

## **Bibliotecna. La Biblioteca Digital de la UPC. Drupal como gestión de contenidos 2.0.**

Jordi Prats ([jordi.prats@upc.edu](mailto:jordi.prats@upc.edu)), Toni Prieto ([antonio.juan.prieto@upc.edu](mailto:antonio.juan.prieto@upc.edu))  
Oriol Rico ([oriol.rico@upc.edu](mailto:oriol.rico@upc.edu)), Jordi Serrano ([j.serrano@upc.edu](mailto:j.serrano@upc.edu))

Servei de Biblioteques i Documentació.  
Universitat Politècnica de Catalunya

En 2004 Dale Dougherty de O'Reilly Media acuñaba el término web 2.0 para hacerse eco de la revolución que estaba empezando en esos momentos en la red. En general, la red, deja de ser un mero aparador de contenidos para hacer partícipe de ellos a los usuarios y hacerlos dueños de la información que depositan.

Para conseguir ese objetivo se debe dotar al site de herramientas que faciliten la extracción e introducción de información de forma sencilla y clara, respetando para ello todos los estándares existentes, de forma que el usuario sea capaz de poder realizar todas las acciones únicamente con la ayuda de un navegador.

Es en este nuevo contexto en el que los servicios y recursos ofrecidos por las bibliotecas deben evolucionar hacia nuevos entornos y abrirlo definitivamente al usuario.

Es aquí donde el Servei de Biblioteques i Documentació de la Universidad Politècnica de Catalunya comienza un profundo cambio en todos sus sistemas de información. Un antiguo VTLS se verá renovado por un nuevo y moderno catálogo procedente de la empresa Innovative y denominado Millennium, así como su biblioteca digital, que realiza una reorientación del producto centrándolo en servicios hacia el usuario.

Pero llevar a término una tarea como la reorientación de la biblioteca digital, en la que se entrelazan diferentes elementos como el catálogo, recursos electrónicos, depósitos institucionales, servicios, etc. hace necesario disponer de una herramienta que sirva para englobar, interconexionar y ofrecer servicios transversales a todos ellos, servicios, por otro lado, que facilitaran convertir nuestro web en un entorno 2.0.

Esta tarea se debe dejar en manos de un gestor de contenidos. No es objetivo de este artículo repasar y explicar todos y cada uno de los existentes en el mercado. No obstante si remarcar que entre los tres destacados bajo el paradigma del software libre, Joomla, Typo y Drupal, es este último el que mejor se adapta a entornos bibliotecarios.

Las principales características que hacen a Drupal ideal para ese cometido son su facilidad de lenguaje de implementación, casi sin tener conocimientos técnicos se puede implantar y mantener un sitio web, disponer de una comunidad de desarrollo muy potente, que facilita disponer el sitio constantemente actualizado, y proveer diferentes funcionalidades que se obtienen de los módulos que se publican en el sitio oficial de Drupal.

Disponer de múltiples taxonomías<sup>1</sup> para categorizar contenidos, posibilidad de

---

<sup>1</sup> Se considera taxonomía como el sistema de clasificación que agrupa a toda la diversidad de elementos bajo términos que los caracterizan.

ofrecer el site en varios idiomas, así como sistemas para añadir comentarios son algunas de las características nativas a Drupal, cosa que de entrada facilita y mucho el hecho de montar los tan deseados servicios 2.0.

La tarea más tediosa, a priori, puede resultar la importación y explotación de los recursos que se encuentran fuera de la plataforma, como pueden ser el catálogo o los depósitos institucionales, cada uno de ellos con sus propios protocolos de comunicación.

Es en este punto en el que el implementador debe plantearse si debe catalogar aquellos recursos que se encuentran en un depósito institucional e interconectar únicamente el catálogo con Drupal o bien dejar a manos del gestor de contenidos la implementación de los protocolos de importación de los metadatos de ambas fuentes.

Ambas opciones tienen sus ventajas e inconvenientes, que intentaremos repasar a continuación.

La primera de ellas, catalogar absolutamente todo, impide tener cierto control sobre que partes del catálogo se quiere mostrar en el sitio web, dado que el protocolo Xrecord<sup>2</sup> no permite filtraje de contenidos en base a campos del registro, por no hablar de los costes de catalogación directamente proporcionales a la cantidad de registros encontrados en los depósitos institucionales. Remarca también que la disponibilidad de un cierto elemento se verá influenciada por el tiempo necesario para su catalogación, y la imposibilidad de mantener cierto control sobre elementos constantemente en cambio, como por ejemplo, una página web.

