

Software libre para repositorios institucionales: propuestas para un modelo de evaluación de prestaciones

Por Jesús Tramullas y Piedad Garrido Picazo



Jesús Tramullas Saz es profesor titular en el Depto. de Ciencias de la Documentación de la Universidad de Zaragoza. Miembro del grupo de investigación sobre Gestión de Recursos de Información en las Organizaciones (GRIO). Coordinador de la Red temática sobre Documentación Digital (Plan Nacional de I+D+I, 2004-2005). Investigador principal del proyecto «Web semántico y bibliotecas digitales: desarrollo de servicios de información basados en RDF y Topic Maps» (2006-2007). Sus líneas de investigación se centran en bibliotecas digitales y servicios de información digital, y herramientas de software libre para la gestión de información.

Piedad Garrido Picazo es profesora asociada en el Depto. de Informática e Ingeniería de Sistemas de la Universidad de Zaragoza. Miembro del grupo de investigación sobre Gestión de Recursos de Información en las Organizaciones (GRIO). Perteneció a las redes temáticas sobre Recuperación de Información en Textos y Bibliotecas Digitales; Documentación Digital; y Sistemas de Acceso a la Información en la Web basados en Soft-Computing. Sus líneas de investigación son bases de datos xml, software libre para gestión de información, bibliotecas digitales, RDF y Topic Maps (XTM) en el contexto del web semántico.



Resumen: Los repositorios institucionales se han convertido en la principal forma de publicar, preservar y difundir la información digital de las organizaciones. La mayoría de ellos están soportados por software libre, elegidos tras la evaluación de sus prestaciones. Sin embargo, los análisis que se han realizado han sido comparativos, atendiendo a las funciones que ofrecen, antes que a otros factores. Para superar esta limitación, este trabajo analiza los diferentes modelos publicados de evaluación de software para repositorios institucionales, estudia los enfoques adoptados y propone un modelo orientado a la definición de procesos informativo-documentales, a la comunidad de usuarios, a las características de las colecciones, y al contexto del proyecto en el que se enmarcan.

Palabras clave: Repositorios institucionales, Software libre, Evaluación de prestaciones.

Title: Open source software for institutional repositories: proposal for a model to evaluate features

Abstract: Institutional repositories have become an important means of publishing, disseminating, and preserving the digital information of organisations. The majority of them run on open software tools. Institutional repository projects are being carried out following an evaluation of the software's features. Nevertheless, the evaluations to date have been comparative, focussing on the features offered by each tool, rather than on other factors. In order to overcome this limitation, this work analyses the different published models for evaluating institutional repository software, studies the approach adopted in each case, and proposes a model. This model is oriented toward the study and definition of informational-documental processes, the users, the characteristics of the collections, and the context of the project in which it is framed.

Keywords: Institutional repositories, Open source software, Evaluation of features.

Tramullas, Jesús y Garrido Picazo, Piedad. «Software libre para repositorios institucionales: propuestas para un modelo de evaluación de prestaciones». En: El profesional de la información, 2006, mayo-junio, v. 15, n. 3, pp. 171-181.

Repositorios institucionales

Los repositorios de documentos digitales se están convirtiendo progresivamente en los lugares en los que numerosas organizaciones almacenan y organizan el resultado de sus actividades. A partir de ellos pueden

construir sistemas de publicación de revistas digitales, bibliotecas de tesis doctorales o sistemas más complejos de soporte a actividades de gestión del conocimiento. La importancia de este tipo de recursos de información comienza a considerarse como estratégica, lo que explica la creciente implantación de estas plata-

