



Los repositorios digitales en el ámbito universitario

Clara López Guzmán
clara@servidor.unam.mx

Dirección General de Servicios de Cómputo Académico
Universidad Nacional Autónoma de México

Francisco J. García Peñalvo
fgarcia@usal.es

Facultad de Informática y Automática
Universidad de Salamanca

Resumen

Los modelos de educación en línea y la prolifera producción de los docentes y los investigadores han sobrepasado las funciones que las bibliotecas digitales pueden proveer para la organización y el almacenamiento de los contenidos digitales. Esto a dado paso al uso de los repositorios digitales como las herramientas *ad-hoc* para la gestión de contenidos que se integren más fácilmente con diversos usos y aplicaciones. Las universidades están trabajando intensamente en el desarrollo de repositorios y la UNAM, con una complejidad particular por sus dimensiones, a través del proyecto 3R, está dando marcha a la creación de una red de repositorios que permita mayor uso y visibilidad de la producción intelectual de los miembros de la comunidad.

1. Introducción

La integración de las TIC al ámbito de la formación universitaria ha implicado cambios radicales en la forma de gestionar los contenidos producidos y utilizados en las actividades de investigación y de docencia. Los recursos de enseñanza y de investigación se producen con mayor fluidez y las posibilidades de la transmisión del conocimiento se amplían gracias a modelos flexibles de la educación a distancia, como el *e-learning* y a herramientas que se han generado facilitadas por la tecnología Web.

Hace ya casi una década, las bibliotecas digitales comenzaron a dar sus primeros pasos para almacenar los contenidos digitales, que se encontraban disponibles en la Web de forma catalogada y organizada (López, 2000). Sin embargo, los modelos de los programas de educación en línea y la prolifera producción de los investigadores, han requerido organizar los contenidos de una forma más dinámica, en la que no sea



necesario contar de expertos documentalistas para catalogar un recurso. También se hace necesario almacenar distintos tipos de contenidos, diferentes a los que han sido editados de manera formal, por una casa editora, y que ahora se producen por profesores e investigadores que desean compartirlos y hacerlos públicos. Ahora, ya no sólo se producen libros y artículos, hay otros tipos de recursos que aportan información o conocimiento significativo y que tiene un valor relevante para la formación de una comunidad universitaria.

Las actividades de gestión de la información ya no recaen forzosamente en el bibliotecario, ya que no sólo se requiere organizar y almacenar los recursos, se requiere, entre otras cosas, facilitar su acceso, reutilización (importación y exportación a otras aplicaciones) y hacerlos visibles en un ambiente competitivo y complejo como lo es Internet. Así, se ha llegado al desarrollo de herramientas que no únicamente almacenan el recurso, si no que le da una serie de servicios, a los usuarios y a las aplicaciones informáticas, para que los recursos sean más fácilmente utilizados. Estas nuevas herramientas se han llamado repositorios digitales (JISC, 2005).

Dentro de la UNAM (Universidad Nacional Autónoma de México), la problemática de la gestión de contenidos no es diferente. Por su amplio universo de actividades y la alta población de miembros, la Universidad tiene una abundante producción de contenidos y una demanda excesiva de acceso a éstos a través de los medios digitales. Por ello, se ha emprendido un proyecto para conformar una red de repositorios que facilite su comunicación, la visibilidad de los recursos y el acceso a los miembros de su comunidad.

2. Los repositorios universitarios

Universidades e instituciones de todo el mundo han dado marcha a proyectos de repositorios para la organización y difusión de sus contenidos. Según el directorio de repositorios abiertos (OpenDOAR, 2007), se encuentran registrados 853 repositorios académicos, distribuidos de la siguiente manera (tabla 1):

Area geográfica	Número de repositorios	Representación
Europa	419	49%
EU y Canada	279	33%



Centro y Sudamérica	43	5%
Asia	44	5%
Africa y Australia	68	8%
total	853	100%

Tabla 1. Repositorios académicos por área geográfica

Es notable que casi la mitad de los repositorios son europeos y que Latinoamérica (representada por Centro y Sudamérica), aportan apenas un 5%.

Del total de los 43 latinoamericanos, 24 son proyectos de universidades. México cuenta con 4 repositorios, de los cuales 3 pertenecen a universidades.

Las universidades están trabajando intensamente en la creación de repositorios, quizá porque es ahí en donde se genera la mayor cantidad de trabajo intelectual, que se desean difundir y mantener organizados.

