particular del *e-learning*, y el uso intensivo de las tecnologías como estrategia indispensable para la creación del EEES facilita la producción de materiales electrónicos, nuevos objetos complejos, dinámicos. Objetos de aprendizaje como, por ejemplo, datos en bruto, simulaciones, anotaciones, pequeñas unidades de información en formato multimedia (muy alejadas del tipo documental libro, revista, artículo o capítulo) que permiten al estudiante comprender una idea, mostrar un problema y su solución. El *e-learning* requiere atomizar y distribuir la información hasta límites insospechados.

Desde el punto de vista del trabajo bibliotecario, estos objetos digitales deberían ser tratados, organizados, almacenados y preservados con la misma rigurosidad que los documentos en soporte papel u otros formatos clásicos, pero las condiciones en que esa información se crea, se usa y se comunica son radicalmente diferentes. Como respuesta a estas expectativas, en el ámbito de la Biblioteca Digital se crean nuevas plataformas tecnológicas como los repositorios institucionales y redes internacionales de repositorios que soportan ese trabajo de colaboración e integración. Es necesario establecer las condiciones para que los repositorios digitales sean capaces de tratar con objetos heterogéneos y complejos, así como para permitir que estos objetos puedan ser accesibles y reutilizables en escenarios también diversos, de una forma transparente para el usuario.

Las iniciativas promovidas desde el mundo académico relacionadas con el acceso abierto a la información y los resultados científicos, y el desarrollo de proyectos de “software libre” se están extendiendo prácticamente en todas las universidades y también en las bibliotecas. Las cuestiones legales de derechos de autor y copyright en el entorno del acceso abierto a la información científica se perfilan como esenciales para la Biblioteca Digital. Una de las líneas de actuación más afectada por el tema del copyright son los proyectos de digitalización, y cómo se deberían tratar aspectos tales como el uso justo o las obras huérfanas, qué modelos económicos se requieren en la era de la digitalización en masa, o cómo afectará la economía de la digitalización al movimiento de acceso abierto a la información.

Tras el último Workshop sobre Proyectos Digitales se ha puesto de manifiesto además la necesidad de garantizar el acceso permanente a la información digital mediante la creación de políticas de preservación institucionales, tal y como lo demuestran diversas iniciativas internacionales y europeas. Teniendo en cuenta que en el seno de REBIUN aún no se han diseñado estrategias que

traten de favorecer esta actividad, se ha considerado necesario llevar a cabo diversas actuaciones, que han sido tratadas durante una sesión de trabajo del Grupo de la Línea 2 del Plan Estratégico de REBIUN y que han sido incorporadas dentro de los objetivos operacionales del año 2009.

Las bibliotecas universitarias, en fin, están bien posicionadas y tiene una idea clara de su papel, pero el futuro, aunque prometedor, no está exento de amenazas. Es esencial adoptar una visión estratégica que se anticipe a los cambios y transformaciones que se están produciendo, no sólo en las bibliotecas sino en el entorno de la ciencia y la educación, y que garantice que será posible seguir cumpliendo una función relevante en la sociedad.

Integración e interconexión de recursos electrónicos

La gestión de las colecciones electrónicas es un gran reto que actualmente las bibliotecas y centros de documentación deben de afrontar. La incorporación de estos materiales ha traído consigo la necesidad de replantearse la forma en que las bibliotecas los seleccionan, adquieren, describen y ponen a disposición de sus usuarios. Si bien a grandes rasgos existen unos principios comunes en la gestión de los documentos tradicionales y los electrónicos, las características de éstos han obligado a reconsiderar algunos aspectos. Temas como la posesión frente al acceso, la complejidad de las políticas de licencias y las dificultades de establecer su precio han ocupado numerosas páginas en la bibliografía académica y profesional. Los aspectos relacionados con su descripción y la aplicación de esquemas tradicionales de metadatos como MARC 21 o alternativas como Dublin Core o MODS también han sido ampliamente tratados en distintas monografías y encuentros de trabajo. Otro tema debatido es la responsabilidad sobre la preservación futura de los contenidos y cómo deben los proveedores garantizar el acceso futuro a los materiales cuando la suscripción no se renueva.

Sin embargo, la integración e interconexión del impresionante volumen de recursos electrónicos heterogéneos, dinámicos y difusos, recursos aislados sin posibilidad de interconexión y navegación entre unos y otros, resulta crítica para seguir cumpliendo con las funciones de la biblioteca en el siglo XXI: integrar la información en el entorno del usuario como lo hace el catálogo.

Puesta a disposición

La contratación del acceso a un recurso electrónico (o grupo de ellos) debe seguirse de su puesta a disposición para los usuarios. Esto se alcanza mediante una serie de actividades:

- Catalogación
- Publicación de estas descripciones en un OPAC tradicional, mediante páginas o directorios web
- Resolución de aspectos relativos a la configuración del acceso.

Sobre la configuración del acceso, caben distintas alternativas, como acceder según nombres de usuario y contraseñas, o el sistema más habitual basado en el filtrado de direcciones IP. Este modelo hace que sólo pueden acceder los usuarios conectados a un ordenador de la red institucional (o vinculado a la red institucional a través de una VPN). El sitio web del proveedor donde se alojan los contenidos comprobará desde qué ordenador se está cursando la petición; si ésta se hace desde un ordenador conectado a la red informática de la institución que lo ha contratado, el proveedor establecerá la sesión y permitirá la consulta. En caso contrario, la petición será denegada. La forma en que se ha configurado y negociado el acceso a los recursos electrónicos remotos debe ser conocida en todo momento por el personal de la biblioteca encargado de su gestión. Si la biblioteca opta por mantener los contenidos contratados dentro de su red institucional (en este caso no se haría a través de una conexión remota sino mediante la red corporativa del centro en la que se mantienen copias electrónicas de los documentos), también se debe proceder a la organización de la infraestructura técnica necesaria.

La catalogación y publicación de las fuentes electrónicas ha venido realizándose de dos formas:

- a) directorios publicados en la web del centro (listas A-Z, por ejemplo) e
- b) integrando las descripciones dentro del OPAC tradicional.

Partidarios y detractores de una u otra opción han expuesto sus ventajas e inconvenientes, aunque como conclusión general suele afirmarse que no son opciones alternativas excluyentes sino complementarias, y que la combinación de las dos modalidades de acceso ofrece mejores resultados para los usuarios.

En el caso de los directorios, los usuarios pueden consultar las llamadas listas A-Z, páginas web donde las fuentes electrónicas contratadas se presentan en orden alfabético. Para cada título en el listado, se suele indicar una breve descripción de su contenido, materias de las que trata, y su cobertura cronológica (la que tiene contratada el centro). Desde aquí el usuario puede hacer clic en el título en cuestión para navegar al sitio web del proveedor y comenzar su sesión de trabajo. Normalmente se complementa con otra lista de recursos agrupados por categorías temáticas (que no materias) y ofrece la posibilidad de buscar por palabras contenidas en los títulos o descripciones de los recursos-e. Las listas A-Z ofrecen un mecanismo de acceso rápido y eficiente aunque se debe considerar la posibilidad de incorporar al catálogo tradicional registros para estas fuentes, para que las búsquedas en el OPAC

muestren al lector todos los títulos disponibles a través del centro, electrónicos o no, sean parte de las colecciones locales o de las disponibles remotamente. Esto obliga a incorporar en el catálogo registros MARC para los contenidos digitales. Como conclusión, para ofrecer al usuario las mayores garantías de acceso, la biblioteca debe realizar el doble trabajo de mantener las listas A-Z a la par que los registros catalográficos del OPAC.

