

Re-ingeniería y Modernización de la Biblioteca Virtual Galega

Delfina Ramos-Vidal, Alejandro Cortiñas, Miguel R. Luaces,
Óscar Pedreira y Angeles Saavedra Places

Universidade da Coruña, Centro de Investigación CITIC, Laboratorio de Bases de Datos, Facultade de Informática, Elviña, 15071 A Coruña, Spain
{delfina.ramos,alejandro.cortinas,luaces,
opedreira,asplaces}@udc.es

Resumen La Biblioteca Virtual Galega (BVG) es una biblioteca digital de referencia en literatura gallega de todos los tiempos, para el público general y especialmente para centros de educación de enseñanza obligatoria de la comunidad. Esta plataforma almacena y da acceso a información bibliográfica de autores clásicos y actuales, así como información multimedia correspondiente a ediciones digitales de sus obras. En este artículo se presenta el proyecto de re-ingeniería y modernización de esta biblioteca digital (desarrollada en 2002) utilizando una línea de producto software para la generación de bibliotecas digitales desarrollada por investigadores del Laboratorio de Bases de Datos (LBD).

1. La Biblioteca Virtual Gallega

Las bibliotecas digitales permiten almacenar y gestionar grandes colecciones de recursos digitales junto con su información bibliográfica asociada. Su uso más destacado se da en el ámbito cultural. Las bibliotecas digitales juegan un papel fundamental a la hora de hacer accesibles a través de la red grandes colecciones de documentos ya que, en algunos casos, esos documentos proceden de digitalizaciones de fondos bibliográficos de gran valor cultural a los que el público no podría acceder de otro modo.

La Biblioteca Virtual Galega¹ (BVG) es una biblioteca digital dedicada a la difusión de literatura gallega. La BVG nació en el año 2002 como un proyecto impulsado por la Universidade da Coruña, con financiación también de la Diputación da Coruña y la Real Academia Galega, y fue desarrollada por investigadores del ámbito de la filología gallego-portuguesa y por investigadores del Laboratorio de Bases de Datos (LBD). La BVG se ha convertido en un recurso de referencia en la literatura gallega, ofreciendo a la ciudadanía, principalmente profesorado y estudiantado de primaria y secundaria, datos bibliográficos de autoras/es, obras completas o fragmentos de ellas, recursos lingüísticos y complementos como imágenes y archivos de audio y vídeo con entrevistas y reportajes. Actualmente tiene catalogados 353 autoras/es y un total de 4.770 obras, de las

¹ <https://bvg.udc.es/>

cuales alrededor de 1.000 se pueden consultar en formato digital de forma pública en formato imagen, texto, audio y/o vídeo.

Tras casi 20 años en funcionamiento, varios aspectos de la BVG han quedado obsoletos. Además de su aspecto gráfico, otros elementos como la experiencia de usuario, determinadas funcionalidades y la tecnología utilizada en la plataforma requieren una actualización. Así, la Universidade da Coruña plantea un proyecto para realizar el desarrollo de una nueva plataforma que gestione el catálogo bibliográfico soportado actualmente y que permita que este crezca.

2. Línea de producto software para Bibliotecas Digitales

La BVG no ha sido la única colaboración de los investigadores del LBD con otros grupos de investigación e instituciones en el ámbito de las Humanidades. El LBD ha participado en numerosos proyectos de desarrollo de bibliotecas digitales que lo han llevado a formar parte de la *Red Aracne: Red de Humanidades Digitales y Letras Hispánicas*², de la que forma parte como socio tecnológico desde 2013. Tres de las bibliotecas más recientes que ha desarrollado son: Symbola: Divisas o Empresas Históricas³, CBDRS: Catálogo y Biblioteca Digital de Relaciones de Sucesos⁴ o el Archivo Digital Valle-Inclán⁵.