Artículo recibido el 27-02-06
Aceptación definitiva: 21-03-06

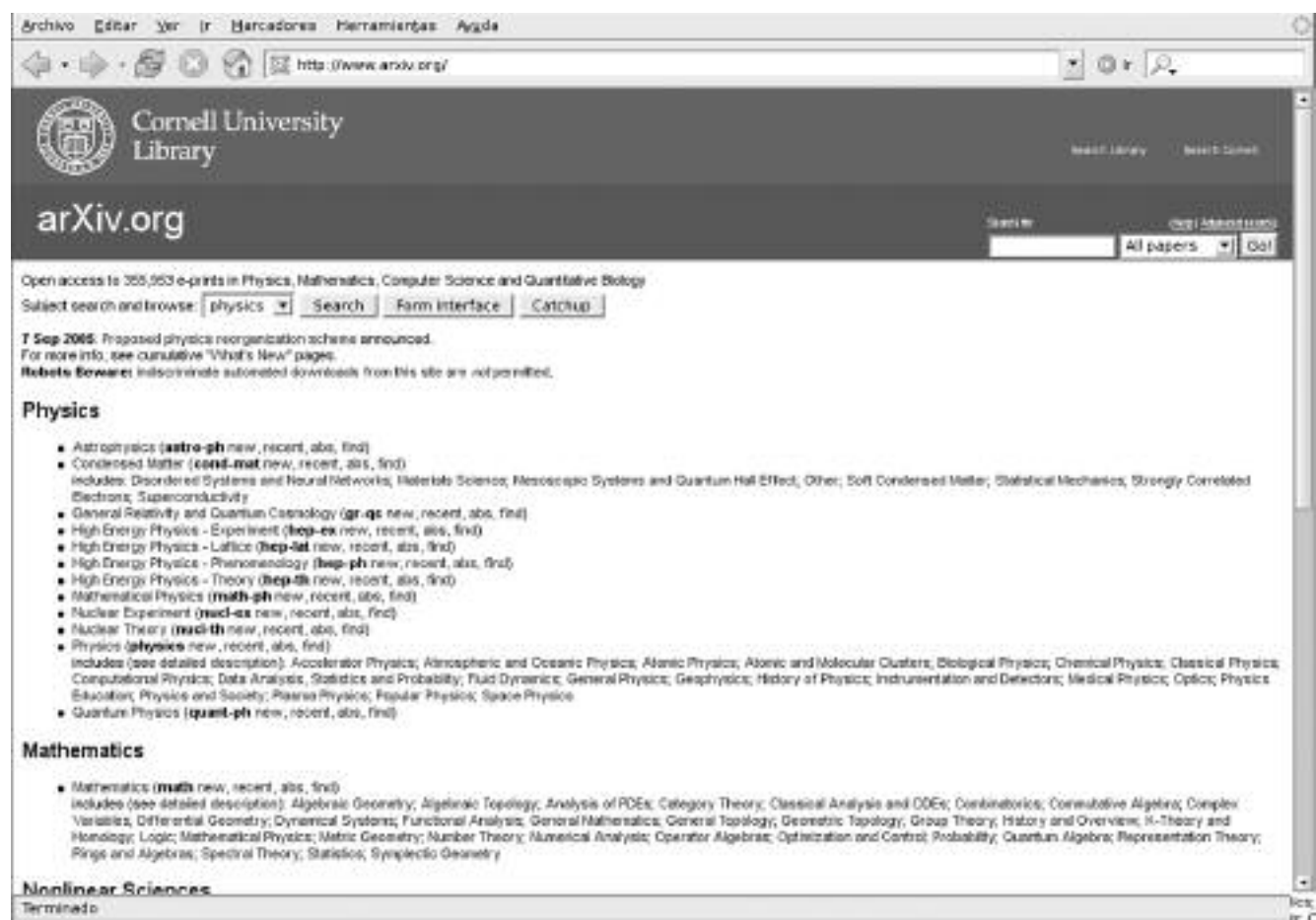


Figura 1. arXiv.org

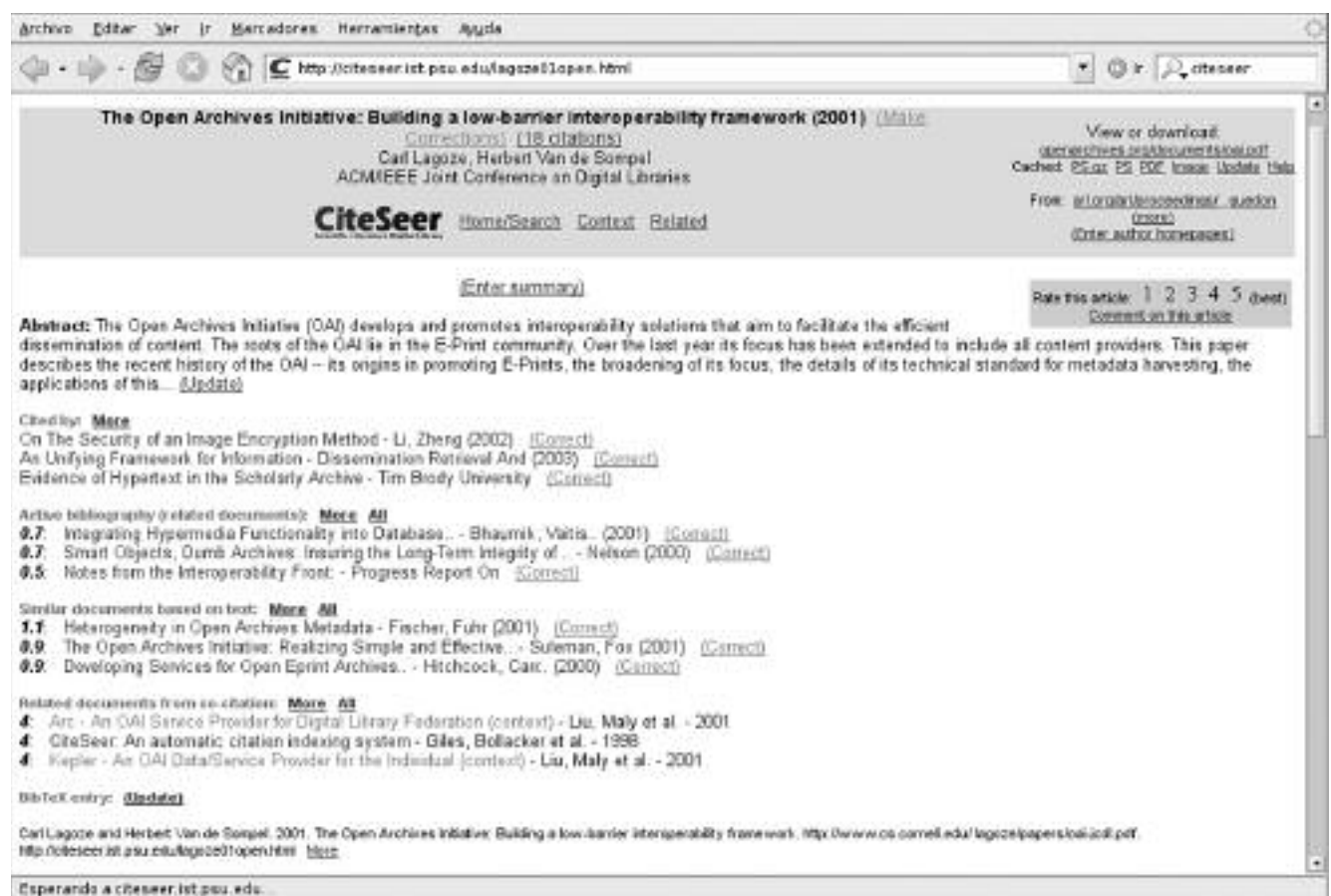


Figura 2. CiteSeer

formas en entidades dedicadas a la investigación, desarrollo e innovación (I+D+I). Véase, por ejemplo, la importancia que ha adquirido *arXiv.org* en los campos de las matemáticas, física, informática y otros relacionados. Mantenido por la *Cornell University*, es el continuador del seminal archivo de *preprints* de *Los Alamos National Laboratory*. Merced a su interoperabilidad sirven como fundamento para la realización de otros servicios de valor añadido, como la anotación e interrelación de citas e índice de impacto entre documentos de diferentes repositorios distribuidos, de lo que es ejemplo *CiteSeer*.

<http://www.arxiv.org>

<http://citeseer.ist.psu.edu>

La extensión de esta clase de fondo de información, y su integración en diferentes tipos de organizaciones, especialmente científicas y bibliotecarias, ha permitido sustituir progresivamente la expresión “digital repositories” por otra más acorde con sus objetivos y misión, como es “institutional repositories”, o repositorios institucionales. De acuerdo con **Crow** (2002) son “digital collections that capture and preserve the intellectual output of university communities”, pero resulta evidente que es posible ampliar su campo de implantación y uso a cualquier organización cuya materia de trabajo principal sea la información científico-técnica recogida en cualquier clase de documento. Este autor ha señalado, además, una serie de componentes que lo definen. En primer lugar, la existencia de tres capas de actividad (**Crow**, 2002):

— De contenido: el conjunto de prestaciones que hacen posible la incorporación y la publicación de documentos en el repositorio, así como su intercambio y acceso por parte de otros, gracias a la interoperabilidad que aportan los protocolos *OAI*.