Mantenimiento de la colección electrónica

En el caso de las revistas electrónicas, los bibliotecarios también deben mantener información actualizada sobre el estado de la colección y la disponibilidad de nuevos números conforme éstos se publican (con las bases de datos de revistas-e se presenta una situación similar). El control ejercido con el tradicional kárdex utilizado en las bibliotecas para controlar la recepción de los números de las publicaciones impresas también es necesario con las digitales. De esta forma no se registrará la llegada de los nuevos números al centro, sino la disponibilidad en línea de un nuevo número en el sitio remoto del proveedor. En centros donde se trabaja con un alto número de revistas-e y bases de datos de distintos proveedores, mantener esta información actualizada se convierte en un esfuerzo y costes sumamente elevados. Es también una actividad realmente complicada incluso si se cuenta con una estrecha colaboración por parte de los proveedores. Por otra parte puede darse el caso de que éstos modifiquen los títulos distribuidos a través de una base de datos, con el consiguiente problema para conocer qué está realmente disponible para los usuarios del centro.

En respuesta a las dificultades que acarrea el mantenimiento de la cobertura de los recursos electrónicos y controlar su evolución, en los últimos años han surgido empresas cuya principal actividad es mantener y distribuir a los centros usuarios de este material, información sobre la cobertura de las bases de datos de distintos proveedores y de las actualizaciones que se producen en sus contenidos y títulos. Facilitan listas A-Z con los contenidos que tienen suscritos, y se encargan de su actualización, liberando a las bibliotecas de esta actividad. Normalmente, las listas A-Z son páginas personalizadas para cada centro (tanto en sus contenidos como en su aspecto y formato), hospedadas en los servidores web de esas empresas, que actúan como mediadoras entre proveedores y clientes. También suelen facilitar registros en formato MARC para su integración en los OPACs. Entre esas empresas destacan Serials Solutions, Gold Rush o TDNet. Obviamente los proveedores de recursos-e también ofertan servicios similares para las bibliotecas, como por ejemplo el Ebsco A-to-Z Service. En numerosos sitios web de centros bibliotecarios universitarios españoles podemos ver que sus listados de recursos electrónicos nos redirigen a páginas alojadas en algunas de estas empresas. También existen desarrolladores de software como Innovative Interfaces Inc., que ofrece complementarios para sus herramientas, como CASE (Content Access Service)

tanto para datos de cobertura (Coverage Data Service) incluyendo el soporte a los servicios A-Z, como para datos MARC (Marc Record Service). También Ex Libris Inc. ofrece servicios como Citation Linker de SFX, que permite acceder al texto completo digital y/o generar búsquedas de colecciones impresas en catálogos colectivos o locales a partir de una simple referencia.

Uso de los recursos-e

Una vez se han puesto a disposición de los usuarios, la biblioteca debe esforzarse por conocer el uso real que se hace de los mismos y evaluar la rentabilidad de las suscripciones. Tienen un coste elevado y su uso constituye un aspecto decisivo a la hora de decidir renovarlas o cancelarlas. Los estudios de utilización también permiten evaluar la rentabilidad de suscribir determinados paquetes de revistas-e o servicios. Los proveedores suelen ofrecer acceso conjunto a publicaciones sumamente atractivas frente a otras que pueden resultar menos interesantes para la comunidad de usuarios. Disponer de un alto número de títulos a través de un mismo paquete puede resultar práctico para la biblioteca, pero quizás el uso real que se haga de ellos no justifique una inversión que inicialmente resultaba atractiva.

En resumen, se debe conocer con la mayor exactitud posible el uso que los usuarios hacen de la colección de recursos digitales. En este área también se han desarrollado importantes iniciativas, como el proyecto *Counter* o el estándar *Sushi* (Standardized Usage Statistics Harvesting Initiative), dos proyectos complementarios.

Por una parte, el primero se constituye como una forma de medir el uso que se hace de los recursos-e. Para ello establece una definición precisa de las actividades que se van a medir (búsquedas, accesos, descargas, etc.) y la forma en la que éstas acciones deben totalizarse. Se definen también unos informes normalizados para medir el uso que se hace en distintos períodos de tiempo. En *Counter*, la responsabilidad de obtener las métricas recae en los proveedores. Para evitar posibles manipulaciones de los datos (se podría caer en la tentación de informar a las bibliotecas de un uso mayor del que realmente están haciendo los usuarios), *Counter* establece un proceso de auditoría realizado por un tercero, independiente del proveedor. En este proceso de auditoría se verificará la forma en que éste obtiene las medidas para asegurar que son correctas y se han conseguido siguiendo las pautas dictadas por esta iniciativa. En la actualidad los principales proveedores de bases de datos y recursos-e han declarado su conformidad con *Counter* y ofrecen a sus clientes la posibilidad de obtener este tipo de informes de uso descargándolos del sitio web del proveedor (normalmente en formato Microsoft Excel). La biblioteca cuenta así con la información necesaria para conocer qué títulos se consultan con mayor asiduidad, y qué bases de datos constituyen las principales fuentes de información para sus usuarios. Una de sus grandes ventajas es que todos

los que aplican este método totalizarán los datos de la misma forma, lo que asegura que son comparables y se han obtenido siguiendo un mismo procedimiento.

Respecto a *Sushi* se trata de una iniciativa complementaria a *Counter* que establece un mecanismo para descargar los informes *Counter* de los sitios web de los proveedores de forma automática y desatendida. Su desarrollo se inició en junio de 2005 por un grupo de trabajo formado por la Cornell University, Ebsco, Ex Libris, Innovative, Swets y la Washington University, a los que posteriormente se añadió Thomson Scientific. Su finalidad es que sea una aplicación informática la que se encargue de conectarse regularmente a los sitios web de distintos proveedores y descargar los informes correspondientes al centro. De esta forma el bibliotecario no tendrá que dedicar tiempo a descargar uno tras otro los informes de los distintos proveedores, y puede centrar su atención a realizar el análisis de los datos. Se convirtió en la norma norteamericana Z39.93 en julio de 2007.

Tecnologías para la integración e interconexión de RE

La gestión de colecciones electrónicas ha supuesto la necesidad de desarrollar tecnologías complementarias a las que tradicionalmente venían usándose en las bibliotecas. Hoy es frecuente oír hablar de sistemas de resolución de enlaces (Links Resolvers), metabuscadores (Federated Search Engines) y de protocolos o estándares como OpenURL, sistemas de gestión de recursos electrónicos (Electronic Resource Management) y portales de descubrimiento de información (Discovery Tools).

Sistemas de resolución de enlaces (Links Resolvers) - OpenURL

Los sistemas de resolución de enlaces son sin duda alguna una de las ideas más atractivas desarrolladas en los últimos años en el área de la documentación. Solucionan uno de los aspectos más complejos de la gestión de contenidos electrónicos: la multiplicidad de accesos a un mismo título y el problema de la copia apropiada.

Para introducir brevemente el propósito de estos sistemas, plantearemos un escenario habitual en muchos centros. Supongamos que una biblioteca cuenta con distintas suscripciones a bases de datos de revistas electrónicas y que una de estas publicaciones está disponible a través de tres bases de datos diferentes, posiblemente con distintas coberturas y derechos de uso en cada una de ellas. Un sistema de resolución de enlaces permitiría al usuario de un OPAC (o de una lista A-Z), hacer clic sobre el título y obtener una nueva página donde se le indica a través de qué proveedores y de qué bases de datos tiene acceso, su cobertura y derechos de uso. Con esta información el usuario puede seleccionar la opción que le resulte más conveniente y navegar al sitio web del proveedor seleccionado. Esta interacción, que a priori puede parecer algo trivial, encierra una importante complejidad técnica. En primer lugar es necesario conocer a través de qué bases de datos y proveedores está disponible la revista en cuestión, su cobertura y condiciones de acceso en cada caso (esta

información es muy similar a los datos que facilitan los proveedores de listas A-Z antes citados: Serials Solutions, TDNet, CASE de Innovative Interfaces, etc.). Por otra parte el sistema desde el cual el usuario realiza su consulta debe ser capaz de acceder a la información anterior y de alguna forma se le tienen que presentar los enlaces correspondientes. Finalmente el sistema del proveedor debe tener la opción de responder a la petición de acceso que recibe cada vez que el usuario hace clic en un enlace que le dirija a su sitio web. Es necesario contar con una utilidad informática que medie entre la aplicación origen con la que interactúa el usuario final, y las aplicaciones y bases de datos de los proveedores que ofrecen acceso al texto completo de los artículos y revistas-e. Esa herramienta intermedia es el sistema de resolución de enlaces. La dificultad estriba en que tanto el origen de los mismos (es decir, los OPACs o listas A-Z en los que busca el usuario) como los sistemas de los proveedores son dispares, y a priori desconocidos unos de otros, por lo que establecer enlaces directos sería imposible.