La experiencia acumulada en el dominio ha hecho patente la dificultad y el coste de desarrollar una biblioteca digital desde cero, tanto en términos temporales como monetarios. Ante la perspectiva de nuevas iniciativas en este sentido, se llevó a cabo el diseño y desarrollo de una línea de producto software (LPS) para la generación de bibliotecas digitales. El objetivo de las LPS es generar una familia de sistemas software similares que, como en el caso de las bibliotecas digitales, comparten un conjunto de características y funcionalidades. Para el modelado de las características de la LPS y la arquitectura y tecnologías de los productos que genera se tuvieron en cuenta también las principales bibliotecas digitales en ámbitos nacionales e internacionales, en concreto, la Biblioteca Virtual Miguel de Cervantes⁶ y la Library of Congress⁷.

La LPS para Bibliotecas Digitales cuenta con una base de datos que ofrece la posibilidad de almacenar cualquier tipo de obra y sus ediciones, colecciones, etc.; registrando información sobre las personas involucradas en su creación y publicación; con información geográfica de donde ocurre la obra, donde nace/vive quien la escribe o donde es publicada; recopilando los diferentes formatos digitales en que se presenta la obra: imagen, texto, audio y/o vídeo; y guardando información sobre los diferentes usuarios y permisos con los que pueden acceder. La LPS también ofrece funcionalidades de navegación y búsqueda (generalista, avanzada/compleja, por contenido, por localización geográfica) en diferentes idiomas a

² <http://www.red-aracne.es/>

³ <https://www.bidiso.es/Symbola/>

⁴ <https://www.bidiso.es/CBDRS/>

⁵ <https://www.archivodigitalvalleinclan.es>

⁶ <http://www.cervantesvirtual.com/>

⁷ <https://www.loc.gov/>

diferentes perfiles de usuario públicos y privados, presentando la información de manera conveniente sobre listados y mapas; permite exportar la información encontrada a diferentes formatos como CSV, XLS o PDF; e implementa el protocolo OAI-PMH para el intercambio de información con agregadores como Galiciana, Hispana y Europea. Además, da soporte a todo el flujo de trabajo de gestión y revisión de datos, permitiendo gestionar la información almacenada a través de interfaces específicamente diseñadas para el tipo de información de que se trate, facilitando y automatizando la alimentación, y facilitando los procesos de revisión por los que las obras pasan antes de ser publicadas definitivamente.

Aunque en un primer momento se pudiera pensar que todas las bibliotecas digitales son prácticamente iguales, en realidad presentan una variabilidad importante, tanto en el tipo de datos que almacenan y en cómo esos datos se relacionan entre sí, como en las funciones que los usuarios necesitan. Así, la LPS soporta un total de 162 características, que se estructuran en las categorías de objetos de la biblioteca, explotación de los datos, gestión de usuarios, acceso a la biblioteca e idioma. La implementación de la LPS se basó en el motor de derivación basado en *scaffolding* presentado en [1]. La Figura 1 presenta los primeros niveles del árbol de características. El detalle de la línea de producto fue presentado en [2]. Como parte del desarrollo de la LPS para bibliotecas digitales se generaron dos proyectos de prueba. El primero fue una biblioteca “teórica” que cubría una selección de características plausible, pero artificial. El segundo fue la re-ingeniería de las *Colecciones y Recursos Digitales de BIDISO*, miembro de la Red Aracne, que incluye a su vez bibliotecas y catálogos menores que encajan en el marco de la variabilidad de la LPS construida, permitiendo realizar un amplio conjunto de pruebas.