—De archivo: el conjunto de utilidades dedicadas a la preservación y permanencia en el tiempo de los documentos, muy relacionada con la capa de contenido.

—De servicio: ofrece servicios de valor añadido para el usuario final, como el proceso de registro, la certificación de la calidad del contenido y la notificación o información de su disponibilidad.

En segundo lugar, identifica cuatro elementos básicos que definen a un repositorio institucional (**Crow**, 2002):

—Ha sido institucionalmente definido y establecido.

—Su contenido es académico y/o científico.

—Es acumulativo y perpetuo.

—Es interoperable y de acceso abierto (*open access*).

Este planteamiento de trabajo ha sido generalmente aceptado como modelo por la comunidad especializada (**Van de Sompel**, et al., 2004; **Yeates**, 2003; **Anuradha**, 2005). El estudio de las profundas implicaciones que tiene la instalación de repositorios institucionales con acceso abierto no es el objetivo de este texto, como tampoco se abordan cuestiones referidas a los nuevos modelos de publicación y comunicación científica (véanse, por ejemplo, **Lynch**, 2003; **Dill**; **Palmer**, 2005; **Velterop**, 2005). Una completa recopilación de fuentes puede encontrarse en la reciente monografía de **Bailey** (2005).

Repositorios institucionales y software libre

Corrado (2005) ha indicado cómo la conjunción de acceso libre, software libre y estándares abiertos supone un notable beneficio para los sistemas bibliotecarios, gracias al ahorro de costes, la flexibilidad de los sistemas, la independencia de proveedores privados de software propietario y la posibilidad de implementar programas de intercambio de información y de acceso a otros sistemas similares. A todo ello hay que añadir el favorable desarrollo de revistas y publicaciones electrónicas de calidad, mediante las cuales la comunidad investigadora puede comunicar los avances obtenidos en su campo de actividad.

«El software libre aporta las principales soluciones disponibles para la creación, desarrollo, mantenimiento e interoperabilidad de estos depósitos»

El software libre aporta las principales soluciones disponibles para la creación, ejecución, mantenimiento e interoperabilidad de repositorios institucionales. Las aplicaciones de software libre para la gestión de información, tanto en un sentido amplio, como desde una perspectiva estricta y canónica de las ciencias de la información y la documentación, son muy numerosas (**Tramullas**, 2005), como permite afirmar una revisión de los proyectos referenciados en *Open Source Systems for Libraries*. La recopilación de herramientas que cumplen los requisitos establecidos por la *Open Archives Initiative* (2005) se encuentra disponible en la web y, de todas ellas, interesan al objetivo de este trabajo aquéllas que son capaces de dar soporte a las tres capas señaladas por **Crow** (vid. supra). Ello excluye, por ejemplo, a las que se limitan a actuar como reco-



Figura 3. E-LIS



Figura 4. Eprints en la Universidad Complutense

piladores o *harvesters* para crear *mirrors* o copias locales de otros repositorios, como puede ser *Celestial*.
<http://www.oss4lib.org>
<http://www.openarchives.org/tools/tools.html>
<http://celestial.eprints.org>

Las aplicaciones de software libre para repositorios institucionales no responden a un modelo tecnológico que se ha tenido por genérico para este tipo de programas, y cuyas características serían la de estar soportados por pequeños grupos de desarrollo, con escaso apoyo institucional, y con una actividad discontinua, lo que puede incidir en la calidad y la sostenibilidad del software resultante. **Chawner** (2005) ha demostrado que, en el caso de los repositorios institucionales, precisamente, el modelo es diferente. Las múltiples herramientas han surgido, y se están llevando a cabo, en un entorno corporativo, gracias a proyectos de investigación que han obtenido financiación de administraciones públicas, principalmente. Por esta razón, disponen de equipos de desarrollo de software bastante estables, lo que asegura la mejora e innovación de las prestaciones del software. Las organizaciones que apoyan los proyectos son de alto nivel, e implican a universidades, centros de investigación y bibliotecas universitarias.

Alrededor de los diferentes proyectos se han ido creando comunidades de usuarios, precisamente del mismo tipo, que aseguran la sostenibilidad de los sistemas a corto y medio plazo. Otro de los factores a tener en cuenta ha sido la ausencia de software propietario capaz de realizar este tipo de tareas, de tal forma que incluso puede apreciarse la participación de iniciativa privada en algún caso, o el uso de los mismos como plataforma para el desarrollo de soluciones comerciales propietarias. Como debilidad cabe señalar la falta de coordinación entre los diferentes proyectos, a los que puede calificarse de “locales”. Sin embargo, en el caso concreto de los repositorios institucionales, la interoperabilidad que permite la adscripción a estándares, en especial a *OAI-PMH*, asegura la posibilidad de acceder al contenido de cualquiera de ellos, independientemente del software utilizado para su construcción y mantenimiento.

Como ejemplo de aplicaciones libres para repositorios institucionales, deben citarse los dos principales proyectos específicos para el ámbito de la información y la documentación. *E-LIS* (**De Robbio; Subirats**, 2005) es un fondo centralizado, soportado por *Eprints 2*, multilingüe. *Dlist* (**Coleman; Roback**, 2005), en cambio, utiliza un versión propia de *PKP Harvester*

