

Desarrollo de docencia universitaria de informática documental con herramientas de software libre

Jesús Tramullas Saz*, Piedad Garrido Picazo**

*Dpto. de Ciencias de la Documentación
Universidad de Zaragoza

Pedro Cerbuna 12, 50009 Zaragoza

**Dpto. de Informática e Ingeniería de Sistemas
Universidad de Zaragoza
Ciudad Escolar s/n, Teruel
tramullas@unizar.es, piedad@unizar.es

Resumen

La informática documental es una especialización de la informática, que se ocupa del tratamiento y gestión de la información recogida en documentos digitales. Este campo ha sido dominado durante largo tiempo por software privativo, lo que se reflejaba en las prácticas de las asignaturas. La creciente disponibilidad de herramientas de software libre, capaces de dar soporte a los procesos y tareas de la informática documental ha hecho posible la integración de las mismas en el proceso de aprendizaje. En este trabajo se detalla la integración de estas herramientas en el contexto de la adaptación de una asignatura especializada al Espacio Europeo de Educación Superior.

1. Introducción

La enseñanza universitaria de los métodos y técnicas de tratamiento y gestión de la información documental se organiza en las titulaciones Diplomatura en Biblioteconomía y Documentación y Licenciatura en Documentación, que se ven complementadas con los correspondientes programas de doctorado. Esta organización va a variar a causa del desarrollo del Espacio Europeo de Educación Superior, que estructura la enseñanza de la disciplina en grado, máster y doctorado.

La gestión de la información tiene como parte esencial la aplicación de tecnologías de la información y la documentación, de tal forma que no pueden entenderse una sin la otra. Ambas han dado forma a la disciplina, que ha recibido denominaciones como Documentación Automatizada o Informática documental. Las directrices establecidas por los Reales Decretos que regulan estos estudios así lo contemplan,

incluyendo específicamente las Tecnologías de la Información y las Bases de Datos. El libro blanco sobre el Título de Grado en Documentación, punto de partida para los futuros planes de estudio, incide en los contenidos digitales, y en los sistemas y herramientas para su creación, implementación y gestión.

Cada universidad ha establecido en sus planes de estudio la forma en que se organizan los créditos correspondientes, generalmente en asignaturas anuales o cuatrimestrales. La variedad de títulos y de contenidos de las mismas es elevada, aunque puede rastrearse un núcleo común en todas ellas, cuyas materias corresponden a:

- Bases de datos documentales y bibliográficas
- Edición electrónica
- Técnicas y herramientas de recuperación de información
- Sistemas electrónicos de gestión documental
- Bibliotecas digitales y repositorios institucionales
- Sistemas de informatización de bibliotecas y archivos
- Servicios de información documental en Internet: motores de búsqueda, esquemas de metadatos...

Estos contenidos básicos pueden verse complementados mediante asignaturas optativas, que inciden en diferentes aspectos especializados.

2. La asignatura “Sistemas electrónicos de tratamiento documental”

La asignatura *Sistemas electrónicos de tratamiento documental* se incardina en el contexto indicado en el apartado 1. Se trata de una

asignatura troncal de seis créditos (dos teóricos y cuatro prácticos), que se imparte durante el primer cuatrimestre del segundo curso de la Diplomatura en Biblioteconomía y Documentación de la Universidad de Zaragoza.

El contenido teórico de la misma viene dado por las directrices generales de la titulación, por una parte, y por la organización interna de la docencia en la Diplomatura, lo que ha facilitado la coordinación de contenidos y actividades con el resto de asignaturas. El contenido práctico, sin embargo, venía determinado en cada momento por varios factores:

1. La orientación, demandada por el mercado, a formar usuarios finales de las herramientas informáticas
2. El dominio del ámbito de trabajo informativo-documental por parte de productos de compañías de software propietario o privativo
3. La disponibilidad gratuita, o por convenio, para la docencia de las herramientas de software privativo
4. La escasez de documentación de libre acceso para esas herramientas

La actividad práctica se llevaba a cabo en laboratorio, en el cual los alumnos desarrollaban prácticas individuales guiadas con las diferentes herramientas de software privativo.