Además se debe contar con una forma normalizada de enlazar con los títulos ofertados por los distintos proveedores de los recursos-e. Aquí entra en juego el estándar OpenURL. Su origen está en el trabajo de Herbert van de Sompel iniciado en 1999, y se aprobó como estándar norteamericano ANSI Z39.88 en el año 2005 (antes de esta fecha, OpenURL se había convertido en un auténtico estándar *de facto*). Establece una forma de hacer referencia a un recurso-e disponible en la web, usando para ello sus metadatos. Podemos decir que un vínculo basado en este sistema es un enlace web en el que en lugar de utilizar la dirección URL se usan sus metadatos para hacerle referencia.

Un ejemplo para una revista electrónica podría ser éste:

<http://sfxserver.uni.edu/sfxmenu?issn=1234-5678&date=1998&volume=12&issue=2>

Vemos que junto al nombre del servidor (sfxserver.uni.edu) al que se dirige la petición (sfxmenu?) se añaden metadatos sobre el recurso-e: su código ISSN (1234-5678), fecha (1998), volumen (12) y número (2). Como estos metadatos constituyen el contenido básico de un enlace OpenURL, se suele decir que este estándar constituye una forma de "transportar metadatos a través de la web". El enlace OpenURL anterior no se dirigirá directamente al sitio web del proveedor del recurso, sino a una aplicación informática complementaria llamada sistema de resolución de enlaces (link resolver). El sistema de resolución de enlaces interpretará los metadatos incluidos en la petición, comprobará la disponibilidad del recurso solicitado en los distintos proveedores de recursos-e, y devolverá al usuario una página web con el listado de accesos (esta vez, los enlaces apuntarán a los sitios web y bases de datos de cada uno de los proveedores que se hayan identificado). El sistema de resolución de enlaces establece así un puente entre una aplicación origen (normalmente un OPAC o una lista A-Z, aunque podría ser cualquier otra, incluso una base de datos de un tercero) y las bases de datos de los distintos proveedores que ofrecen acceso a los recursos-e o a algún servicio relacionado con ellos (texto completo, obtención de documentos, préstamo interbibliotecario, etc.). Su funcionamiento se basa en:

- la capacidad de recibir e interpretar peticiones OpenURL procedentes de la aplicación origen,
- la disponibilidad de información sobre la cobertura de los títulos en los distintos servicios y bases de datos contratados por el centro, y

- la capacidad de presentar al usuario enlaces que apunten a los recursos-e hospedados en los sitios de los proveedores remotos, por lo que el servidor de los mismos debe conocer la forma que usa cada proveedor para enlazar a sus recursos.

OpenURL es un mecanismo genérico para enviar metadatos a través de la web entre servicios de información. El estándar permite no sólo enviar información sobre los metadatos del recurso. También se podría enviar en una petición OpenURL datos sobre el tipo de usuario y el contexto en el que está realizando su trabajo. De esta forma el servidor de enlaces contaría con más información sobre el tipo de usuario, y podría así descartar –por poner un ejemplo– aquellos recursos a los que no tiene capacidad de acceder. Una práctica intuitiva cuando se codifica una referencia bibliográfica en una página web sería hacer que el enlace apuntase a la página del editor donde se encuentra el texto completo. Sin embargo, puede darse el caso de que el usuario del centro no tenga acceso a ese documento, aunque sí puede tener acceso a ese mismo contenido a través de la base de datos de un agregador (ésta sería la copia apropiada). Con OpenURL, el servidor de enlaces recibe los metadatos necesarios para poder identificar cuál es la más apropiada para cada usuario. Debemos considerar también que el destino de un enlace puede ser un servicio gratuito o de pago, un servicio donde el texto del documento esté disponible en línea o un servicio de obtención de documentos. Los servidores de enlaces deben ofrecer la posibilidad de discriminar en estos casos cuál es la opción preferible, y hacer que los usuarios seleccionen la adecuada.

Los sistemas de resolución de enlaces hacen que la aplicación origen únicamente tenga que saber cómo generar un enlace en la sintaxis que establece la norma OpenURL, y la dirección en internet del sistema de resolución. Disponemos de numerosos programas servidores de enlaces. Quizás el más popular sea SFX, de Ex-Libris, ya que fue el sistema informático realizado por el creador de esta idea (Van de Sompel) durante su trabajo experimental en la Universidad de Gante. La oferta de Innovative Interfaces Inc. incluye uno llamado WebBridge LR (Link Resolver) con su complementario PathFinder Pro para recursos de destino. Serials Solutions también ofrece servicios de resolución de enlaces OpenURL a través de 360 Link.

Búsquedas federadas (Federated search)

La segunda gran iniciativa tecnológica asociada a la gestión de los recursos-e son los sistemas de búsqueda federada o simultánea. Conforme una biblioteca aumenta el número de fuentes digitales a las que da acceso, surge la necesidad de ofrecer a los usuarios un mecanismo por el que puedan recuperar documentos relevantes de distintas bases de datos y proveedores. Los sistemas de búsqueda federada tienen como interfaz una página de consulta que permite consultar simultáneamente bases de datos y revistas-e hospedadas en sitios web de distintos proveedores, recursos internos, etc. Los usuarios pueden interactuar y consultar múltiples bases de datos desde un único punto. El objetivo es redirigir la petición del usuario a múltiples sitios, recibir los resultados e integrarlos en una única lista en la que “idealmente” se eliminarán posibles duplicados. Para que esto sea posible, los proveedores deben exponer al exterior sus catálogos y dar la posibilidad

de recibir búsquedas a su sistema desde otras aplicaciones informáticas. Su funcionamiento puede verse como algo similar a un cliente Z39.50 usado para consultar simultáneamente varios catálogos, aunque existe una diferencia importante: los proveedores de recursos-e y bases de datos pueden exponer sus datos usando distintos protocolos o normas, no sólo Z39.50. En este sentido contamos con iniciativas que han alcanzado la madurez, como SRW/SRU (Search and Retrieval Web/URL) de la Library of Congress norteamericana y MetaSearch de NISO (US National Information Standards Organization) basada en el anterior. SRW/SRU puede definirse como una adaptación de Z39.50 a la web, basada en el intercambio de peticiones y respuestas XML, que simplifica en gran medida la programación de este tipo de soluciones. Aparte de soportar protocolos normalizados como Z39.50 o SRW/SRU, los metabuscadores pueden implementar otros protocolos de tipo propietario, específicos de cada proveedor, normalmente basados en el intercambio de peticiones y respuestas en XML.

Quizás el metabuscador más conocido sea MetaLIB, de Ex Libris. Innovative comercializa otro denominado Research Pro (que es la evolución de MetaFind). Serials Solutions también ofrece servicios de búsqueda federadas a través de 360 Search.

CrossRef/DOI (Digital Object Identifier)

La gran mayoría de los grandes servicios de revistas digitales establecen enlaces con otros recursos relacionados. Desde las referencias o citas, se establecen enlaces directamente al artículo citado o a los registros relacionados con el tema en una base de datos en línea o a todos los artículos de un autor en dicha base de datos, o incluso a servicios de compra de artículos en línea, el OPAC de la biblioteca, etc. En un contexto de integración de recursos electrónicos, las citas a referencias electrónicas son cada vez más abundantes, pero con frecuencia son poco estables. Para poder llevar a cabo estos proyectos de interconexión multinivel, resulta fundamental utilizar métodos de identificación permanente de los recursos electrónicos como DOI (Digital Object Identifier : <http://www.doi.org/>).