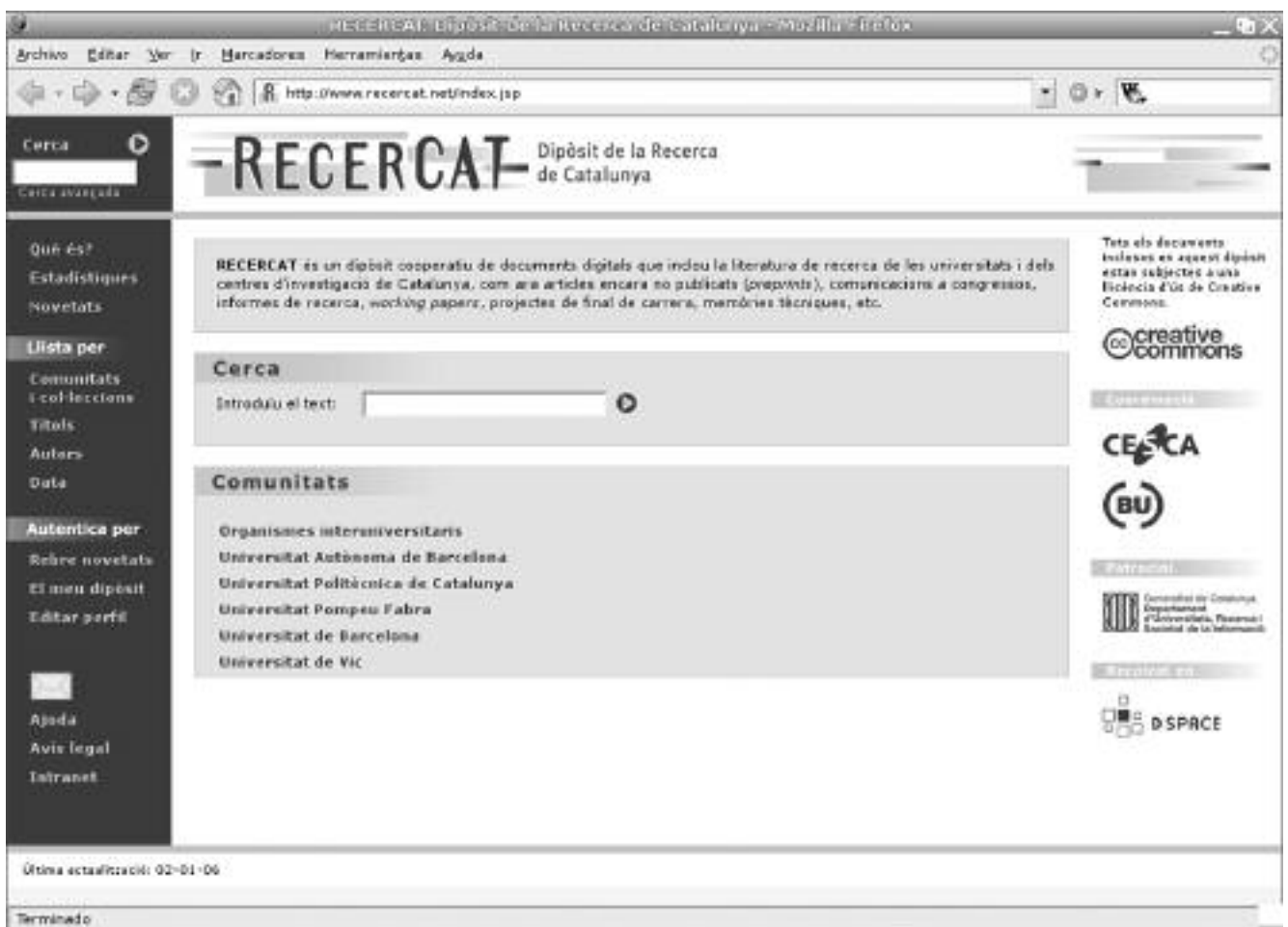


Figura 5. DSpace en Recercat

para confeccionar la colección interoperando con otros servidores *OAI*, aunque la plataforma de gestión y acceso es también *Eprints 2*.

<http://eprints.rclis.org/>

<http://dlist.sir.arizona.edu/>

En el contexto español, los diferentes proyectos, muy escasos, están siendo abordados por bibliotecas universitarias. Los servidores de referencia de las herramientas incluyen un listado de los que operan con las mismas. Cabe destacar que no se indica la existencia de ninguna instalación de *CDSware*, ni tampoco de *Fedora*. La entidad pública *Red.es* aparece como usuaria de *Fedora*, pero no ofrece un repositorio institucional de acceso público. La única instalación de *Eprints 2* operativa es la realizada por la biblioteca de la *Universidad Complutense de Madrid* (figura 4). Cuatro son las instalaciones de *DSpace* existentes: en las universidades de *La Coruña*, *Girona*, *Politécnica de Catalunya*, y *Recercat* (*Dipòsit de la Recerca de Catalunya*), proyecto interuniversitario coordinado por el *CBUC* y el *Cesca* (figura 5).

<http://www.ucm.es/eprints>

<https://dspace.udc.es/>

<http://diobma.udg.es:8080/dspace/index.jsp>

<https://e-revistas.upc.edu/>

<http://www.recercat.net/index.jsp>

Evaluación de software libre para repositorios

El proceso de selección de cualquier herramienta informática debe atender a criterios de calidad, fiabilidad y prestaciones. Si bien es sencillo formular los principios generales que deben regir la misma, la elección real, en situaciones de funcionamiento, es mucho más difícil de realizar. Existen diferentes métricas pero no se han localizado ejemplos de aplicación en la bibliografía especializada. Los estudios disponibles se centran en los resultados obtenidos en proyectos concretos, y en la comparación entre programas atendiendo al estudio de sus prestaciones. También se han publicado informes que comparan las aplicaciones, atendiendo a cuestiones de facilidad de uso.

El trabajo de referencia sobre repositorios institucionales es el publicado por **Crow** (2004) para el *Open Society Institute*. Se trata de un extenso texto descriptivo sobre las prestaciones de nueve herramientas, entre las que se incluyen *DSpace*, *Eprints* y *Fedora*, las tres más extendidas, y *CDSware*. Se estructura en tablas que comparan la presencia o ausencia de funciones que han sido consideradas necesarias por el autor. Las principales categorías que ha empleado se refieren al proceso de instalación, a la gestión de usuarios, a los procesos de incorporación de documentos al repositorio y gestión de los mismos, incluyendo el archivo, y a

las de recuperación de información. Cabe señalar que las versiones estudiadas corresponden a los años 2002–2004 y, como puede imaginarse, se han presentado sucesivas versiones desde entonces que modifican el contenido del informe y su validez en algunos parámetros. Por otra parte, algunos aspectos no son tratados con la profundidad que sería deseable. Por ejemplo, en los aspectos referidos a la integración y procesamiento de formatos de documentos para el repositorio, no se hace mención a la necesidad de instalar o utilizar, o no, programas complementarios para la extracción de contenido, o conversores, con vista a la indización de texto completo que pueda ser utilizado para la recuperación de información. **Wang, Assion y Matthaei** (2003) recopilaron un inventario de 38 aplicaciones para repositorios institucionales, meramente descriptivo, y que incluía en pie de igualdad tanto plataformas completas, como herramientas de *harvesting*.

