

## Del catálogo en línea de acceso público al archivo: búsqueda integrada y bibliotecas digitales con código abierto

*Traducción al español del documento original: From OPAC to Archive: integrated discovery and digital libraries with open source"*

*Traducido por: Federación Española de Sociedades de Archivística, Biblioteconomía, Documentación y Museística (FESABID), Madrid, España.*

**Edmund Balnaves**

Prosentient Systems

E-mail: ebalnaves@prosentient.com.au



Esto es una traducción al español de "*From OPAC to Archive: integrated discovery and digital libraries with open source*" Copyright © 2013 por *Federación Española de Sociedades de Archivística, Biblioteconomía, Documentación y Museística (FESABID)*. Este trabajo está disponible en los términos de la licencia Creative Commons Attribution 3.0 Unported License: <http://creativecommons.org/licenses/by/3.0/>

---

### Resumen:

*Las instituciones han descuidado tradicionalmente el potencial intelectual de sus investigaciones y de su producción editorial. El modelo tradicional de las publicaciones académicas descuida la importancia de que las instituciones contribuyan a la creación y la innovación de la producción intelectual. El movimiento a favor del código abierto representa una alternativa a la visión monolítica de los recursos web **de información** ofrecidos por Google, Amazon y otros gigantes de los contenidos. Desde el cambio de siglo, la innovación en los sistemas de código abierto ha permitido a las bibliotecas ampliar su gestión de la información a las bibliotecas digitales e integrar estos recursos con otras herramientas de gestión y **búsqueda** de recursos. Esto representa una oportunidad para que las instituciones hagan mejor uso del conocimiento inherente **a las publicaciones: deberían consolidar estos recursos en un repositorio institucional gestionado a través de una estructura de metadatos** eficaz y adecuada para **la búsqueda** a largo plazo de contenido.*

*Este trabajo propone un modelo para el desarrollo de recursos electrónicos institucionales unificados, distribuidos a través de modelos de **acceso y de código abiertos**, mediante una combinación de herramientas de búsqueda que usan el OPAC y el potencial del repositorio de las bibliotecas digitales.*

*Se proporcionarán ejemplos de integración entre sistemas de código abierto aplicados en los países tanto desarrollados como en vías de desarrollo.*

*Este enfoque proporciona un modelo sostenible y de bajo coste para implementar los repositorios institucionales. Las bibliotecas —que tienen un conocimiento sólido de la importancia de las normas de los metadatos y servicios— deberían estar en una posición adecuada para gestionar estos recursos, siempre y cuando reorienten sus habilidades para comprometerse activamente en el campo del código abierto. Esta es una visión descentralizada en la que se pueden congregarse de manera significativa tanto colecciones **muy pequeñas** como **muy grandes**. **Es además un modelo / se trata además de una propuesta en la que / el que la biblioteca y el bibliotecario son mediadores de conjuntos de datos internos, colecciones digitales y físicas y recursos externos, todo ello en beneficio de su organización en la gestión eficaz del conocimiento, la educación y el ocio.***

**Palabras clave:** Código abierto, acceso abierto, sistemas de integración

---

## **1 INTRODUCCIÓN**

Las bibliotecas están cambiando: su gestión ha dejado de ser física y ha pasado a ser digital; se han convertido en centros de información, así como en gestoras de colecciones físicas, en riesgo de quedar marginadas en la organización. Sin embargo, como gestoras eficaces de recursos complejos, tanto impresos como electrónicos, las bibliotecas pueden ayudar a las instituciones a hacer un uso más eficaz del descuidado potencial intelectual de sus investigaciones y producción editorial. El modelo tradicional de las publicaciones académicas desatiende la importancia de que las instituciones contribuyan a la creación e innovación de la producción intelectual. El movimiento a favor del código abierto representa una alternativa a la visión monolítica de los recursos web de información ofrecidos por Google, Amazon y otros gigantes de los contenidos. Junto con el acceso abierto, el movimiento de código abierto de las bibliotecas ha prosperado desde el cambio de siglo. La innovación en los sistemas de código abierto ha facilitado que las bibliotecas amplíen su gestión de la información al ámbito digital e integren estos recursos con otras herramientas de gestión y búsqueda de recursos. Esto representa una oportunidad para que las instituciones hagan mejor

uso del conocimiento inherente a las publicaciones: deberían consolidar estos recursos en un repositorio institucional gestionado a través de una estructura de metadatos eficaz.

Muchas instituciones tienen ahora un deseo continuo de capitalizar su producción editorial y de hacer mejor uso de su investigación e información, y ~~este~~ es preciso ~~para~~ que las bibliotecas estén a la altura y se involucren en este campo. Para algunas bibliotecas, esto supone convertirse en una especie de editorial y una transición importante para el bibliotecario como "conservador" de la información. Las herramientas para apoyar esta transición ya están disponibles y accesibles para organizaciones pequeñas y grandes. Los sitios web de organizaciones suelen ser de corta duración y están sujetos a grandes cambios, mientras que los sistemas bibliotecarios tienen un sólido historial en la gestión de contenido a largo plazo y son eficaces como repositorios documentales y de metadatos.