El DOI fue creado en 1998 para identificar textos, audios, vídeos, imágenes, software, etc. un ambiente digital. La información acerca de un objeto digital puede cambiar a lo largo del tiempo, incluso dónde encontrarlo. Sin embargo, su número DOI permanece.

El DOI es una cadena alfanumérica única que tiene un doble objetivo:

- Identificar o nombrar de forma única un objeto digital
- Servir como un vínculo persistente y estable a la localización de ese objeto en el web

Es una implementación de los conceptos Uniform Resource Name (URN) y Uniform Resource Identifier (URI).

El DOI es el único identificador persistente adoptado de forma general para los trabajos académicos en línea. El DOI persiste por encima de los cambios en la propiedad intelectual (cambios de editor) o la localización porque es solamente un nombre para buscar una dirección en un directorio fácilmente actualizable. La funcionalidad principal del sistema DOI es resolver el DOI como una URL registrada y actualizada.

La sintaxis de los nombres DOI tiene rango de norma americana NISO Z39.84:2005 Syntax for the Digital Object Identifier. El sistema DOI lo desarrolla y mantiene la International DOI Foundation (IDF). IDF no asigna directamente números DOI, sino que utiliza varias “agencias de registro” (actualmente 8) que participan en la IDF. Las agencias están especializadas según regiones o áreas de interés.

La tecnología DOI se basa en el programa Handle System, desarrollado por la norteamericana Corporation for National Research Initiatives. El sistema DOI utiliza un servidor directorio de registro centralizado: cuando un usuario hace click sobre un DOI éste remite un mensaje al directorio central, devolviendo la ubicación actual asociada con el contenido descrito en el DOI redireccionando al navegador del usuario a la dirección correcta del contenido. En el momento en que un proveedor de información digital decida cambiar la ubicación de sus archivos, o que por ejemplo una revista científica pase a ser publicada por otra editorial, el productor inicial ha de informar a la Agencia de Registro de los cambios realizados para así poder actualizar debidamente el directorio de registro y evitar pérdidas de información.

CrossRef es una de las 8 agencias de registro antes mencionadas, y se presenta como la “agencia oficial de registro DOI para publicaciones académicas y profesionales, incluyendo revistas científicas, libros y otro tipo de contenidos”. Se trata de una iniciativa sin ánimo de lucro llevada a cabo por los principales editores de información científica, técnica y médica, agrupados desde el año 2000 en la Publishers International Linking Association (PILA). Cualquier editor de documentos primarios de investigación en formato digital (tanto si es grande como si es pequeño, comercial o sin ánimo de lucro, tradicional o no tradicional) puede registrar sus contenidos con CrossRef. También pueden participar otras organizaciones como productores de bases de datos referenciales y de resúmenes y agentes de suscripciones, a parte de bibliotecas y consorcios bibliotecarios, que pueden hacer uso de la base de datos de Crossref, para la mejora de la gestión de sus colecciones digitales. La afiliación a CrossRef es, además, gratuita para las bibliotecas.

Dicha iniciativa consiste, según sus propias palabras, en un servicio cooperativo de enlazado de referencias. Es decir, CrossRef proporciona un servicio que permite acceder a un artículo a texto completo, desde la cita o referencia mencionada en otro artículo diferente en formato digital, con un solo click y dentro de los fondos digitales de las entidades asociadas a la iniciativa.

En este contexto, podemos encontrar la tecnología DOI aplicada en:

- Documentos impresos (artículo de revista impresa)
- Sumarios (enlaces al Full Text)
- En revistas electrónicas a texto completo
- En citas y referencias
- En Google (sitio de un editor o la referencia a un artículo concreto)

Para terminar, diremos que DOI y OpenURL no son tecnologías contrapuestas, sino complementarias, y que los sistema de resolución de enlaces locales interactúan perfectamente con CrossRef, permitiendo al usuario localizar la copia adecuada en el contexto de la biblioteca a partir de un enlace DOI verdaderamente persistente.

Sistemas de gestión de recursos electrónicos (ERM)

Los sistemas de gestión de recursos electrónicos permiten gestionar la base de datos de proveedores, recursos-e y títulos, con toda la información necesaria sobre su cobertura, derechos y condiciones de uso. La información se organiza en torno a un registro para cada proveedor y base. Forma parte de este registro toda la información necesaria para la gestión del recurso-e a lo largo de su ciclo de vida: adquisición, términos de la licencia, suscripciones, personas de contacto, peticiones cursadas al soporte técnico, títulos que incluyen y su cobertura, información sobre embargos, duración de períodos de prueba, URL para la descarga de informes y estadísticas de uso, etc. Una función interesante son los informes para comprobar el solapamiento de títulos y cobertura entre bases. A estas funciones pueden unirse otras características como el acceso a la base de datos mediante listas A-Z y de categorías temáticas. Millennium, por ejemplo, permite al usuario consultar el catálogo simultáneamente con las base de datos desde la interfaz web del opac, pero no todos los ERM permiten su plena integración con el catálogo u otros módulos como adquisiciones. Estos módulos suelen interactuar con los sistemas de resolución de enlaces para que los usuarios puedan ser redirigidos a los distintos servicios desde los que tengan acceso a los títulos. Los datos sobre recursos-e gestionados por un ERM tienen también visibilidad a través de los metabuscadores.

Entre los productos disponibles destacamos el ERM de Innovative Interfaces Inc., desarrollado conforme los requisitos establecidos por ERMI (Electronic Resource Management Initiative) de la DLF (Digital Library Federation) norteamericana. Innovative fue una de las organizaciones participantes en esta iniciativa junto con otros grupos de usuarios y proveedores de sistemas informáticos, y la primera en comercializar un producto de estas características. Otros son ERMS de Sirsi/Dynix (pero desarrollado por Serials Solutions), y Verde, de Ex Libris. Serials Solutions también ofrece servicios de gestión de recursos electrónicos a través de 360 Resource Manager.

Portales de descubrimiento de información (Discovery Tools)

Los portales de descubrimiento de información, basados en el concepto de la Web 2.0, son una interfaces de consulta sobre un único OPAC o búsquedas federadas. Entre sus principios se encuentra el hecho de que un usuario no tiene por qué ser un experto en recuperación de información, y la necesidad de ofrecerle una experiencia basada en la inmediatez de la búsqueda, algo similar a lo que ofrecen sistemas como Google o Amazon. Estos portales establecen las bases para una mayor participación de los usuarios en torno al interfaz de consulta, al permitirles registrar comentarios y valoraciones sobre las revistas y bases de datos. Se incorpora también un sistema de *tagging* para que puedan etiquetar o asignar categorías o descriptores a los contenidos y mejorar así las opciones de recuperación (la asignación de descriptores o categorías por parte de los usuarios puede ser revisada por el personal de la biblioteca). Los resultados de una búsqueda también se pueden acotar o redefinir aplicando distintas facetas o criterios (como son el formato del documento, la colección, el idioma o los descriptores que se les ha asignado). Los mecanismos usados para ordenar los resultados aplica algoritmos propios distintos y más creativos a los factores usados tradicionalmente para calcular la relevancia. Se trata de un paso necesario en la evolución de los sistemas integrados de gestión bibliotecaria, que además de incorporar nuevas funciones y capacidades para gestionar distintos tipos de recursos y metadatos, también deben explorar métodos para aumentar la participación de las comunidades de usuarios y resolver los problemas derivados de la sobrecarga de información que vivimos en la actualidad.

Entre los productos disponibles destacamos el AquaBrowser de Browker, Encore de Innovative Interfaces Inc. o Primo de Ex Libris. Serials Solutions también ofrece un servicio similar a través de Summon (beta).