Nixon (2003), en un artículo muy citado, detalló el proyecto *Daedalus*, desarrollado en la *University of Glasgow* en el cual se optó por combinar *Eprints 2* y *DSpace*. Sin bien se había comenzado con la primera plataforma, la aparición de *DSpace*, y su orientación a grupos y comunidades de trabajo hizo aconsejable incorporar ambas. Más que una comparación evaluativa, **Nixon** hace una exposición de las funcionalidades que interesaban al proyecto en una y otra. Sin entrar en cuestiones específicamente tecnológicas, **Nixon** concluye que “They have much in common and the choice of which, or both, or neither, will hinge on a range of local factors. It is not a question of which software is better but rather which is appropriate for the institutional services which you are building, their purpose and the content... The choice of software is only one component in a larger collection of issues for the implementation of an institutional repositories service. There is a range of policy decisions which must be made and from that will flow the decisions on assets, advocacy, access and audience».

«Existen diferentes métricas que pueden utilizarse para la selección, pero no se han localizado en la bibliografía especializada ejemplos de aplicación»

Han (2004) ha publicado el proceso de evaluación de la *University of Arizona Library* para la selección de una herramienta, atendiendo a criterios de gestión de contenido digital, entre lo que incluye preservación, metadatos, acceso y características de los sistemas,

comparando *Greenstone*, *DSpace* y *Fedora*. Para ello se adoptó un enfoque basado en el análisis de sistemas. La comparación se estructuró en dos grandes bloques, correspondientes a requerimientos funcionales y viabilidad, respectivamente. *Greenstone* resultó el más sencillo y rápido de poner en producción, pero *DSpace* le aventajaba en administración de usuarios y definición de flujos de trabajo. *Fedora* implicaba laboriosidad en la adecuación del sistema, además de mostrarse como un producto todavía inmaduro. La opción elegida fue *DSpace*. Sin embargo, debe tenerse en cuenta que este análisis se llevó a cabo en 2003, y que varias de las limitaciones han sido resueltas por nuevas versiones de las plataformas.

Kim (2005) ha comparado *Eprints* y *DSpace* usando un enfoque basado en la aplicación de técnicas de usabilidad. Este autor sólo ha evaluado las interfaces estándar de recuperación de información de ambos sistemas, mediante la asignación de tareas concretas a un grupo de usuarios, el estudio de tiempos y errores, incluyendo el estudio posterior de los niveles de satisfacción de los mismos, así como del éxito en las tareas encargadas. Al tratarse de las interfaces estándar tiene el valor de poder establecer comparaciones entre las tareas en uno y otro. Las conclusiones que se obtienen no permiten afirmar una clara superioridad de uno sobre otro, ya que ambos se muestran superiores recursivamente en una o en otra tarea, sin poderse establecer una preferencia específica por ninguno de ellos.

Otros trabajos han pretendido analizar prestaciones de los repositorios, pero resultan parciales o incompletos, tanto en herramientas como en enfoques. **Jones** (2004) ha comparado las aplicaciones de *DSpace* y de *ETD-db* para la creación de un depósito de tesis, en el contexto de la *University of Edinburgh*. Pondera las funciones en lo concerniente a la remisión de documentos (en este caso tesis y disertaciones doctorales), y al conjunto de metadatos de los mismos, prestando atención tanto a los metadatos de descripción y de identificación, como a los de gestión y preservación. El autor concluye que *DSpace* es superior en los aspectos de metadatos, al igual que en la definición de sus ciclos de vida y flujos de trabajo para los documentos. A ello une la argumentación de que es un proyecto vivo, en auge, con actualizaciones, y que es un producto muy desarrollado y fiable en los aspectos tecnológicos.

Warner (2003) comparó la adecuación al estándar *OAI*, pero lo hizo contraponiendo *Eprints* con repositorios institucionales, no con software similar, lo que limita severamente su utilidad. **Prudlo** (2005) ha confrontado recientemente *Lockss*, *Eprints* y *DSpace*. Para este autor, *Lockss* es la solución adecuada para la preservación a medio plazo de colecciones digitales de

revistas, principalmente de terceros. *Eprints* se muestra como la opción correcta para *preprints*, mientras que *DSpace* ofrece una plataforma más potente, que se ajusta mejor a las necesidades generales de un repositorio institucional.

La revisión de los estudios citados permite apreciar que la evaluación de las plataformas de software libre para repositorios digitales se lleva a cabo, casi por completo, en el marco de proyectos llevados a cabo por instituciones académicas. En consonancia, se realiza dentro de un contexto previo que establece los parámetros de evaluación, adecuándolos a las capacidades, disponibilidades y limitaciones existentes en el citado contexto. A pesar de ello, es posible afirmar que los diferentes trabajos coinciden en unas afirmaciones genéricas, que pueden resumirse en los siguientes puntos:

—*DSpace* es la solución más adecuada cuando se necesita disponer de un repositorio que va a dar soporte a diferentes tipos de documentos, y a atender a variadas comunidades de usuarios gracias a su versatilidad.

—*Eprints 2* sería la aplicación correcta cuando se necesita implementar una colección de *preprints* o revistas digitales.

—*Fedora* es una plataforma avanzada tecnológicamente, pero su implementación requiere una notable inversión en programación, ya que es más bien un *toolkit*, antes que una solución completa.

—No existe un acuerdo sobre si *Greenstone* puede ser considerado, o no, como software para repositorios institucionales.