Evidentemente, todos estos factores afectaban a la formación práctica que los alumnos podían desarrollar, a causa de la variable disponibilidad de las herramientas necesarias para ello. Si bien la enseñanza impartida hasta el momento había asegurado la formación del alumno con la calidad exigible, esta dependencia de herramientas privativas de terceros podía hipotecar la actividad docente en un entorno cada vez más cambiante y dinámico.

3. Adaptación al Espacio Europeo de Educación Superior

El Espacio Europeo de Educación Superior (EEES) propone un cambio efectivo en la docencia universitaria, primando el trabajo del alumno y su proceso de aprendizaje, frente a modelos docentes que podrían calificarse como más tradicionales. Son numerosas las publicaciones y trabajos que recogen las

características y enfoques del EEES, por lo que no es necesario insistir sobre estas cuestiones en este trabajo.

En este contexto, y en el marco del proceso de adaptación al EEES que se está llevando a cabo en la Diplomatura en Biblioteconomía y Documentación de la Universidad de Zaragoza, se ha llevado a cabo un rediseño de la asignatura *Sistemas electrónicos de tratamiento documental*. Para ello, ha sido necesario replantear los objetivos de la misma, así como las competencias y habilidades que los alumnos deben desarrollar durante su desempeño. En consecuencia, se ha considerado pertinente llevar a cabo un completo proceso de revisión, planificación y adecuación de la actividad práctica, de forma que no se viese constreñida por los vaivenes y variaciones del mercado de software privativo y de sus productos. Inmersos en este proceso, también se rediseñó la propia actividad práctica, mediante la integración y uso de modelos de trabajo del alumno no contemplados hasta este momento.

La definición de competencias y habilidades trajo consigo un cambio del enfoque de la asignatura, que supone también cambios estratégicos en la concepción de las actividades prácticas que se llevan a cabo con herramientas informáticas. En primer lugar, se han establecido cuatro principios generales:

1. La adquisición de competencias y habilidades no debe verse limitada por el contexto comercial en el que se desarrolle la actividad profesional
2. La disponibilidad de herramientas software no debe restringir la adquisición de competencias y habilidades
3. El alumno debe poder desarrollar su actividad de aprendizaje en diferentes contextos: aula, laboratorio, entorno doméstico, otros entornos informáticos...
4. El alumno debe poder localizar y acceder a documentación externa suficiente para llevar a cabo su propio proceso de aprendizaje

El respeto de los principios generales lleva, a limitar la excesiva influencia del entorno profesional sobre la docencia, y a superar la orientación meramente instrumental de este tipo de asignaturas. Si bien no excluye la integración de herramientas de software privativo, la

dificultad para que éstas puedan cumplir los principios tercero y cuarto limita realmente su uso potencial. Puede argumentarse que la exclusión de estas herramientas privativas, adecuadas a lo que está demandando el mercado, es, en realidad, un perjuicio para el alumno, al alejar su formación de las demandas del mercado laboral. La respuesta a esta cuestión clave se abordará en el apartado 4.

El segundo aspecto de adaptación, una vez definidos los principios generales, ha sido el cambio de modelo de las actividades prácticas. Como se ha indicado anteriormente, éstas se llevaban a cabo principalmente en laboratorio. Sin embargo, el EEES propone un cambio en la actividad de aprendizaje, insistiendo en la actividad propia del alumno, independientemente de su localización y disponibilidad espacial y temporal. Además, incluye una clara orientación al desarrollo de habilidades de aprendizaje autónomo y trabajo en equipo y en colaboración.

En consecuencia, y aprovechando las posibilidades que ofrece el cambio en las herramientas informáticas, se ha optado por cambiar el modelo de práctica presencial e individual en laboratorio, or el modelo de trabajo en equipo en colaboración, en diferentes localizaciones y con flexibilidad horaria. De la misma forma, se ha producido también el paso de la actividad de práctica guiada hacia el modelo basado en el desarrollo de proyectos. Como se analizará en el apartado 4, esto no hubiese sido posible sin el cambio estratégico a herramientas de software libre. Las posibilidades que ofrece este tipo de cambio ya han sido señaladas por Faber [1].