No obstante esta opción puede resultar interesante si se quiere disponer de una única fuente de datos, aunque quizá en ese caso Drupal pueda suponer una herramienta demasiado grande para tal acometido.

La segunda opción, implementar los motores de interconexión en el propio gestor de contenidos, puede resultar más laboriosa inicialmente, pero a la larga aportará una serie de beneficios. Mayor control sobre los contenidos importados e inmediatez en el sitio web una vez publicado un determinado recurso en el depósito son solo algunas de ellas. No hay que perder de vista que las taxonomías deben ser las mismas que en el catálogo, requerimiento no demasiado difícil de cumplir si se tiene en cuenta que en la mayoría de sitios los metadatos son validados contra el mismo catálogo y revisados por el propio bibliotecario.

Finalmente el Servei de Biblioteques i Documentació opta por hacer una estrategia mixta. La segunda opción ofrecerá mayor control sobre los contenidos importados, mientras que una catalogación de aquellos contenidos de los depósitos que ya disponen de su versión en papel se presenta como una opción razonable y a bajo coste, entre otras cosas porque es algo que ya se realiza antes de la planificación de este nuevo proyecto. En la figura 1 se puede observar cual será la estrategia general a seguir.

Tal y como se apuntaba al principio del artículo, Drupal dispone de una potente comunidad de desarrolladores, que junto con su facilidad de implementación, hace que sea cada vez más grande, y que como se va a ver, va a facilitar y mucho la implementación final.

---

<sup>2</sup> Xrecord es el servicio web que ofrece el catálogo Millennium para poder consultar y extraer datos de un cierto registro del catálogo.

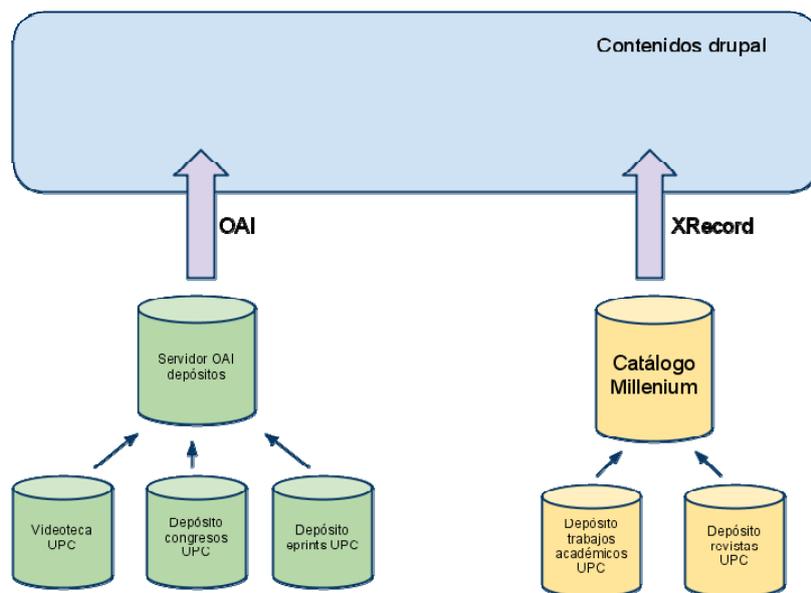


Figura 1. Los recursos externos se importaran como contenidos a Drupal

El primer motor de importación resulta sencillo de implementar, para ello hay que fijarse en un módulo denominado Millennium(<http://drupal.org/project/millennium>). Entre sus características principales destacan mostrar en tiempo real la disponibilidad de ese elemento en la biblioteca, mapear los diferentes campos del registro MARC a taxonomías de Drupal creadas previamente, mostrar la cubierta del libro configurando un servicio de imágenes de carátulas tipo Amazon, la posibilidad de ver una previa del libro si este se encuentra en Google Books, así como poder reservar el libro a través de una pasarela con el catálogo.

También permite tener un control sobre que forma se deben importar los registros de Millennium a Drupal, si de forma manual o automática, siendo esta última la mas interesante, ya que el sistema es capaz de detectar cuando ha llegado al final de todos los elementos a importar y empezar de nuevo la reimportación para mantener sincronizados todos los datos que provienen del catálogo.