Propuesta de modelo genérico de evaluación

Las iniciativas sobre modelos o marcos de evaluación de herramientas de software para repositorios son escasas, y además no presentan un enfoque homogéneo. Las más elaboradas se centran exclusivamente en la presencia o ausencia de prestaciones y funcionalidades que se consideran básicas. En este aspecto, lo común es ofrecer una lista de comprobación, contra la cual confrontar las opciones informáticas que se analizan. **Powell** (2005) ha presentado unas breves anotaciones, basadas en la experiencia de *Ukoln*, y, como indicaciones, señala que se deben contemplar las prestaciones para objetos digitales complejos, sistemas de metadatos, identificadores e interfaces para servicios de interoperabilidad y acceso a la información. Más completa es la propuesta de **Barton y Waters** (2004–2005), en el marco de un modelo general de repositorios institucionales. En lo que respecta a la evaluación del software, ambas autoras proponen la ela-

Software	Versión	Url	Documentación
<i>Eprints 2</i>	2.3.13.1	http://www.eprints.org/software	http://www.eprints.org/documentation/tech/php/intro.php
<i>DSpace</i>	1.3.2	http://www.dspace.org	http://dspace.org/technology/system-docs/
<i>CDSware</i>	0.7.1	http://cdsware.cern.ch	http://cdsware.cern.ch/cdsware/documentation.html
<i>Fedora</i>	2.1	http://www.fedora.info	http://www.fedora.info/documentation/

Tabla 1. Principales herramientas de software libre para repositorios institucionales

boración de una lista de comprobación de prestaciones, que debe contemplar los siguientes apartados:

- Tipo de distribución.
- Programación y adecuación propia.
- Formatos de ficheros.
- Características técnicas.
- Estándares de metadatos.
- Interoperabilidad.
- Administración del sistema.
- Configuración del sistema.
- Soporte técnico.
- Documentación técnica.
- Otros factores adicionales.
- Revisión de instalaciones ya operativas.

La segunda recopilación más exhaustiva de criterios de evaluación está contenida en el trabajo de **Crow** (2004), anteriormente citado, para el cual las categorías a considerar serían:

- Especificaciones técnicas.
- Administración del sistema y del repositorio.
- Gestión de contenidos.
- Difusión (interfaz de usuario y funcionalidad de recuperación de información).
- Archivo.
- Mantenimiento del sistema.

Estos modelos de lista de comprobación no pueden ser considerados como la única aproximación para realizar el examen. La puesta en marcha y desempeño de una repositorio institucional es una iniciativa a largo plazo, en el cual la herramienta informática de soporte es posible que permanezca o varíe, y puede ser necesario llevar a cabo migraciones para asegurar la preservación de la colección digital. Este planteamiento debe llevar a considerar el modelo de evaluación des-

de un enfoque basado en el proyecto, antes que en cuestiones meramente informáticas. En cualquier caso, previamente a la evaluación, es necesario ser conscientes de que ésta sólo puede llevarse a cabo adecuadamente en tres contextos de actividad:

—En el marco de la elaboración de un proyecto para la creación de un repositorio institucional para una organización. Evidentemente, este punto de partida supone que es necesario analizar colecciones, usuarios y flujos de trabajo (tanto de la institución y del repositorio, como de los usuarios).

—En la evaluación de prestaciones reales de la herramienta, con carga de trabajo, siempre y cuando se lleve a cabo usando una colección de documentos, contra la cual comprobar las capacidades reales y los límites de aquélla.

—Atendiendo a la disponibilidad real de soporte técnico, sea mediante la documentación existente, a través de servicios comerciales, o con comunidades de usuarios. Este criterio suele recibir poca atención en los procesos de evaluación, y en cambio resulta un factor clave en el desarrollo de proyectos con software libre.

«Debe establecerse un marco general, dentro del cual definir qué características técnicas ha de ofrecer el software»

En todos los casos el análisis debe llevarse a cabo atendiendo a la colección de documentos, como elemento nuclear del repositorio, la cual existe para ser usada por los usuarios, que son la segunda piedra de toque. Podría argumentarse que la revisión de la colección de un repositorio institucional, o del software que lo soporta, son aspectos que no requieren ni hacen necesaria la participación de los usuarios, sin embargo éste es un error que suele afectar a los procesos de estimación de recursos de información digital. Las colecciones se construyen para ser usadas por los usuarios, y la forma que éstos tienen de acceder y conocer

su contenido es mediatizada por la herramienta, la cual establece, mediante sus funcionalidades y prestaciones, lo que los usuarios pueden o no hacer con la colección. Los sistemas de metadatos son la base para la construcción de estructuras de navegación, y la posibilidad de modificar la interfaz final permite ajustar el nivel de presentación y visualización de información a las características de los usuarios, por ejemplo. La mayoría de la bibliografía consultada en este artículo precisamente ofrece evaluaciones desde un punto de vista estrictamente de profesional de la información, insistiendo en enumeraciones comparativas de prestaciones, y sólo en uno de los trabajos se ha acudido a un estudio de tareas de usuario (Kim, 2005). Contrasta esta situación con la creciente importancia de los estudios de evaluación orientados a usuarios que se están desarrollando en el campo de las bibliotecas digitales (Tramullas; Garrido, 2005).

Debe establecerse un marco general, dentro del cual definir qué características técnicas debe ofrecer el software para repositorios. Como han señalado Gonçalves, Fox, Watson y Kipp (2004) para las bibliotecas digitales, “A formal model abstracts the general characteristics and common features of a set of systems developed for similar problems, explains their structures and processes, and strengthens common practice. Furthermore, formal models for information systems can be used for the design of a real system, providing a precise specification of requirements against which the implementation can be compared for correctness». Dado que el repositorio debe dar soporte a procesos de publicación, descripción, acceso y recuperación, y preservación en una organización, debe tenderse a un modelo conceptual de referencia de actividades, que contemple como la base de las mismas la noción de colección de documentos digitales. En este sentido, cabe señalar dos documentos de trabajo que pueden ser utilizados como modelo conceptual:

—*Open archival information system, OAIS*, norma ISO 14721:2003 (ISO, 2003; Lavoie, 2004). Sin entrar en aspectos tecnológicos concretos, establece el mode-

lo que debe seguir un archivo para asegurar su correcto desempeño y la preservación de sus contenidos. Ofrece modelos de referencia para procesos y actividades, de los cuales pueden deducirse rápidamente las prestaciones a las que el software debe dar soporte.

—*NISO Framework of guidance for building good digital collections* (NISO, 2004). Esta guía complementa a la anterior en cuanto introduce, alrededor de la colección, nociones de orientación a usuario, integración de información heterogénea, reutilización de contenidos, persistencia y confianza. Además, incide en la importancia del “proyecto” como marco del desarrollo de las colecciones.