Tal es la evolución de estos sistemas, con un historial de desarrollo de más de 15 años, que no sólo son sólidos y estables, sino también cada vez más capaces de integración. Esto ofrece la posibilidad de generar una nueva integración entre los sistemas de código abierto y entre estos sistemas y los recursos externos basados en la web.

## 2 UN MODELO DE CÓDIGO ABIERTO

Uno de los grandes binomios que sustentan las bibliotecas modernas ha sido el auge de los sistemas de código abierto (OSS en su sigla inglesa) y del acceso abierto (OA en su sigla inglesa) (Keast and Balnaves, 2009). Los sistemas de código abierto para bibliotecas han prosperado en los últimos 15 años, especialmente en el ámbito de las bibliotecas digitales y del software del sistema de gestión de bibliotecas. Aunque ya en 2008 el estudio del JISC puso de manifiesto que "el sistema de gestión de bibliotecas (LMS en su sigla inglesa) no es viable para la mayoría de las instituciones" (Adamson et al., 2008), este hallazgo ha sido desmentido por la extraordinaria adopción del código abierto, no sólo en el espacio del LMS, sino también para las bibliotecas digitales y, más recientemente, para las herramientas de búsqueda. El código abierto puede facilitar la adopción del acceso abierto en una institución. Del mismo modo, un proyecto exitoso de acceso abierto puede justificar la mejora continua de la implementación de sistemas de código abierto. En segundo lugar, los sistemas de código abierto pueden proporcionar certeza a una institución en los costes de funcionamiento cuando se combina con el código abierto. Cuanto más grande es la comunidad de los usuarios de código abierto, más sólido es el apoyo general. En tercer lugar, los sistemas de código

abierto pueden proporcionar un cierto nivel de seguridad, al no haber bloqueo propietario y ser el código visible (y por lo tanto poderse corregir). La profundidad funcional de esta seguridad se puede mejorar mediante el trabajo de aquellos que adoptan el modelo de código abierto.

Una confusión común es pensar que el código abierto significa "gratis". Si bien puede tener un coste bajo, ninguna operación del sistema de la tecnología de la información es gratis. La ampliación continua de un sistema, las actualizaciones de software a lo largo del tiempo, el apoyo a la personalización y mejoras, la administración del servidor y los costes de la red son sólo algunos de los elementos básicos de la gestión de un sistema de información. No obstante, la amortización del soporte de software, a través de una amplia base instalada, da lugar a un modelo rentable para instituciones más pequeñas. Los departamentos de tecnología de la información (TI) son normalmente la principal barrera para la aceptación de un modelo de código abierto y son reacios a implementar sistemas de código abierto. Esta barrera se ha superado en muchos casos con la aparición de un gran número de proveedores de alojamiento que proporcionan una solución de código abierto enpaquetada. Esto ha estimulado por sí mismo una innovación considerable en el campo del código abierto, ya que ha mejorado los elementos comerciales de la comunidad.

Los proyectos de código abierto, como Koha, han tenido cientos de colaboradores por todo el mundo que han mejorado el proyecto de manera continua.

El aumento del acceso a los recursos digitales implica una creciente complejidad en la gestión de los recursos de la biblioteca, incluidas las bibliotecas o colecciones digitales, fuentes de noticias digitales o la digitalización de los recursos de la organización. Esta complejidad adicional puede conducir a requerir una opción de búsqueda federada (la integración en un único portal de los principales recursos de información disponibles para los clientes de la biblioteca), y sistemas de gestión de distribución del trabajo (para gestionar los procesos complejos en el desarrollo de la colección electrónica) y a un único inicio de sesión (para ocultar la complejidad del acceso a los múltiples recursos de las bases de datos subyacentes). La rapidez del desarrollo tecnológico conlleva dificultades a largo plazo en la gestión de la producción intelectual y creativa en formato digital. Las bibliotecas y los museos tienen un papel clave en la preservación de los esfuerzos analíticos y creativos a largo plazo. Sin embargo, la mayoría de las bibliotecas no están bien equipadas para llevar a cabo

investigaciones sobre la preservación de artefactos y creaciones de los nuevos medios. La creación de la colección de las bibliotecas digitales se ha asociado con los riesgos inherentes a la obsolescencia tecnológica. Además de los riesgos sistemáticos asociados a la arquitectura esencial de las tecnologías de la información, están los problemas del software y de obsolescencia del hardware. Las cuestiones de obsolescencia no son obstáculos inherentes a la transición a la gestión de recursos electrónicos, pero son cuestiones que deben ser abordadas por la institución en la gestión de los recursos dispares que constituyen una colección electrónica. Los sistemas de la información sufren inevitablemente, con el tiempo, una serie de transformaciones, como ocurre con los objetos digitales almacenados en un sistema de información, y una comunidad de código abierto ofrece cierta diversidad en la innovación del sistema y su potencial para gestionar los recursos a largo plazo.