Por último, debe destacarse que el cambio y adaptación también ha afectado a los mecanismos de evaluación. Los alumnos desarrollan un proceso personal de evaluación de su propio desempeño en los proyectos, así como del resto de sus compañeros, de las herramientas utilizadas y de la documentación de soporte disponible en cada caso.

4. Software libre para tratamiento y gestión de información documental

Como se ha señalado en el apartado anterior, la cuestión clave es la disponibilidad de herramientas de software no privativo que ofrezcan las prestaciones técnicas necesarias para

que los alumnos puedan desarrollar las competencias y habilidades establecidas para la asignatura. En la bibliografía especializada pueden encontrarse trabajos que recogen experiencias en diferentes campos de aplicación de la informática [2] La experiencia docente e investigadora previa sugería que existían todas las herramientas necesarias para ello dentro del campo del software libre [3].

En consecuencia, se llevó a cabo un estudio de campo para dilucidar la cuestión. Para ello, se exploraron sistemáticamente los principales servidores para proyectos de software libre, se analizaron los contenidos de las sedes web de referencia sobre software libre aplicado al tratamiento y recuperación de información en bibliotecas y servicios de información y documentación, y se revisó detenidamente la bibliografía especializada sobre este tema [4]. El resultado del estudio ha identificado herramientas de software libre, especializadas en procesos y tareas que se llevan a cabo en los servicios de información, que pueden agruparse en las siguientes categorías:

1. Bases de datos bibliográficas: sistemas de bases de datos para la construcción de repertorios de bibliografía, tanto en soportes tradicionales como en entornos digitales. Ejemplos: Connotea (www.connotea.org), WIKINDEX (wikindx.sourceforge.net/), RefBase (refbase.sourceforge.net),
1. Motores de indización y recuperación: herramientas para la extracción de información textual de colecciones de documentos, y para su indización automática, así como para la creación de interfaces de recuperación de información. Ejemplos: Lucene/Nutch (lucene.apache.org), Xapian (www.xapian.org), SWISH-E (swish-e.org)
2. Sistemas de gestión de documentos: sistemas completos para la definición de ciclos de vida de documentos, así como flujos de trabajo, control de versiones y de acceso, etc. Ejemplos: Alfresco Content Management (www.alfrescosoftware.com), KnowledgeTree (www.knowledgetree.com)
3. Sistemas de informatización de bibliotecas: herramientas de software libre que dan soporte a todas las funciones y procesos técnicos que se llevan a cabo en una biblioteca, como

- catalogación, circulación, gestión administrativa, etc. Ejemplos: Koha (www.koha.org), PMB (www.sigb.net)
4. Directorios de recursos digitales: permiten crear portales de acceso a los recursos de información digital, de todo tipo, que pueda necesitar una organización. Ponen especial énfasis en las tareas de administración y control de metadatos, ya que suelen usar Dublin Core. Ejemplos: iVIA (ivia.ucr.edu) Scout Portal (scout.wisc.edu/Projects/SPT/)
 5. Bibliotecas digitales: aplicaciones para la construcción y mantenimiento de colecciones de documentos completos (textuales, gráficos, etc), organizados y estructurados alrededor del concepto de "colección". Dan prioridad al acceso de los usuarios. Ejemplos: Greenstone (www.greenstone.org)
 6. Repositorios digitales: aplicaciones para el desarrollo de repositorios de documentos digitales, que contengan el acervo científico de una institución. Dan prioridad a los procesos de publicación. Ejemplos: Eprints (eprints.org), DSpace (dspace.org)

Junto a todos los anteriores cabría añadir las herramientas para edición de documentos digitales y para gestión de metadatos. Dado que las posibles elecciones son muy numerosas, no se han incluido como una herramienta específica de gestión de información documental. Igualmente, no se han considerado los sistemas de gestión de contenidos o de portales de tipo generalista, ya que se entiende, precisamente, que son herramientas básicas de gestión de información, y que su utilidad dependerá, precisamente, de las características del proyecto en el que se encuadre su utilización. No debe dejar de tenerse en consideración que la mayor parte de los procesos que tienen lugar en cualquier sistema de gestión de contenidos corresponden, en un setenta u ochenta por ciento, a tareas básicas de tratamiento y gestión de información.