Las herramientas de software que se utilizan para la operación de los repositorios es diversa, aunque hay una fuerte tendencia por el uso de DSPACE (MIT Libraries, 2007), un sistema de software libre para la creación de repositorios que destaca por su sencillez. El tipo de contenidos que los repositorios incorporan son también variados: libros, conferencias, objetos de aprendizaje, multimedios, patentes, tesis, publicaciones y preprints, entre otros.

De manera conceptual, estos proyectos fundamentan la estructura de sus repositorios en alguna de estas tipologías:

- Repositorios de eprints y temáticos. Contienen documentos de un mismo tema, usualmente artículos científicos.
- Repositorios de materiales académicos. No contiene sólo documentos científicos arbitrados, sino que alberga todo tipo de materiales que apoyen la enseñanza y el aprendizaje, que pueden o no corresponder al mismo tema.
- Repositorios de objetos de aprendizaje. Basan su contenido en unidades de aprendizaje, que tienen como principal objetivo transmitir un conocimiento concreto y técnicamente estos contenidos pueden ser fácilmente reutilizables en otras aplicaciones.
- Repositorios institucionales. Incluyen material académico diverso, tienden a ser organizados por una institución más que por áreas temáticas.



3. Las necesidades de una mega-universidad

La UNAM registró en 2006 más de 280,000 alumnos de posgrado, licenciatura, técnico y bachillerato (SDEU, 2007). La cantidad de información que produce y demanda esta comunidad es un problema constante y latente, tanto en el aspecto administrativo como en el académico.

En el ámbito de investigación, en 2005 se reportan más de 9,000 proyectos y una producción científica de más de 4,000 publicaciones de centros e institutos de diferentes áreas. Sin embargo esta producción no se ve reflejada en las páginas Web de la UNAM, así como tampoco en la Web de cada organización.

Un elemento de preocupación mayor, en cuanto a la demanda de contenidos, es el hecho de que las opciones de educación en línea han aumentado considerablemente y son un objetivo claro en la UNAM, como una alternativa para atender la educación de forma masiva. Así, algunos programas, tanto del bachillerato, como de las licenciaturas y los posgrados ofrecen esta alternativa para la formación, mismos que requieren contenidos educativos en línea.

Dada la amplia producción intelectual de la UNAM y la diversidad de recursos que esto genera, así como la complejidad administrativa que una universidad de grandes dimensiones implica, no es fácil adoptar un esquema tradicional para el desarrollo de un repositorio centralizado que congregue variados contenidos universitarios.

Por ello, ante la necesidad de registrar, almacenar, organizar y difundir los contenidos que se producen en la UNAM, se ha gestado una iniciativa para la creación de una red de repositorios que facilite la gestión de estos recursos para la comunidad universitaria.

Al tener requerimientos particulares con una gran complejidad por la descentralización, el volumen y la pluralidad de la información que se genera, la primera tarea se enfocó a delimitar un concepto de repositorio realista y funcional a las particularidades de la UNAM. Se propuso, entonces, que:

un repositorio universitario es un depósito de contenidos digitales informativos, educativos y de investigación, con un conjunto de servicios en línea para su difusión, uso y visibilidad, producido y administrado por la comunidad académica de la UNAM, de forma institucional o grupal.

Con esta definición se trata de dar cabida a la mayor cantidad posible de tipos de recursos que transmitan información o conocimiento y su ámbito de desarrollo no esté



ligado en particular a ninguna figura institucional, dando cabida incluso a la participación de proyectos universitarios conjuntos.

4. La iniciativa 3R

Conceptualmente se tienen claro lo que se pretende con un repositorio universitario, sin embargo, ponerlo en marcha no es simple y los esfuerzos se están llevando a cabo a través del proyecto 3R (Red de Repositorios Universitarios de Recursos Digitales)¹.

3R tiene como objetivo crear el prototipo de una red de repositorios de la UNAM, que permita mayor uso y visibilidad de la producción intelectual de los miembros de la comunidad.

Es un proyecto multidisciplinario en el que se involucran investigadores y académicos del ámbito de las ciencias y de las humanidades que discuten el tema de los repositorios en el ámbito global y perfilan un modelo de solución para una universidad de gran magnitud.