Este modelo se puede ilustrar a través de la combinación de tres sistemas de código abierto. No hay una solución única de "código abierto" para todas las necesidades de la biblioteca. Sin embargo, hay formas en las que las combinaciones de código abierto se pueden implementar para lograr resultados altamente funcionales. Por ejemplo, hemos combinado los siguientes sistemas con eficacia para proporcionar soluciones valiosas para la biblioteca.

- VuFind: una herramienta de búsqueda que puede recoger información a través de múltiples fuentes. Con Apache SOLR, extensible a grandes conjuntos de búsqueda. VuFind es relativamente nuevo, pero proporciona una buena búsqueda para el usuario y puede recoger e importar datos de una cierta variedad de fuentes.
- Koha: sistema de gestión de bibliotecas, que puede realizar varias búsquedas a través de varios millones de objetos y eficaz como sistema de gestión de tamaño mediano y de búsqueda federada.
- DSpace: un sistema para bibliotecas digitales, modificable que puede apoyar colecciones muy grandes, pero diseñado como sistema de un solo propósito.

De entre estos, Koha tiene la comunidad de apoyo más diversa, con cientos de desarrolladores y colaboradores de todo el mundo. DSpace tiene un fuerte respaldo de las universidades para el desarrollo institucional y un proceso de recuperación de información más lento.

El siguiente es un análisis de funciones de las formas en que estos sistemas de código abierto se pueden implementar:

	Bibliotecas digitales	Gestión de bibliotecas	Búsqueda federada	Inicio único de sesión
Koha		Koha	Koha se puede utilizar para agregar diferentes códigos que proporcionen un único punto	Koha puede ser a la vez el objetivo y el código para un único inicio de sesión.
DSpace	Sistema de biblioteca digital extensible que apoya una función tanto para la gestión de bibliotecas digitales como para la gestión de registros.			DSpace puede ser el objetivo de inicio de sesión único.
VUFind			VUFind, a diferencia de Koha, se diseña desde cero como herramienta para búsqueda de conjuntos	

### 3 INTEROPERABILIDAD Y METADATOS

Los metadatos son la información que describe los objetos en las bibliotecas digitales. El título del artículo, autor, dimensiones y el formato son ejemplos de metadatos. Estos cumplen tres objetivos en las bibliotecas digitales:

- Los metadatos descriptivos: al igual que en la catalogación tradicional, necesitan describir e identificar los objetos digitales para que puedan ser localizados dentro de las bibliotecas digitales. Las normas de metadatos de las bibliotecas digitales para describir objetos desempeñan la misma función que las normas AACR2 y MARC en los catálogos tradicionales. Ejemplos de normas de metadatos descriptivos de uso común en las bibliotecas digitales son la Iniciativa de Metadatos Dublín Core (IMDC), Metadata Object Description Schema (MODS), and Metadata Encoding & Transmission Standard The (METS). Aunque el IMDC es probablemente el más utilizado por las bibliotecas digitales, las normas MODS y METS proporcionan un marco descriptivo más completo como sucesores del MARC. DSpace y Greenstone utilizan el IMDC

como estructura de metadatos descriptivos.

- Metadatos semánticos: los metadatos semánticos proporcionan la clasificación por temas y la información de relación de los objetos de las bibliotecas digitales. Si bien esto puede estar basado en el binomio tradicional de identificadores nombre/valor (tema='Historia parlamentaria'), la tendencia actual es el cambio a la Estructura de Descripción de Recursos. La Estructura de Descripción de Recursos sustenta muchos proyectos, que se están dando cuenta de las posibilidades de la Web Semántica para una descripción de metadatos más sólida de los documentos en la web (y en los archivos). La descripción semántica de metadatos va más allá del binomio nombre/valor y describe los metadatos mediante una serie de "enunciados", enunciados de sujeto, objeto y predicado (el título del libro es 'La historia de los Parlamentos'). Para el concepto de Estructura de Descripción de Recursos, resulta fundamental la posibilidad de unificar conceptos a través de muchos recursos de una manera significativa. Fedora Commons implementa la Estructura de Descripción de Recursos como esquema subyacente.
- Recogida de metadatos. Hay muchos sistemas de bibliotecas digitales: comerciales, de código abierto y hechos a medida (de cosecha propia). Independientemente del enfoque interno de los metadatos para la descripción y clasificación temática de los objetos de la biblioteca, el apoyo de una norma de recogida de metadatos proporciona un medio para la interoperabilidad entre los sistemas de bibliotecas digitales. El sistema de recogida más implementado es el Open Archives Initiative Protocol for Metadata Harvesting (OAI/PMH). Este esquema soporta los metadatos "recogidos" entre las bibliotecas digitales, permitiendo la búsqueda de recursos digitales entre sistemas. Kete utiliza el OAI/PMH para su esquema interno. DSpace, Greenstone, Fedora Commons y Kete admiten una interfaz de recogida de metadatos OAI/PMH.
- A largo plazo, la interoperabilidad de unos recursos con otros recursos digitales, desarrollados en un país y una zona del mismo, se verá reforzada u obstruida por el nivel y la calidad de los metadatos recogidos y vinculados mediante recursos digitales. La selección de una estructura de trabajo con metadatos debe realizarse en referencia a los proyectos existentes a nivel nacional y local.