Para el segundo motor de importación existen diferentes proyectos en marcha en la comunidad Drupal, pero ninguno de ellos tiene aun publicado el código de manera oficial, o bien aquellos que lo tienen publicado no se acaban de adaptar a lo que se necesita. Algunos de estos proyectos son Extensible Catalog (<http://www.extensiblecatalog.org/>), OAI-PMH Module (<http://drupal.org/project/oai2>) o Drupal OAI-PMH ([http://drupal.org/project/oai\\_pmh](http://drupal.org/project/oai_pmh)).

El primero de estos proyectos tiene publicada una versión preliminar del código, no obstante, la aproximación a la solución difiere un poco de lo que se necesita, siendo el apartado en el que se mapean los metadatos a las taxonomías aquello que mas se encuentra a faltar. El segundo tiene desarrollado un servidor de metadatos OAI, pero no así un recolector. Es el tercero de ellos el que se aproxima a lo que se necesita, aun así el código ni tan siquiera ha llegado a publicarse de forma oficial.

A pesar de ello el Servei de Biblioteques i Documentació ha desarrollado una serie de funcionalidades sobre este código en fase embrionaria para dotarlo de funcionalidades parecidas a las del módulo de Millennium y próximamente se liberará en la comunidad Drupal. Entre sus características principales destacan la de mapear los diferentes campos

del registro OAI a taxonomías de Drupal creadas previamente, mostrar la cubierta del libro configurando un servicio de imágenes de carátulas tipo Amazon o la posibilidad de seleccionar que colecciones deben importarse a Drupal.

Llegado a este punto se dispone ya de todos los elementos necesarios dentro de Drupal para empezar a mostrar y/o agrupar los contenidos según las necesidades. En ese sentido la comunidad Drupal vuelve a sorprender poniendo a disposición un módulo que ofrece una serie de interfaces para realizar una navegación de contenidos dirigida en base a las taxonomías. Esto es, ofrecer filtros de entrada, de forma que el usuario vaya acotando cada vez más los elementos que pueden resultar de su interés. Seleccionar Audiovisuales de Calatrava en inglés ahora es relativamente sencillo gracias al módulo Faceted Search ([http://drupal.org/project/faceted\\_search](http://drupal.org/project/faceted_search))

Esta aplicación en si mismo no realiza nada, pero en combinación con otros módulos supone una herramienta muy potente para ayudar al usuario a acotar los contenidos. No obstante existen otras aproximaciones al mismo problema con el añadido de disponer de un buscador integrado. Una de ellas es el módulo Apache Solr Search Integration (<http://drupal.org/project/apachesolr>), que es finalmente el escogido para implementarse.

El núcleo de este módulo es el buscador Lucene, ampliamente usado en la comunidad DSpace. Entre otras características el módulo permite acotar los resultados, de forma que después de una primera búsqueda, el usuario es capaz de refinarla en base a los términos de las taxonomías que mas aparecen en los resultados de la búsqueda. En la figura 2 podemos observar una pantalla ejemplo con resultados de búsqueda.

The screenshot shows the search results page for 'calatrava' on the Bibliotècnica UPC website. The page features a search bar at the top with the query 'calatrava' and a 'Search' button. A 'Current search' box indicates 'Search found 53 items'. On the left, there are several filter panels: 'Filter by Tipus de continguts' (Book: 44, Manuscript: 4, Projected medium: 3, Audiovisual: 1, Graphic: 1), 'Filter by Tags' (Arquitectes: 33, Calatrava, Santiago: 29, Espanya: 29, Arquitectura moderna: 7, Art: 4, Italia: 5, PFC: 5, Ponts: 5, S. XX: 5, Art del Renaixement: 3), 'Filter by Idiomes' (Spanish: 31, English: 11, Catalan: 3, German: 3, Italian: 3, French: 1, Multiple languages: 1), and 'Filter by Autors' (Calatrava, Santiago: 16, Calatrava Escobar, Juan: 11, Frampton, Kenneth: 4, Rosselli, Paolo: 5, Nolla, Quim: 3, Tischhauser, Anthony: 3, Escola Técnica Superior d'Arquitectura del Vallès: 2, Nicolin, Pierluigi: 2, Polano, Sergio: 2, Tzonis, Alexander: 2). The main content area displays a list of search results, including 'Santiago Calatrava : obras y proyectos 1980-1985', 'Arquitectura de Santiago Calatrava Valls : 1980-1985', 'Santiago Calatrava : sculpture and drawing = esculturas y dibujos : 31 de mayo - 26 de agosto', 'Escale Satolas : Rhône-Alpes à Grande Vitesse', 'Santiago Calatrava', 'Santiago Calatrava : secret sketchbook', 'De la magnificencia y arquitectura de los romanos y otros escritos', and 'Santiago Calatrava : 1983-93 : catálogo de la exposición antológica en la Lonja de Valencia del 31 de mayo al 30 de junio de 1993'. Each result includes the author, title, year, and a brief description.