4.1 Funcionamiento

3R es una red que pretende ligar los diferentes repositorios que se produzcan en la UNAM, y ofrece servicios a dos niveles: uno local y otro global (figura 1). A nivel local se encuentran los repositorios universitarios, que son alimentados con la producción académica y científica de las dependencias de la UNAM, que se dedican a la docencia y a la investigación. A nivel global se encuentra una aplicación centralizada que ofrece una interfaz para la consulta de los repositorios universitarios locales, además de otros servicios generales.

¹ <http://www.3r.unam.mx>



Modelo para los Repositorios Universitarios

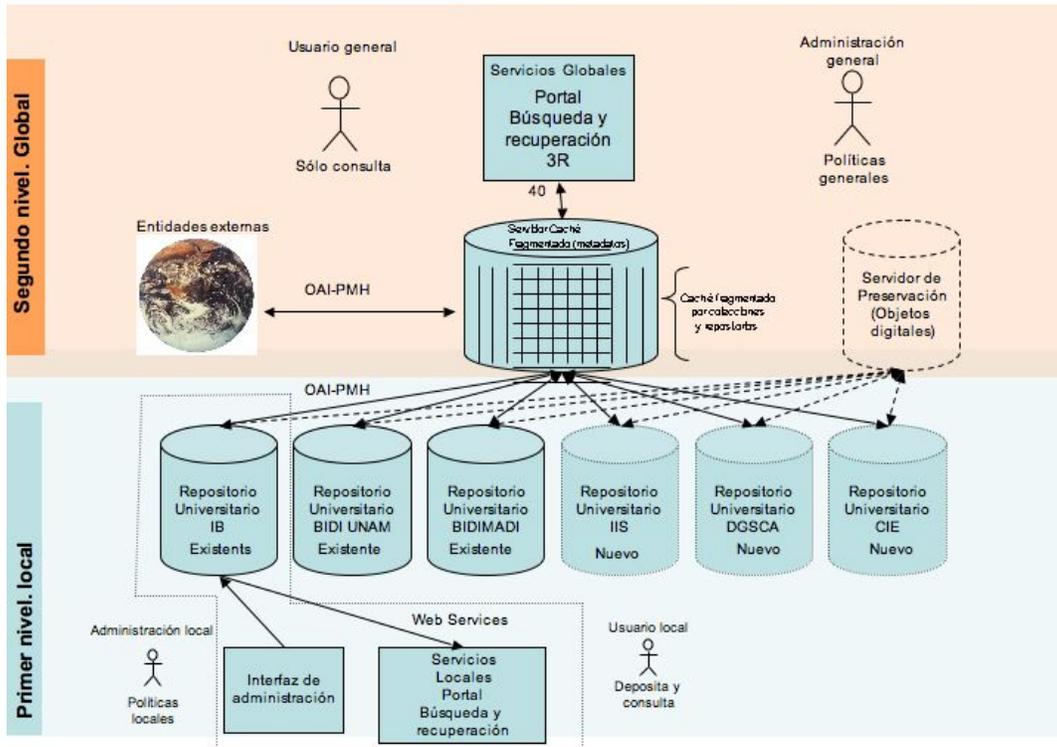


Figura 1. Modelo conceptual de 3R

El primer nivel tiene como función específica capturar, administrar y validar la información asociada a los objetos digitales depositados en cada uno de los repositorios de las Dependencias que integran el prototipo de la Red, así mismo expone los metadatos para su cosecha por el servidor caché del servidor central de 3R.

La interoperabilidad de ambos niveles esta dada por medio del protocolo para cosecha de metadatos OAI-PMH (OAI, 2004). Los repositorios locales se convierten en proveedores de datos que canalizan sólo metadatos referenciales hacia el cosechador OAI-PMH del nivel global, que los almacena y actualiza en el servidor caché, que tendrá funciones de proveedor de servicios.



4.2 Contenidos

Los contenidos se manejan como objetos digitales, ya que tendrán un tratamiento y una organización que los transforma en algo más que contenidos y se convierten en componentes de un repositorio que los administrará como objetos.