Otro factor de suma importancia, inherente a las herramientas de software, es la utilización de estándares. Todas estas herramientas implementan estándares, generalmente públicos y abiertos, como XML, Z39.50, MARC, Dublin Core, OAI, etc. En el campo de la información y la documentación los estándares tienen una máxima importancia, hasta el punto de tener su propio

comité específico en ISO y en las diferentes agencias nacionales de normalización, como AENOR. Los alumnos de información y documentación deben tener un claro conocimiento de las normas y estándares que afectan a su actividad, y la utilización de herramientas de software libre refuerza esta habilidad. Es este sentido destaca el recientemente aprobado estándar para documentos ofimáticos OpenDocument (ISO 26300:2006), que establece, como su propio nombre indica, el único estándar internacional para los documentos más comunes en los entornos digitales, como son los resultantes de la actividad ofimática de las organizaciones. En consecuencia, la adscripción a estándares es un requerimiento que se establece para las herramientas software que se vayan a utilizar. Igualmente, es el formato ofimático que se exige a los estudiantes para la presentación e intercambio de trabajos escritos, informes, formularios de evaluación, etc. Las herramientas seleccionadas incorporan, en sus funcionalidades, los estándares para el tratamiento y recuperación de la información señalados anteriormente en este párrafo, lo que asegura su conocimiento y utilización por parte del estudiante.

Como se ha indicado en el apartado 3, cabe discutir si este enfoque limita las posibilidades de aprendizaje del estudiante. En primer lugar, el cambio hacia un modelo basado en el aprendizaje activo de habilidades permite que el estudiante adquiera las pautas básicas para saber plantear y desarrollar actividades frente a cualquier herramienta con la que se encuentre en su actividad profesional. En segundo lugar, la libre disponibilidad de las herramientas facilita la instalación y replicación de plataformas de actividad en cualquier máquina, sin estar obligatoriamente localizada en un aula universitaria, como era el caso con el software privativo. Ambas cuestiones son complementadas por los procesos de documentación que el estudiante debe llevar a cabo para conocer y utilizar las herramientas libres, y que inciden en su capacidad de autoaprendizaje.

El desarrollo de las actividades mediante proyectos en colaboración, en entornos distribuidos, también es posible gracias a la utilización de herramientas libres, ya que las características de las herramientas privativas

obligaban al planteamiento individual de las prácticas.

Por último, debe considerarse que la mayoría de los alumnos disponen, particularmente, de plataformas Windows, de la misma forma que las salas de usuarios de la Universidad de Zaragoza ofrecen exclusivamente este tipo de sistemas operativos. Esta incomprensible restricción impone que las herramientas que sean seleccionadas pueden instalarse sin complicaciones en plataformas GNU/Linux, Windows y Mac OS X. De esta forma se asegura el acceso de los alumnos a los medios necesarios.

5. Integración de software libre como herramienta de aprendizaje

El desarrollo de la asignatura *Sistemas electrónicos de tratamiento documental* se lleva cabo mediante la integración de herramientas de software libre, que dan soporte al proceso de aprendizaje de los estudiantes. Las herramientas se han organizado en tres grupos, cada uno de los tiene diferentes objetivos y tareas, en el marco de la planificación global de la docencia.