En consecuencia con los razonamientos previos, el proceso de evaluación de herramientas para repositorios institucionales debería responder a las siguientes fases:

—Descripción del proyecto: objetivos y capacidades tanto técnicas como humanas.

—Definición de la colección digital: características de los documentos, niveles de procesamiento de la información. Requerimientos técnicos.

—Identificación de las comunidades de usuarios implicadas. Análisis de objetivos, pautas de comportamiento y flujos de trabajo.

—Identificación y definición de las operaciones que se llevarán a cabo por las comunidades de usuarios sobre la colección. Conocimiento de los requerimientos técnicos necesarios para soportar los flujos.

—Política y programa de preservación digital. Requerimientos técnicos.

—Diseño y aplicación de una matriz de análisis de prestaciones. Evaluación y resultados obtenidos.

Con el planteamiento indicado se obtiene un completo listado de los requerimientos técnicos necesarios ajustados, además, a las tareas y flujos de trabajo con la que se vinculan. La presencia o ausencia de una funcionalidad se considera en el contexto en el cual es ne-



Matriz de evaluación	Publicación	Acceso a la información	Interoperabilidad	Administración
<i>Colección</i>				
<i>Objetos digitales</i>				
<i>Metadatos</i>				
<i>Usuarios</i>				

Tabla 2. Ejemplo de matriz de análisis

Funcionalidad	Área de metadatos	Otras áreas relacionadas
Búsqueda a texto completo	Integración de metadatos en la interfaz de interrogación	Interoperabilidad
Creación automática de estructuras de navegación	Uso de metadatos descriptivos para la generación de estructuras	Administración: generación de interfaces
Servicios de alerta para usuarios	Disponibilidad de metadatos	Publicación, revisión y control Administración de usuarios

Tabla 3. Ejemplo de desarrollo de la intersección «Acceso a la información/Metadatos»

cesaria, lo que permite valorar también la forma en la que se implementa. La matriz de análisis y las listas de comprobación se obtienen en el contexto general del proyecto del repositorio.

Esta propuesta es lo suficientemente flexible como para que pueda ser aplicada a cualquier organización que instale y opere un repositorio institucional. Evidentemente, exige un esfuerzo previo de definición de colecciones, usuarios y procesos administrativos e informativo-documentales, pero sólo tras el mismo podrá discernirse cuál es la herramienta que más se ajusta a las necesidades de cada caso. Las funciones vendrán definidas según los procesos identificados y modelados. Por ejemplo, la cuestión no sería si soporta metadatos Dublin Core, sino si el esquema Dublin Core es el más adecuado para el fondo en cuestión; si es necesario disponer de un servidor *OAI*; la pregunta no es sólo si ofrece la funcionalidad, sino si es posible crear una tabla de equivalencias entre el conjunto de metadatos de la colección y Dublin Core. Es posible incorporar un indicador numérico, de manera que se pueda asignar un peso específico a cada uno de ellos, pero ese peso, en consonancia con el modelo propuesto, puede ser flexible y depender del contexto de cada proyecto.

Como ejemplo, véase la propuesta de matriz de análisis simple incluida en la tabla 2. Las prestaciones técnicas se valoran en virtud de las necesidades definidas en las fases previas de estudio de tratamientos, colecciones y flujos de trabajo. Se han definido cuatro procesos básicos, cada uno de los cuales se ve participado por cuatro áreas principales. En las intersecciones es posible asignar las prestaciones y funcionalidades necesarias o asignar, a su vez, un documento de valoración específico, lo que dotaría a la matriz de dos niveles de profundidad de análisis (tabla 3). La matriz obtenida podría entonces utilizarse como lista de comprobación con la que confrontar las soluciones disponibles y las requeridas en los flujos de trabajo definidos en el proyecto.

Conclusiones

Las diferentes propuestas de evaluación de herramientas software para repositorios institucionales revi-

sadas en este artículo muestran un enfoque centrado en la valoración de la presencia o ausencia de prestaciones. Esta orientación es insuficiente cuando se trata de abordar un proyecto a medio y largo plazo, y obliga a tomar en consideración otros parámetros de valoración, que integren el programa en el contexto del proyecto en el que se va a utilizar. Para superar el enfoque parcial de los modelos existentes, se propone una matriz de análisis que incluya el proyecto y la comunidad de usuarios, considerando las prestaciones no como un fin, sino como el soporte necesario a los procesos y flujos de trabajo que los usuarios van a llevar a cabo en el repositorio.

Al igual que las bibliotecas digitales, los repositorios institucionales, además de su misión de difundir y preservar los documentos de una institución, deben orientarse a satisfacer las necesidades de información de sus usuarios. Los procesos de trabajo y de acceso y descubrimiento de información han cambiado sustancialmente en los últimos años. La operación en la Red (redes sociales, trabajo en colaboración, ubicuidad, integración de información en escritorios remotos, etc.) está modificando profundamente la manera en que se usa la información. Si bien los repositorios están proporcionando una necesaria labor, en su mayoría siguen enfoques tradicionales, y es necesario que se ajusten a las nuevas demandas de los usuarios. En consecuencia, los modelos de evaluación deben reflejar el cambio de patrones y estudiar las funcionalidades necesarias para soportar nuevas necesidades. La afirmación de **Dempsey** (2006) para las bibliotecas digitales es igualmente válida para los repositorios institucionales: "Historically, library users have adapted their workflow to the library. As the network becomes more important, libraries need to adapt their services to the network workflows of their users».

Bibliografía

Anuradha, K. T. "Design and development of institutional repositories: a case study». En: *The international information & library review*, 2005, v. 37, n. 3, 2005, pp. 169-178.

Bailey, C. W. *Open access bibliography: liberating scholarly literature with e-prints and open access journals*. Annapolis: Association for Research Libraries, 2005.