El movimiento “Open Access” y los repositorios de investigación

Introducción a OA

El movimiento “*Open Access*” o acceso abierto a las publicaciones científicas, o mejor dicho, a la producción científica en general (en la que se englobarían no sólo los artículos publicados en revistas sino otro tipo de documentos como objetos de aprendizaje, imágenes, datos o documentos audiovisuales, etc.), es aquel que permite el libre acceso a esos recursos digitales sin barreras económicas o restricciones derivadas de los derechos de copyright sobre los mismos. Este acceso es online a través de Internet, y salvo limitaciones tecnológicas y de conexión a la red del usuario no debería estar restringido por otro tipo de imposiciones.

Existen iniciativas pioneras en este sentido, como las primeras revistas online que aparecieron a principios de la década pasada, difundidas a través del correo electrónico como por ejemplo la publicación *The Public-Access Computer Systems Review*. En el año 1991 también se creó “Arxiv”, el reconocido repositorio temático en física, matemáticas y computación. Sin embargo, es en este último lustro cuando se inicia un verdadero movimiento internacional en torno al acceso abierto, en el que se empieza a fraguar el consenso para definir el concepto y marcar las vías para alcanzar sus fines.

El movimiento “*Open Access*” no sólo se fundamenta en aspectos económicos como la desproporcionada subida de precios de las revistas científicas, iniciada en la década de los 80, conocida como la *serials crisis* (crisis de las publicaciones seriadas), sino también a la pérdida de control de los derechos de copyright de los autores sobre sus trabajos publicados, a las imposiciones de las editoriales en políticas de acceso y comercialización (*big deals*), al avance de las TICs e Internet. Se puede decir que el movimiento “*Open Access*” es la respuesta de la comunidad científica y de la sociedad ante determinados abusos editoriales y, sobre todo, la asunción del potencial de Internet para el tráfico y comunicación libre de información.

Existe un compromiso social avalado por declaraciones de ámbito internacional que sostienen y perfilan la definición del *open access*. Las tres más importantes y de referencia obligada en la bibliografía, son la Declaración de Budapest (*Budapest Open Acces Initiative*, BOAI) de 2002, seguida de la Declaración de Bethesda (2003) y la Declaración de Berlín (*Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities*), también del año 2003.

- Budapest: La primera define *open access* como el libre acceso a través de Internet a la literatura científica, respetando las leyes de copyright

existentes, aunque se aboga porque sean los autores o las instituciones quienes retengan los derechos de copyright.

- Bethesda: En la segunda, además se menciona el archivo inmediato de los trabajos para facilitar este acceso en abierto.
- Berlín: La tercera fue suscrita en octubre de 2003 por diferentes representantes políticos y científicos, y en ella se manifiestan explícitamente las grandes posibilidades que brinda internet en la difusión del conocimiento, avala el paradigma de *open access*, y recoge los términos de las dos declaraciones anteriores. La Declaración de Berlín es especialmente relevante porque la adhesión a la misma significa un apoyo al paradigma del movimiento “*Open Access*” y compromete a las instituciones a favorecer las vías open para alcanzar el acceso abierto.

A partir de éstas se han sucedido numerosas declaraciones emitidas tanto por grupos de trabajo, sociedades profesionales o instituciones internacionales de apoyo y reconocimiento del valor intrínseco del *open access*, no sólo como vía de difusión del conocimiento y reconocimiento del crecimiento del impacto de los recursos digitales, sino por el valor que tiene para la preservación de los mismos a través de repositorios. Entre estas instituciones se encuentra la UNESCO, la IFLA, la OCDE o la Comisión Europea. En la *Timeline of the Open Access Movement*, que compila Peter Suber (última revisión 7 de noviembre de 2008) es posible consultar una relación exhaustiva de las declaraciones, iniciativas y apoyos que recibe este movimiento en todo el mundo. Ejemplo del avance del *open access* en España son las 21 instituciones españolas que han firmado la Declaración de Berlín, de las cuales 17 son universidades. La última, el pasado octubre, fue la Universitat de València.

Vías para alcanzar el acceso abierto

En la Declaración de Budapest se establecen dos rutas para alcanzar el *open access*:

- la “ruta dorada” o la de publicación en revistas *open access*, o
- la “ruta verde” que alude al archivo o depósito de recursos digitales en repositorios institucionales o temáticos.

Cuando el depósito de los ficheros lo hace el autor hablaríamos de auto-archivo (*self-archiving*). Es necesario hacer esta observación porque cuando se trata de repositorios temáticos el depósito generalmente lo hace el autor, mientras que en el caso de los institucionales, además de los autores, en muchos casos el depósito está ligado a personal relacionado directamente con el repositorio, a personal de la biblioteca o a personal de apoyo del grupo de investigación al que están vinculados los autores; en estos casos, hablaríamos de depósito “delegado”.

Las tipología de revistas *open acces* cada vez es más diversa debido a los nuevos modelos híbridos que adoptan las empresas editoriales. El caso más puro según la definición de *open access* sería aquella revista que ni lector ni autor paga por publicar y los autores retienen el copyright sobre sus trabajos, cediendo los derechos no exclusivos de publicación a la revista. Obviamente, alguien “tiene que pagar” la publicación ... En estos casos, las revistas generalmente pertenecen a instituciones académicas o sociedades profesionales cuyos recursos incluyen también el sostenimiento de estas publicaciones. Ejemplos de este tipo de revistas se recogen en el directorio de revistas *open access* (DOAJ, Directory of open access journals), que en este momento cuenta con casi 4.000 revistas en su base de datos.

Existen casos de revistas totalmente *open access* en que el autor o la institución paga por su publicación, como las revistas de BioMed Central, editor online especializado en todas las áreas de la medicina y la biología, o de Public Library of Science (PLoS), una organización sin ánimo de lucro. Entre uno y otro modelo se dan casi todas las posibles combinaciones entre pago y derechos de copyright.

Cuando coexisten artículos *open access* junto con los de pago por suscripción en una misma revista, estaríamos hablando de “modelos híbridos”. Hoy en día las grandes editoriales como Elsevier, Springer, Blackwell, etc. ofrecen este tipo de modelos híbridos en algunas de sus revistas; es decir, ofrecen la posibilidad de pagar por publicar a cambio de que el artículo quede en acceso abierto entre otros que no lo están, lo que comunmente se conoce como “*open choice*”. Los precios varían de editorial a editorial y de revista a revista, aunque normalmente están sobre los 2.500 €: hablamos de iniciativas, programas, servicios o modelos como “Sponsored article” de Elsevier, “Oxford open” de Oxford University Press o “Funded access service” de John Wiley & Sons.

Todo parece indicar que las dos vías que ha encontrado el *open access* para expresarse, los repositorios y las revistas, más que vías alternativas (como algunos debates parecen dar a suponer), probablemente serán complementarias. Ni las revistas OA ni los repositorios parecen tener el camino totalmente despejado para su desarrollo futuro.

Repositorios de investigación

Los repositorios, entendidos como archivos donde se almacenan recursos digitales (textuales, de imagen o sonido, en general llamados objetos digitales) surgen de la llamada comunidad *e-print*, preocupada por maximizar la difusión y el impacto de los trabajos depositados en los mismos. Si el repositorio responde a unas áreas de conocimiento en particular, hablaremos de repositorios temáticos, y si los objetos digitales almacenados responden a los de una institución hablaremos de repositorios institucionales.

Un Repositorio Institucional es una base de datos compuesta de un grupo de servicios destinados a capturar, almacenar, ordenar, preservar y redistribuir la documentación académica de la Universidad en formato digital. La organización SPARC Europe define los Repositorios Institucionales como sigue:

- Pertenecen a una institución.
- Son de ámbito académico.
- Son acumulativos y perpetuos.
- Son abiertos e interactivos.

Cuando decimos que los Repositorios Institucionales son abiertos e interactivos queremos decir que cumplen con el OAI y permiten acceso abierto a la documentación académica.