Figura 2. Resultados de la búsqueda realizada mediante el término "calatrava"

Otra característica que resulta interesante es la posibilidad de dar mas peso a unos tipos de contenidos sobre otros, excluir ciertos elementos de la búsqueda o la recomendación de otros elementos parecidos al que se está visualizando. De esta forma no solo se consigue que el usuario sea capaz de encontrar información relacionada con la temática buscada, si no que es capaz de filtrarla. En la figura 3 se puede ver la recomendación que realiza el sistema cuando se está consultando un resultado de la búsqueda.

The screenshot shows a library website interface. At the top left is the logo for 'Bibliotècnica UPC' with the tagline 'Catàleg de les Biblioteques de la UPC'. Below the logo is a search bar with the text 'Search this site:' and a search button. To the right of the search bar is a navigation menu with 'Frequently Asked Questions' and a 'User Login' section with a 'Login' button. The main content area displays a search result for 'Santiago Calatrava : obras y proyectos 1980-1985'. The result includes the author 'by Calatrava, Santiago' and the type 'Audiovisual'. A table lists the following details: Title: Santiago Calatrava : obras y proyectos 1980-1985; Type: Audiovisual; Author(s): Calatrava, Santiago; Language(s): Spanish; Description: Conferència amb diapositives a l'ETSAB, 17 de desembre de 1985.; Related links: http://hdl.handle.net/2099.2/467. Below the table, it says 'Submitted by Anonymous on Thu, 05/21/2009 - 19:27' and lists tags: Arquitectes, Arquitectura, Audiovisual, Calatrava, Santiago, Calatrava, Santiago. At the bottom left is a 'Drupal' logo. On the right side, there is a 'More like this' recommendation block with a list of related items: Santiago Calatrava obras y proyectos 1980-1985; Arquitectura de Santiago Calatrava Valls : 1980-1985; Arquitectura obra publica a Catalunya = Obra publica en Catalunya : 1980-1985; Intervencions en el patrimoni : 1980-1985; Arata Isozaki in Japan part two: 1980-1985.

Figura 3. El bloque "More like this" nos ofrece elementos interesantes parecidos al que se está visualizando

Alguien podría pensar que un punto crítico del site debería ser su rendimiento, teniendo en cuenta que se está cargando mas de 350000 elementos dentro de Drupal y que existirán taxonomías con mas de 400000 elementos. En esa dirección la comunidad Drupal nos brinda un módulo para la gestión de enormes taxonomías. Se trata de LeftandRight - Nested Set Taxonomy (<http://drupal.org/project/leftandright>). Es capaz de trabajar con taxonomias de más de 2000000 de elementos sin que para ello se penalice el rendimiento del sistema.

A partir de este punto se trata de buscar módulos dentro de la comunidad que sirvan para añadir funcionalidades y explotar los contenidos del site. A título de ejemplo destacar Content Recommendation Engine (<http://drupal.org/project/cre>) que sirve para recomendar elementos de Drupal al más puro estilo de Amazon.

Otros módulos, como User-to-user recommendation (<http://drupal.org/project/uurec>), utilizan filtrados colaborativos para recomendar elementos que pueden ser de interés para el usuario. El módulo Affinity pondrá sobre aviso a los usuarios que tienen intereses parecidos (<http://drupal.org/project/affinity>) en base a los elementos que se están consultando.

Existen otros módulos que sin ser herramientas para explotar los contenidos ayudaran en la gestión de la biblioteca. Como ejemplo cabe destacar Reservation module (<http://drupal.org/project/reservation>) que permite ofrecer un sistema de reservas dentro del site. Prestamos de portátiles, reserva de sala de estudio son sólo unos ejemplos.

En definitiva Drupal pone al alcance de todas las bibliotecas actuales herramientas, soluciones y técnicas que hasta ahora solo estaban al alcance de unas pocas empresas. Los usuarios llaman a la puerta de la biblioteca virtual y se nos presenta una oportunidad única para hablar, escuchar, dialogar e intentar hacerle la vida más fácil.