- Barton, M. R.; Waters, M.** *Creating an institutional repository: leaders workbook*. MIT Libraries, 2004–2005.
- Coleman, A.; Roback, J. R.** “Open Access Federation for Library and Information Science: dLIST and DL–Harvest”. En: *D–lib magazine*, 2005, v. 11, n. 12.
<http://www.dlib.org/dlib/december05/coleman/12coleman.html>
- Corrado, E. M.** “The importance of open access, open source, and open standards for libraries». En: *Issues in science and technology librarianship*, 2005, n. 42. Consultado en: 10–01–06.
<http://www.isrl.org/05-spring/article2.html>
- Chawner, B.** “F/OSS in the library world: an exploration». En: *Open source application spaces: fifth workshop on open source software engineering* (5–Wosse). ACM: St. Louis, 2005.
- Crow, R.** *The case for institutional repositories: a Sparc position paper*. Washington: The Scholarly Publishing & Academic Resources Coalition, 2002. Consultado en: 12–01–06.
<http://www.arl.org/sparc/IR/ir.html>
- Crow, R.** *A guide to institutional repository software*. New York: Open Society Institute, 2004.
- De Robbio, A.; Subirats, I.** “E–LIS: an international open archive towards building open digital libraries”. En: *High energy physics libraries webzine*, 2005, n. 11.
<http://library.cern.ch/HEPLW/11/papers/1/>
- Dempsey, L.** “The (digital) library environment: ten years after». En: *Ariadne*, 2006, n. 46.
<http://www.ariadne.ac.uk/issue46/dempsey/intro.html>
- Dill, E.; Palmer, K. L.** “What’s the Big IDeA? Considerations for implementing an institutional repository». En: *Library hi tech news incorporating online and cd notes*, 2005, v. 22, n. 6, pp. 11–14.
- Gonçalves, M. A.; Fox, E. A.; Watson, L. T.; Kipp, N. A.** “Streams, structures, spaces, escenarios, societies (5s): a formal model for digital libraries». En: *ACM transactions on information systems*, 2004, v. 22, n. 2, pp. 270–312.
- Han, Y.** “Digital content management: the search for a content management system». En: *Library hi tech*, 2004, v. 22, n. 4, pp. 355–365.
- ISO, Space data and information transfer systems–open archival information system–reference model*. Ginebra: International Organization for Standardization, 2003.
- Jones, R. D.** “DSpace vs. ETD–db: choosing software to manage electronic theses and dissertations». En: *Ariadne*, 2004, n. 38.
<http://www.ariadne.ac.uk/issue38/jones/intro.html>
- Kim, J.** «Finding documents in a digital institutional repository: DSpace and Eprints». En: **Grove, A.** (ed.). *Proceedings 68th Annual meeting of the American Society for Information Science and Technology (Asist)*, 2005, n. 42, 2005, Charlotte, NC (US). Consultado en: 22–01–06.
http://eprints.rclis.org/archive/00005189/01/Kim_Finding.pdf
- Koutsomitropoulos, D. A.** [et al.]. “Towards the development of a general-purpose digital repository». En: *Iceis*, 2004, n. 5, pp. 271–278.
- Lavoie, B. F.** *The open archival information system reference model: introductory guide*. Dublin: OCLC, 2004. Consultado en: 22–01–06.
http://www.dpconline.org/docs/lavoie_OAIS.pdf
- Lynch, C. A.** «Institutional repositories: essential infrastructure for scholarship in the digital age». En: *ARL bimonthly report*, 2003, n. 226.
<http://www.arl.org/newsltr/226/ir.html>
- NISO Framework Advisory Group. A framework of guidance for building good digital collections*. Bethesda: National Information Standards Organization, 2004.
- Nixon, W.** “Daedalus: initial experiences with Eprints and DSpace at the University of Glasgow”. En: *Ariadne*, 2003, n. 37.
<http://www.ariadne.ac.uk/issue37/nixon/intro.html>
- Open Archives Initiative, OAI Tools*, 2005. Consultado en: 22–11–05.
<http://www.openarchives.org/tools/tools.html>
- Peters, T. A.** «Digital repositories: individual, discipline–based, institutional, consortial, or national?». En: *The journal of academic librarianship*, 2002, v. 28, n. 6, pp. 414–417.
- Powell, A.** *Notes about possible technical criteria for evaluating institutional repository (IR) software*. Ukoln, University of Bath, 2005. Consultado en: 14–01–06.
<http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/ir-software.pdf>
- Prudlo, M.** “E–archiving: an overview of some repository management software tools». En: *Ariadne*, 2005, n. 43. Consultado en: 16–01–06.
<http://www.ariadne.ac.uk/issue43/prudlo/intro.html>
- Tramullas, J.** “Open source tools for content management”. En: *Hipertext.net*, 2005, n. 3. Consultado en: 12–11–05.
<http://www.hipertext.net/english/pag1013.htm>
- Tramullas, J.; Garrido, P.** “Los estudios de usuario en proyectos de biblioteca digital: una revisión de técnicas”. En: *Actas de las 9as Jornadas españolas de documentación Infogestión*, 2005, pp. 169–179.
- Van de Sompel, H.** [et al.]. “Rethinking scholarly communication. Building the system that scholars deserve». En: *D–lib magazine*, 2004, v. 10, n. 9. Consultado en: 06–05–05.
<http://www.dlib.org/dlib/september04/vandesompel/09vandesompel.html>
- Velterop, J. M.** *Open access publishing and scholarly societies*. New York: Open Society Institute, 2005. Consultado en: 14–12–05.
http://www.soros.org/openaccess/html/open_access_publishing_and_scholarly_societies.htm
- Wang, J. Y.; Assion, M.; Matthaeci, B.** *Inventories–open archives software tools*. Open Archives Forum, 2003. Consultado en: 12–05–04.
<http://www.oaforum.org/otherfiles/tv-tools.pdf>
- Warner, S.** “Eprints and the Open Archives Initiative”. En: *Library hi tech*, 2003, v. 21, n. 2, pp. 151–158.
- Yeates, R.** “Institutional repositories». En: *VINE*, 2003, v. 33, n. 2, pp. 96–101.
- Jesús Tramullas**, Depto. de Ciencias de la Documentación, Universidad de Zaragoza.
tramullas@unizar.es
- Piedad Garrido Picazo**, Depto. de Informática e Ingeniería

Versión online de EPI

Existe una versión electrónica de *El profesional de la información*, de uso gratuito para todos los suscriptores de la revista, que pueden acceder a través de internet a los textos completos y materiales gráficos publicados desde el año 2000.

Más información en:

<http://www.elprofesionaldelainformacion.com/contenidos.html>