Clifford Lynch define los RIs de la siguiente forma: “Un Repositorio Institucional universitario es un conjunto de servicios que ofrece la Universidad a los miembros de su comunidad para la dirección y distribución de materiales digitales creados por la institución y los miembros de esa comunidad. Es esencial un compromiso organizativo para la administración de estos materiales digitales, incluyendo la preservación a largo plazo cuando sea necesario, así como la organización y acceso o su distribución” (“Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age” ARL, no. 226 (February 2003): 1-7).

Open Archives Initiative: OAI

La Iniciativa de Archivos Abiertos (OAI) proporciona la arquitectura y especificaciones técnicas necesarias para que productores y distribuidores de documentos de acceso abierto pongan a disposición de agregadores de contenidos metadatos sobre los documentos que almacenan con objeto de hacerlos visibles y accesibles a los investigadores.

OAI viene a solucionar el problema de la dispersión de los documentos en múltiples depósitos institucionales, temáticos, revistas individuales, portales de revistas electrónicas, etc. Desde el punto de vista del investigador, hace posible la integración de documentos de una misma temática, área geográfica, tipología, etc. en una única herramienta que es capaz de proporcionarle, además, una serie de valores añadidos, como selección, búsqueda e identificación, evaluación, etc.

OAI surgió a finales de los años noventa a partir de los servidores de documentos en acceso abierto que habían aparecido en distintas disciplinas científicas: arXiv en Física, RePEc en Economía, CogPrints en Psicología, NCSTRL en Informática y NDLT para tesis. Su objetivo inicial fue estudiar la interoperabilidad de los distintos servidores con objeto de facilitar el

intercambio de datos entre los mismos. El nacimiento de la iniciativa se sitúa en la Convención de Santa Fe celebrada en la ciudad norteamericana del mismo nombre en Octubre de 1999.

La iniciativa se concretó en un protocolo para la comunicación de metadatos denominado OAI-PMH (Open Archives Initiative–Protocol for Metadata Harvesting) cuya primera versión apareció en Enero de 2001. Aunque inicialmente se creó para ser aplicado a depósitos de documentos en acceso abierto muy pronto se vio que podía implementarse sobre cualquier sistema que requiriese la comunicación de metadatos. Esto hizo afirmar, con cierta euforia, a alguno de los creadores del mismo que OAI-PMH estaba llamado a ser a las bibliotecas digitales lo que HTTP había sido a la web.

Podríamos señalar tres características fundamentales del protocolo:

- Simplicidad: Se concibió bajo la premisa de la sencillez. Conscientes de los problemas de implementación que habían tenido otras iniciativas anteriores como el Z39.50, los creadores buscaron una fórmula simple que estuviera al alcance de cualquier potencial implementador.
- Normalización: Basado en estándares ampliamente utilizados en Internet como son el protocolo http (HiperText Transport Protocol) para la transmisión de datos y órdenes, y XML (Extended Markup Language) para la codificación de los metadatos.
- Recolección. Frente a otros sistemas de agregación de contenidos como la búsqueda distribuida (Z39.50) o los sistemas de sindicación de contenidos vía RSS, OAI-PMH ha optado por la recolección de metadatos. En este caso, existe una entidad que pone a disposición de los interesados información bibliográfica sobre los documentos que almacena. Estos, normalmente agregadores de contenidos, recogen periódica y sistemáticamente todos o parte de los metadatos expuestos para, localmente, implementar servicios de valor añadido.

OAI-PMH sigue el principio de que existen múltiples proveedores de datos (*data providers*) que comparten información con múltiples proveedores de servicios (*service providers*) a través de un protocolo común. Los primeros son los depósitos de documentos que proporcionan los metadatos de los documentos que almacenan y los segundos son los recolectores o agregadores de contenidos, que toman los datos con el objetivo de incorporarles algún valor añadido y presentarlos a los usuarios finales. Entre los valores añadidos que se pueden ofrecer tenemos: sistema de búsqueda e identificación, filtrado, alertas temáticas, medición del uso e impacto de los documentos, etc.

Es importante matizar que OAI-PMH trata exclusivamente de la comunicación de metadatos, y no de los textos completos de los documentos que se referencian.

La comunicación entre proveedores de datos y proveedores de servicios se realiza utilizando transacciones HTTP para emitir preguntas y obtener respuestas. El *service provider* puede pedir al *data provider* que le envíe metadatos según determinados criterios como por ejemplo la fecha de creación de los registros. En respuesta, el *data provider* envía un conjunto de registros en formato XML. Las peticiones se emiten a través de los métodos GET y POST del protocolo HTTP y constan de una lista de argumentos con la forma de pares del tipo: clave=valor.

Existen solamente seis peticiones que un PS puede realizar a un PD:

- Identify: utilizado para recuperar información sobre el PD: nombre, versión del protocolo que utiliza, dirección del administrador, etc.
- ListMetadataFormats: devuelve la lista de formatos bibliográficos que utiliza el servidor.
- ListIdentifiers: recupera los encabezamientos de los registros en lugar de los registros completos. Tiene un argumento obligatorio (*metadataprefix*) para especificar el formato de metadatos en el que se quiere obtener la respuesta. Permite, además, argumentos como el rango de fechas entre los que queremos recuperar los datos.
- ListRecords: igual que el anterior pero recuperando los registros completos.
- GetRecord: utilizado para recuperar un registro concreto. Necesita dos argumentos: identificador del registro solicitado (*identifier*) y especificación del formato de metadatos (*metadataprefix*) en que se desea obtener.
- ListSets: recupera un conjunto de registros. Estos conjuntos son creados opcionalmente por el servidor para facilitar una recuperación selectiva de metadatos. Se trata de una clasificación de los contenidos según diferentes criterios como materias, lenguaje, tipología documental, etc. Un PS puede pedir que se recuperen sólo los metadatos de una clase específica.

El protocolo requiere además que los metadatos se expresen en formato Dublin Core no calificado. No obstante esto es un mínimo común denominador puesto que es posible además que cada proveedor de datos utilice formatos adicionales que permitan una descripción más rica de los documentos como por ejemplo MARC. Un proveedor de servicios puede pedir que los registros se le sirvan en cualquiera de los formatos soportados por el proveedor de datos.

Plataformas para la creación de repositorios

Existe una amplia comunidad científica involucrada en el desarrollo de plataformas tecnológicas para la creación de repositorios *open access*, la utilización de una o de otra responderá a las necesidades y recursos de la institución que lo ponga en marcha. Existe una revisión bastante exhaustiva sobre software libre para repositorios publicada por el *Open Society Institute*

(OSI) en la cual se analizan y comparan las más utilizadas como son Dspace, Eprints, Fedora, CDS Invenio, entre otras. Son muchas las ventajas que este tipo de software ha aportado a los repositorios pero, sin lugar a dudas, la principal ha sido la adaptabilidad, ya que ha permitido realizar mejoras que han acercado los programas a los requisitos establecidos para cada uno de ellos. A parte de la adaptabilidad y de los costes, los repositorios también se han beneficiado de otras ventajas del software libre, como el escrutinio público, donde el código está sometido a un proceso de revisión que aporta dinamismo a la corrección de errores, o la independencia del proveedor, hecho que garantiza la continuidad teórica del software. Las comunidades internacionales que respaldan los programas *open source* en que se basan los repositorios son importantes, lo cual ya de por sí es una garantía de continuidad. Cabe remarcar que el coste de adquisición del software libre es inexistente pero no así el de desarrollo e implantación. Para adaptar estos programas a las necesidades establecidas de un repositorio es necesario invertir tiempo y recursos tanto materiales como de personal.

La UNESCO recientemente acaba de publicar un estudio sobre implementación de repositorios donde además se aborda el tema de la preservación de los mismos (UNESCO, 2007). Todos estas plataformas tienen en común que cumplen con el protocolo OAI-PMH. De las estadísticas sacadas del directorio OpenDOAR, las dos con mayor número de instalaciones son Eprints y Dspace.