- Soporte ofimático: actividades en las cuales los alumnos llevan a cabo la creación, modificación e intercambio de información mediante documentos ofimáticos. Todos estos documentos deben cumplir el estándar ISO 26300 OpenDocument. El paquete de software elegido para ello es *OpenOffice*, dada su disponibilidad para plataformas GNU/Linux, Windows y Mac OS X.
- Portal de la asignatura: la asignatura pone a disposición de los estudiantes un portal propio, independiente de los sistemas de aula virtual de la universidad. Para ello utiliza un sistema de gestión de contenidos libre, *Drupal* (www.drupal.org, fig.1). Este servicio pone a su disposición foro de discusión, descarga de materiales docentes, noticias, encuestas y sindicación mediante RSS. Las relaciones entre software libre y lo que se ha dado llamado en llamar open courseware pueden verse en [5].

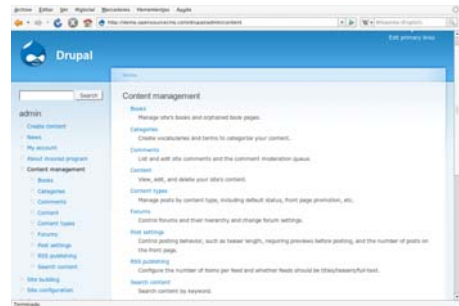


Figura 1. Portal sobre Drupal

- Sistemas de tratamiento documental: son los encargados de soportar el núcleo del proceso de adquisición de competencias y habilidades, ya que son las herramientas con las cuales desarrollarán los proyectos que se les requieran. Los dos paquetes seleccionados son *KnowledgeTree* (para gestión documental, www.knowledgetree.com), y *Greenstone* (para bibliotecas y repositorios digitales, www.greenstone.org). Los paquetes permiten gestionar los documentos digitales desde la perspectiva de tratamiento y recuperación de información documental, incluyen motores de indexación y recuperación, y permiten trabajar con diferentes esquemas de metadatos.

Si bien todas las herramientas, como se ha indicado previamente, pueden instalarse en diferentes plataformas, el portal de la asignatura es instalado y gestionado por los docentes en un servidor propio (disponible en imhotep.unizar.es), sobre la distribución GNU/Linux Ubuntu 7.04 Feisty Fawn. En el mismo se dispone también de versiones de producción de *KnowledgeTree* y de *Greenstone*, para que los alumnos puedan comprobar el funcionamiento de los productos resultantes de su trabajo en un entorno real. El servidor y las herramientas van siendo actualizadas conforme aparecen nuevas versiones.

Una vez seleccionadas las herramientas que van a ser utilizadas en el proceso de aprendizaje, es necesario determinar la forma en que las mismas se van a incorporar a la actividad del alumno. Como se ha indicado previamente (apartado 3), el modelo seleccionado ha sido el desarrollo de proyectos mediante trabajo en grupo

en colaboración. Los proyectos que los alumnos llevan a cabo son los siguientes:

1. Proyecto gestión documental: el proyecto debe definir las características, el ciclo de vida y el flujo de trabajo de un conjunto de documentos digitales, incluyendo control de versiones y de modificaciones, asignación de esquemas de metadatos, e indización del texto completo. Para este proyecto se usa *KnowledgeTree* (fig. 2).

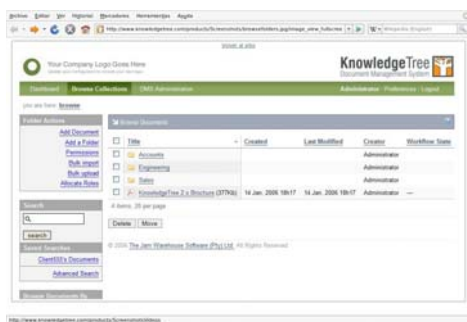


Figura 2. Servidor KnowledgeTree

2. Proyecto colección de documentos digitales: el proyecto debe crear una colección de documentos digitales, en diferentes formatos, incluyendo la indización a texto completo de su contenido, así como la asignación de metadatos y la publicación de la colección en internet para su aprovechamiento por usuarios finales. Para este proyecto se usa *Greenstone* (fig. 3).
3. Proyecto de gestión de contenidos: es el único proyecto que se lleva cabo de manera individual. Para desarrollar las habilidades básicas de gestión de contenidos en entornos digitales, el alumno debe crear, configurar y mantener un blog. A este fin, debe obtener y utilizar una cuenta en cualquiera de los servidores gratuitos de blog disponibles en internet (blogger, blogia, lacocletera, etc.).