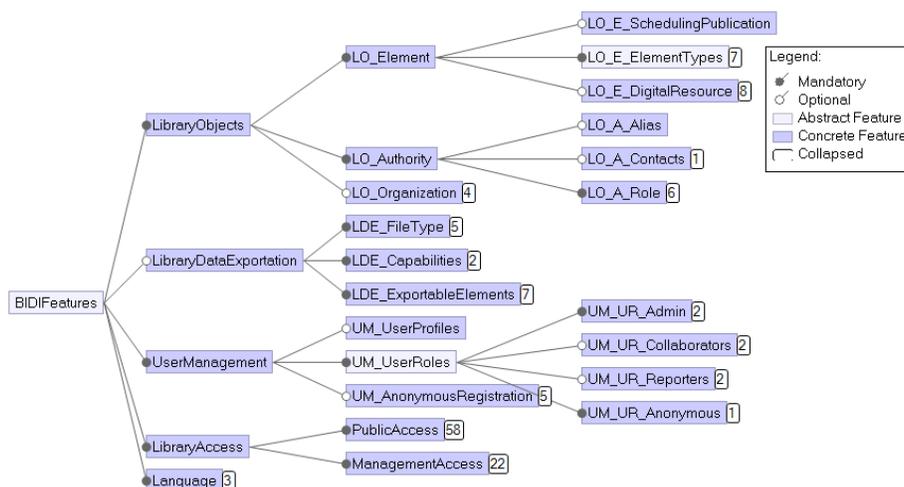


Figura 1. Árbol de características de la LPS para bibliotecas digitales

3. Re-ingeniería de la BVG

En el momento de escribir este resumen del proyecto, se han completado los primeros pasos del mismo. En esta primera fase se ha analizado la variabilidad de la BVG y se ha utilizado la LPS para generar una primera versión. Algunas características relevantes de la LPS que se aplican en el caso de la BVG son la inclusión de alias de autores, la falta de perfiles de usuario más allá de los administradores y la oferta de recursos bibliográficos en distintos formatos, entre ellos: PDF, imagen, vídeo y audio. El producto resultante tiene un total de 161 ficheros, de los cuales 110 son clases Java y los restantes son ficheros JavaScript y/o archivos de configuración del proyecto. Evidentemente, el producto generado no se puede desplegar tal cual en un entorno de producción, pues es necesario refinar la implementación de ciertos aspectos y editar por completo el aspecto gráfico de la interfaz de usuario. Sin embargo, valoramos que el producto generado automáticamente cubre el alcance funcional necesario. La reducción de coste se hace notoria al producir diferentes productos, concretamente a partir del tercer desarrollo software basado en una LPS. Consideramos el tercer producto un punto de inflexión a partir del cual el número de líneas de código generadas automáticamente superan el número de líneas de código escritas manualmente. A partir de los resultados obtenidos en esta primera prueba, valoramos también que aplicar tecnología LPS para el desarrollo de bibliotecas digitales facilitará en gran medida que muchos grupos de investigación del ámbito de las humanidades puedan desarrollar sus propias plataformas con un coste mucho más reducido que el de un desarrollo a medida, lo que hará que esta tecnología sea mucho más accesible para el desarrollo de bibliotecas digitales centradas en ámbitos literarios muy concretos.

Referencias

1. Cortiñas, A. and Luaces, M. R. and Pedreira, O. and Places, A. S.: Scaffolding and in-browser generation of web-based GIS applications in a SPL tool. In: Proc. 21st Int. Systems & Software Product Line Conference (SPLC) vol.2. pp. 46–49 (2017)
2. Ramos-Vidal, D., Cortiñas, A., Luaces, M.R., Pedreira, O., Places, Á.S.: A software product line for digital libraries. In: Procs. of the 16th Int. Conf. on Web Information Systems and Technologies, (WEBIST). pp. 381–389. SCITEPRESS (2020)

Agradecimientos

Parcialmente financiado por: MICIU / FEDER-UE: MAGIST (PID2019-105221RB-C41), EXTRACompact (PID2020-114635RB-I00), Red Aracne-Nodus (RED2018-102755-T), y BIZDEVOPSGLOBAL (RTI-2018-098309-B-C32); Xunta de Galicia / FEDER-UE, ConectaPeme, GEMA (IN852A 2018/14), GRC (ED431C 2017/58); MINECO-AEI / FEDER-UE Datos 4.0 (TIN2016-78011-C4-1-R) y Velocity (TIN2016-77158-C4-3-R); centro de investigación CITIC financiado por la Xunta y la UE a través del Fondo Europeo de Desarrollo Regional - Programa Galicia 2014-2020 (ED431G 2019/01).