Aunque la tendencia mundial es implementar los repositorios institucionales con software libre, existe software comercial adecuado para implementar repositorios, como *Digital Commons* de BePress, *Hyperion Digital Media Archive* de Sirsi/Dynix, *DigiTool* de Ex Libris, o *Symposia* de Innovative Interfaces, Inc.

Directorios de repositorios

Además de existir un repositorio debe poder encontrarse fácilmente, para ello los directorios o listas de proveedores de datos son de gran utilidad, por eso es importante que cuando se cree un repositorio se registre en algún directorio internacional para aumentar su visibilidad y facilitar su localización. Entre los directorios más relevantes y utilizados, aunque no los únicos, para la búsqueda de repositorios tanto institucionales como temáticos están: ROAR, Open Archives Initiative list y OpenDOAR.

Ejemplos de repositorios europeos

UK PubMed Central (<http://ukpmc.ac.uk/>)

Se trata de un repositorio creado a comienzos de 2007, especializado en Medicina y que sigue el modelo de PubMed Central (PMC), el archivo del NIH (National Institutes of Health) gestionado por la National Library of

Medicine (NLM) de los Estados Unidos. A semejanza de éste, en el Reino Unido está financiado por el Wellcome Trust y gestionado por la British Library, la Universidad de Manchester y el European Bioinformatics Institute (EBI). Su objetivo es preservar y facilitar el acceso sin restricciones a la literatura médica de la misma forma que la NLM ha venido haciendo con la literatura impresa.

DSpace@Cambridge (<http://www.dspace.cam.ac.uk/>)

En abril de 2006, tras una fase de proyecto, DSpace@Cambridge se convirtió en un servicio estratégico de la Universidad de Cambridge gestionado conjuntamente por la biblioteca y el servicio de informática. Su objetivo es preservar y difundir los materiales digitales, tanto de investigación como docentes, creados por personal de o relacionado por cualquier motivo con la Universidad. Entre sus contenidos se encuentran artículos, prepublicaciones, tesis, informes técnicos en diferentes formatos como PDF, multimedia, programas interactivos, bases de datos, etc. Hay que destacar que muchas entradas proceden del banco de datos WorldWideMolecularMatrix y que algunos documentos tienen acceso restringido. Los usuarios registrados pueden configurar un sistema de alertas para que se les comuniquen los nuevos materiales añadidos al depósito. El depósito se ha construido utilizando el software DSpace del MIT, uno de los sistemas más populares para depósitos institucionales que existen actualmente

Max Planck Society Edoc Server (<http://edoc.mpg.de/>)

eDoc es el depósito institucional de todos los institutos Max Planck de Alemania. Sus objetivos, como en otros casos de repositorios institucionales, son conservar la memoria científica de la institución, y explorar nuevos modelos de comunicación científica. Fue creado en 2003. El sistema está organizado por colecciones. Cada una de ellas corresponde a contenidos de un instituto concreto. eDoc pretende proporcionar acceso a toda la producción científica de los investigadores que trabajan para el Max Planck, independientemente de los tipos de publicación o formatos de ficheros. Cada instituto decide qué material (revisado por pares o no) quiere almacenar y disseminar a través del repositorio, de acuerdo a sus propias circunstancias y a las prácticas de comunicación que se siguen en su disciplina. Cada colección tiene una política definida que proporciona explicaciones sobre el alcance de la misma. Todos los institutos están invitados más que obligados a usar el servidor eDoc.

Ejemplos de repositorios españoles

TDX/TDR: Tesis Doctorals en Xarxa (<http://www.tdx.cbuc.es/>)

TDR (Tesis Doctorales en Red) es un repositorio cooperativo que contiene, en formato digital, las tesis doctorales leídas en las universidades de Cataluña y de otras comunidades autónomas. Participan actualmente 20 universidades con casi de 6.800 tesis. Es el repositorio con más tradición de nuestro país pues el proyecto nació a partir de un convenio firmado en 1999, aunque no entró en funcionamiento hasta 2001. Permite la consulta remota del texto completo de las tesis, así como realizar búsquedas por autor, director, título, tema de la tesis, universidad y departamento donde se ha leído, año de defensa, etc. Los objetivos de este repositorio, gestionado por el Consorci de Biblioteques Universitàries de Catalunya (CBUC) y el Centre de Supercomputació de Catalunya (CESCA), y patrocinado por la Generalitat de Catalunya, son: difundir, por todo el mundo y a través de Internet, los resultados de la investigación universitaria; ofrecer a los autores de las tesis una herramienta que incremente el acceso y la visibilidad de su trabajo; mejorar el control bibliográfico de las tesis; impulsar la edición electrónica y las bibliotecas digitales; incentivar la creación y el uso de la producción científica propia. Son las propias universidades que participan en TDR las que difunden el repositorio entre sus estudiantes de doctorado, les dan las recomendaciones y la lista de formatos electrónicos aceptados para la redacción de las tesis y (una vez presentadas y aprobadas) las editan y cargan en TDR. Además, a medio plazo, está previsto reconvertir a formato digital algunas tesis ya presentadas durante los últimos años en otros formatos (microfichas o papel). Sobre los aspectos legales, para cada una de las tesis se debe firmar un contrato de edición entre la universidad donde se ha leído y el autor, con el cual este autoriza a la universidad a publicar la tesis en el entorno de Internet y se reserva el derecho de publicarla en otros soportes y/o editoriales. Existe un modelo base de contrato que cada universidad ha personalizado. Actualmente se pueden consultar desde TDR no sólo las tesis de las universidades participantes, sino también todas aquellas tesis de otros repositorios españoles que las hacen accesibles con el protocolo OAI-PMH (actualmente: Dialnet, Universidad Carlos III de Madrid, Universidad Complutense de Madrid, Universidad de Alcalá, Universidad de Burgos, Universidad Politécnica de Madrid, Universidad Rey Juan Carlos, Universitat d'Alacant, Universitat Politècnica de València). Esta búsqueda recibe el nombre, en la interfaz de TDR, de "Búsqueda global". Para poner en marcha esta nueva funcionalidad, se ha utilizado el módulo MetaIndex y las librerías X-Server del software MetaLib de ExLibris, que permite importar los metadatos de los diferentes repositorios que operan el protocolo OAI-PMH. De esta manera, las tesis continúan estando almacenadas en el repositorio original, pero TDR permite consultarlas a través del metadato. Con esta nueva funcionalidad, TDR se ha convertido en el primer repositorio del Estado que, además de dar acceso a sus propias tesis, permite también buscar entre las tesis doctorales de otros repositorios del Estado que ofrecen metadatos OAI-PMH

E-PrintsUCM (<http://www.ucm.es/eprints/>)

Se trata del repositorio institucional de la Universidad Complutense de Madrid. Se trata de uno de los archivos institucionales más antiguos de nuestro país ya que data de 2003. Entre los objetivos del archivo están el incrementar el acceso y la difusión de la investigación desarrollada en la universidad, así como permitir una mayor visibilidad e impacto de la investigación publicada. Actualmente almacena casi 6.000 documentos de los que la mayor parte son tesis depositadas por la Biblioteca. El software sobre el que está implementado es e-prints, igual que en el caso anterior. En la evolución del número de documentos mostrada en el gráfico 5, vemos como ha habido un incremento bastante fuerte al comienzo del archivo con objeto de crear una masa crítica de trabajos y posteriormente se ha estabilizado.

Destacamos además otros repositorios españoles como: Dipòsit Digital de Documents de la UAB (28.397 ítems (2008-12-01)); E-Archivo, Universidad Carlos III de Madrid (2.567 ítems (2008-11-11)); Repositorio de la UNED (10.971 ítems (2008-06-13)); RUA, Repositorio Institucional de la Universidad de Alicante (5.122 ítems (2008-06-05)); RECERCAT, Dipòsit de la Recerca de Catalunya, CBUC (4.186 (2009-02-07)); RACO, Revistes Catalanes amb Accés Obert, CBUC (75.810 (2009-02-07)); UPCommons, Portal d'accés obert al coneixement de la UPC (13.973 (2009-02-07)); DDD, Dipòsit Digital de Documents de la UAB (31.263 (2009-02-07)); Dialnet, Difusión de Alertas en la Red, Universidad de La Rioja; RiuNet, Repositorio Institucional de la Universidad Politécnica de Valencia.