Figura 3. Servidor Greenstone

El planteamiento práctico de la asignatura, como puede apreciarse, busca que los estudiantes desarrollen las competencias y habilidades de la disciplina con herramientas que les obligan a desarrollar sus propios procesos de planificación, documentación e implementación. Los diferentes proyectos tienen como punto de partida un guión de desarrollo y una clase introductoria a la herramienta seleccionada, y a partir de este momento los procesos de documentación, aprendizaje, instalación y configuración de la herramienta quedan bajo la responsabilidad del alumno. Se ha preparado documentación de referencia especialmente orientada a este tipo de usuario [6]. Una vez operativa la aplicación, se inicia la implementación del proyecto de información y documentación digital en sí mismo. Durante todo el ciclo de vida del proyecto los grupos de trabajo disponen de tutorías de grupo y de soporte bajo demanda por parte de los docentes responsables de la asignatura.

6. Conclusiones

El cambio de modelo de desarrollo que ha supuesto el movimiento de software libre influye en cada vez más actividades, y plantea desafíos y soluciones de verdadero interés [7]. La planificación docente expuesta en los apartados anteriores ha sido puesta en marcha y validada durante el curso 2006-2007. Los resultados obtenidos permiten establecer las siguientes conclusiones:

- Las herramientas de software libre especializadas disponibles para la gestión de

información documental cumplen todos los requerimientos necesarios para su uso docente

- Estas herramientas cubren las competencias y habilidades prácticas que los estudiantes deben adquirir
- El uso de las mismas en proyectos refuerza las actitudes y aptitudes de autoaprendizaje deseables para los futuros profesionales de gestión de información y documentación.
- Los problemas encontrados en la aplicación del modelo expuesto se centran en la falta de documentación orientada al usuario final, y en la carencia de habilidades de autoaprendizaje por parte de los estudiantes.

La amplia variedad de herramientas disponibles pueden hacer posible extender la utilización del software libre a un campo más amplio de actuación. La creación de motores especializados de búsqueda para intranet/internet, o la recopilación de recursos de información digital de valor añadido, mediante directorios, o la creación de bases de datos bibliográficas científicas y técnicas, son otras áreas en las que pueden integrarse perfectamente. Además, debe considerarse que en breve plazo las herramientas de software libre especialmente diseñadas para desarrollar el web semántico, en el cual predominan los métodos y técnicas de gestión de la información, potenciarán aún más las posibilidades de formación de especialistas en información y documentación digital, ofreciendo un mayor potencial que las opciones de software privativo.

Referencias

- [1] Faber, B.D. "Educational models and open source: resisting the proprietary university." *Proceedings of the 20th annual international conference on Computer documentation*, 2002, 31-38.
- [2] Zaritski, R.M. "Using open source software for scientific simulations, data visualization, and publishing" *Journal of Computing Sciences in Colleges*, 19, 2, 2003, 218-222.
- [3] Tramullas, J. "Software libre para gestión de recursos de información digital." *Encuentro Internacional sobre Conocimiento Libre, II Conferencia Internacional de Software Libre*, Badajoz, 2006, 346-356.
- [4] Chawner, B. *Open Source Software and Libraries Bibliography*. 2006. (disponible en http://www.vuw.ac.nz/staff/brenda_chawner/biblio.html)
- [5] Baldi, S., Heier, H., Mehler-Bicher, A. "Open Courseware and Open Source Software." *Communications of the ACM*, 2003, 46, 9, 105-107.
- [6] Tramullas, J. Garrido, P. (coords.), *Software libre para servicios de información digital*. Madrid: Prentice Hall, 2006.
- [7] Wolf, M.J., Bowyer, K., Gotterbam, D., Miller, K. "Open source software: intellectual challenges to the status quo." *Proceedings of the 33rd SIGCSE technical symposium on computer science education*, 2002, 317-318.