#### **4 ESTUDIO DE UN CASO REGIONAL**

Las bibliotecas están efectuando una transición progresiva de los recursos físicos a los electrónicos. El sistema de bibliotecas digitales facilita:

- el acceso local y público a la producción intelectual de la organización.
- el acceso local a los recursos creados en soporte digital y a los digitalizados que se almacenan localmente en la biblioteca digital.

El Parlamento Nacional de las Islas Salomón ilustra precisamente esta metodología. En marzo de 2012 pusieron en marcha una iniciativa combinada de código abierto –software de biblioteca digital de código abierto (Dspace) y software de gestión de bibliotecas (Koha)— para crear un servicio electrónico de biblioteca que proporcionase:

- Acceso a legislación, informes de comités, gacetas y registros de debates Hansard.
- Acceso a los recortes y comunicados de prensa de los parlamentarios.

La Biblioteca del Parlamento necesitará gestionar una gama cada vez más amplia de recursos electrónicos, que incluirá:

- Servicios en línea de bases de datos electrónicas: provisión externa con servicios de búsqueda y, posiblemente, de texto completo.
- Bibliotecas digitales: con la posibilidad de gestionar de documentos digitales creados por la biblioteca (ya sean originados en soporte digital o digitalizados).
- Servicios de sindicación: servidores de información que proporcionan recursos actuales basados en preferencias de información (por ejemplo, fuentes de noticias).
- Suscripciones electrónicas: libros electrónicos y revistas electrónicas.

Una estructura construida alrededor de soluciones de código abierto proporciona un marco viable económicamente para acceder a los recursos de información complejos.

Se puede implementar en varios modelos en función de los recursos informáticos de los que disponga la organización:

- a través del alojamiento externalizado y proveedores de alojamiento de código abierto—en la actualidad están surgiendo muchos—.
- a través de TI cooperativa o almacenamiento compartido con otros departamentos u organizaciones para compartir los costes de mantenimiento de la TI

\*A través del alojamiento y gestión directos de la organización. Si la biblioteca tiene un área TI de apoyo, puede ser una manera eficaz de construir un conocimiento interno y de adquirir la propiedad completa de los sistemas. A veces, todavía es posible aprovechar la experiencia



de los proveedores de alojamiento para proporcionar apoyo técnico a los sistemas alojados internamente.

## **5 CONCLUSIÓN**

Los sistemas de código abierto representan una nueva oportunidad para que las bibliotecas participen directamente en la tecnología que proporcionan y para que ofrezcan soluciones valiosas. Esta nueva generación de sistemas de código abierto proporciona capas de servicios abiertos para apoyar a una comunidad de clientes de las bibliotecas cada vez más digital y móvil. Bibliotecas como la del Parlamento Nacional de las Islas Salomón, pueden utilizar esta tecnología para ir por delante de una generación de sistemas y proporcionar nuevos servicios de acceso a través de recursos digitales de manera práctica y rentable.

## **6 REFERENCIAS**

Adamson, Veronica, Paul Bacsich, Ken Chad, and Jane Plenderleith. 2008. "CCSI & SBNU Estudio de los Sistemas de Gestión de Bibliotecas".

[Http://www.jisc.ac.uk/publications/generalpublications/2008/librarymanagementflyer.aspx](http://www.jisc.ac.uk/publications/generalpublications/2008/librarymanagementflyer.aspx)(Abril).

Keast, D., and E. Balnaves. 2009. "Open Source Systems Bring Web 2.0 to Special Libraries". En el Congreso Internacional de Bibliotecas Médicas. Brisbane, 2009. <http://espace.library.uq.edu.au/view/UQ:179870>.

Edmund Balnaves es el Oficial de Información de la Sección de Tecnología de la Información de la IFLA, es director de ALIA y, sin ningún reparo, se aventura con la tecnología de información de la biblioteca en compañía de Prosentient Systems.