Roderic-UV: el repositorio institucional de la Universitat de València

La Universidad de Valencia ha creado un gran depósito digital similar al que tienen otras grandes universidades del mundo. El repositorio institucional recibe el nombre de Roderic-UV (Recursos Abiertos Digitales para la Enseñanza, la Búsqueda y la Cultura) en honor al papa Alexandre VI (Roderic de Borja) que otorgó la bula fundacional a la Universidad de Valencia en 1501.

La creación de este repositorio institucional, aprobado por el Consejo de Gobierno de la Universidad y presentado por el rector Francisco Tomás y el vicerrector de Investigación, Esteban Morcillo, el pasado mes de octubre, se inscribe en el movimiento Open Access. Inicialmente, Roderic nace con más de 4.000 documentos, incorporando los materiales digitales procedentes de dos grandes fuentes de la Universidad de Valencia. De un lado, una selección de más de 500 tesis doctorales y de otro, "Somni", una selección de documentos conservados en la Universidad que se encuentran en dominio público y cuyos derechos de reproducción está en manos de esta Institución, constituyendo una parte importante del patrimonio bibliográfico valenciano. En "Somni" se incluyen todos los incunables existentes en la Universidad, la colección completa de

carteles de la Guerra Civil española y una selección de obras impresas entre el siglo XVI y el XVIII. En todos los casos se ha digitalizado el documento completo.

La Roderic se encuentra adscrito al SBD de la Universitat de València. Atendido a su carácter transversal, colaboran en el desarrollo del proyecto los servicios (Informática, Servicios Jurídicos, Taller de Audiovisuales, Publicaciones, Gabinete de Prensa.) y las estructuras (Centro de Formación y Calidad) implicados directamente en la producción y la gestión de contenidos. Los objetos digitales incluidos en Roderic se almacenarán en un servidor del Servicio de Informática, con todas las medidas de seguridad para permitir una consulta libre, rápida y gratuita desde cualquier lugar del mundo y en cualquier momento y, a la vez, preservar la integridad de estos documentos de una forma permanente. La creación del repositorio digital OA Roderic (Recursos Oberts Digitals per a l'Ensenyament, la Recerca i la Cultura), en el marco del proyecto Open Course Ware, constituye una iniciativa prevista como prioridad estratégica en el plan sectorial para bibliotecas del Plan operativo 2009 que desarrolla el Plan Estratégico de la UV 2008-2011 (PEUV).

A pesar de la presentación pública, hay que reconocer que Roderic –UV no se encuentra en producción. Su creación ha partido de instancias docentes ajenas al SBD, y actualmente se encuentra en fase de definición de sus características y procesos de trabajo.

Retos y consejos

La comunidad internacional tienen tres importantes tareas pendientes en lo que se refiere a los repositorios institucionales:

- Incrementar el autoarchivo en repositorios institucionales: la principal dificultad en la construcción de un repositorio no es la puesta a punto del programa informático, sino establecer un marco legal que permita iniciar el proyecto y, sobre todo, convencer a la comunidad científica y universitaria de que poner documentos en acceso abierto y en un repositorio no solo es fácil, sino que tiene además muchas ventajas. La tarea más difícil es, sin lugar a dudas, establecer mecanismos e inercias de alimentación de los repositorios. La solución más obvia sería que las universidades hicieran obligatorio el autoarchivo de los documentos de investigación en los repositorios institucionales, pero a pesar de las muchas adhesiones a declaraciones como la de Berlín, la verdad es que las políticas de obligación son excepción. No abundan casos como el de la Universidad Complutense de Madrid, que el pasado verano se sumó al mandato de la Comunidad de Madrid para que los grupos de investigación publiquen en abierto sus resultados en su repositorio *E-PrintsUCM*. Ciertamente es también que la tarea de instituciones relevantes en la financiación de la investigación se están posicionando claramente a favor del OA y recomiendan encarecidamente de

introducir en repositorios de acceso abierto la investigación producida con fondos públicos

- Integrar los IR en los instrumentos bibliográficos ya existentes: durante décadas, los investigadores se han acostumbrado a descubrir los documentos que les podían ser de ayuda en su trabajo en dos instrumentos: bibliografías o bases de datos de resúmenes e índices y en los catálogos de las bibliotecas. Pero las posibilidades de descubrir materiales relevantes para una investigación se han incrementado enormemente en los últimos años: los buscadores en Internet (sean Google o Scopus), los ‘paquetes’ de revistas contratados conjuntamente de forma consorciada, las búsquedas federadas a múltiples instrumentos... En este nuevo contexto, ¿dónde situar a los repositorios institucionales? ¿Tenemos que incorporar los metadatos de los documentos de los repositorios en los catálogos actuales o hacer búsquedas federadas entre éstos y los repositorios? ¿La diferencia de granularidad, no conducirá en este caso a la desorientación del usuario?

- Encontrar un equilibrio entre el carácter institucional de los repositorios y su voluntad científica: finalmente, con los repositorios hemos podido constatar la contradicción que se produce entre los legítimos deseos de una universidad o institución de investigación de hacer aparecer los resultados de la investigación producida en su seno bajo su ‘marca’ y las igualmente legítimas necesidades de la comunidad investigadora de prescindir de búsquedas múltiples y de las fronteras institucionales para acceder a la globalidad de la producción de investigación clasificada en cualquier caso solo por criterios de calidad. Los repositorios nacidos como repositorios locales buscan, para ser eficaces, ser consultados conjuntamente, como parecen indicar las iniciativas italiana (PLEIADI), holandesa (DAREnet), australiana (ARROW) o el proyecto europeo DEEP que tiene como objetivo recoger los metadatos de tesis doctorales europeas.

Cualquier iniciativa de repositorio institucional, además, debería reflexionar sobre las siguientes cuestiones:

- La necesidad de fijar especificaciones sobre cómo dar la información de los diferentes elementos descriptivos de Dublin Core, y en especial de los puntos de acceso de los autores y materias, que se deben establecer tomando como referencia fuentes autorizadas como ficheros de autoridad y listas de encabezamientos de materia coherentes con los usados con otros instrumentos en la institución, como el catálogo. Obviar esta cuestión no sólo mermaría la calidad del repositorio, sino que además dificultaría posibles futuras integraciones o proyectos de interoperatividad.

- La necesidad de llevar a cabo acciones básicas de preservación, tanto técnicas (hacer copias de seguridad, comprobación de la integridad de los

documentos (*checksums*), URL permanentes, etc.) como organizativas. Uno de los objetivos de los repositorios institucionales es garantizar el acceso permanente a los documentos, es decir preservarlos a través del tiempo independientemente de su formato, tamaño, etc.

- La necesidad de afrontar estas iniciativas desde una perspectiva conjunta y cooperativa para poder alcanzar mejores resultados, con mayor facilidad, más visibles y con más valor añadido

BIBLIOGRAFÍA

<http://www.rebiun.org/eventos/workshops/workshopsanteriores.html>

Ricardo Eíto-Brun. “La gestión de recursos electrónicos en bibliotecas: la oferta de Innovative”. En: *El profesional de la información*, v.17, n. 3, mayo-junio 2008, p. 347-358

Elisa García-Morales. *Acceso abierto y repositorios de documentos*. Unidad de Autoformación, SEDIC 2002

http://www.sedic.es/autoformacion/acceso_abierto/programa.html

Melero, Remedios. “El paisaje de los repositorios institucionales open access en España”. En: *BiD: textos universitaris de biblioteconomia i documentació*, juny, núm. 20, 2008.

http://www2.ub.edu/bid/consulta_articulos.php?fichero=20meler4.htm