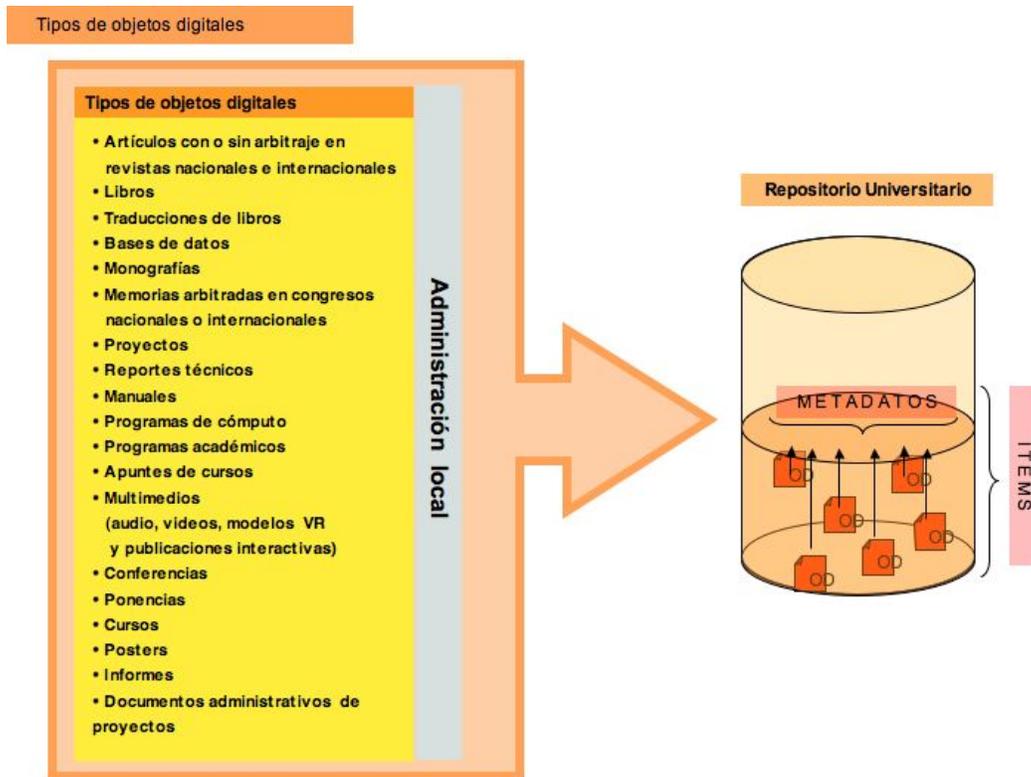


Figura 2. Tipos de objetos digitales.

La figura 2 muestra una lista de los tipos de objetos que pueden alojarse en los repositorios universitarios. Esta lista trata de reflejar las distintas formas que toma la difusión de la investigación en la UNAM, la publicación editorial, la labor docente, y documentos que ayuden al funcionamiento administrativo de las dependencias que establezcan sus repositorios. Los objetos de esta lista están contemplados independientemente de su formato; el depósito de los mismos será realizado de manera autónoma por los miembros autorizados de las dependencias, quienes deberán proveer los metadatos del objeto digital que depositan.

Para asegurar la consistencia de la bases de datos y de los objetos depositados se



debe contar con un proceso de validación (figura 3). En este proceso los metadatos de los objetos digitales depositados son examinados y validados por el administrador local para su depósito definitivo, por cualquiera de estos procedimientos:

a) Auto-depósito. El autor de un recurso accede directamente al sistema mediante una interfaz de captura; ingresa los metadatos y el objeto digital. Antes de que el recurso pueda ser publicado y estar disponible, los metadatos se someterán a un proceso de validación a cargo del administrador local del repositorio para asegurar su consistencia. Este proceso se compone de un contenedor temporal que retiene los objetos hasta su validación y que en caso de ser positiva, los objetos y sus metadatos son depositados de manera definitiva en el repositorio. En caso de que los metadatos del objeto no cumplan con alguno de los requisitos o políticas, el administrador local detendrá su incorporación y contactará con el usuario local depositante para que sean corregidos oportunamente.

b) Depósito a través de un administrador local. El autor o propietario de los contenidos entrega el objeto digital y sus metadatos al administrador local, para que éste se encargue de capturar los metadatos e ingresar el objeto digital validado al repositorio.

Los objetos que integran las colecciones, deben cumplir con estas características:

- Ser producido (autor o coautor) por miembro(s) de la Dependencia que lo registra.
- Resultar de actividades de investigación, difusión o enseñanza.
- Contar con los datos básicos para su identificación y clasificación.
- No ser efímero.
- Estar en formato digital.
- Estar completos.
- No tener problemas respecto a los derechos de propiedad intelectual y/o que estos no afecten a terceros.

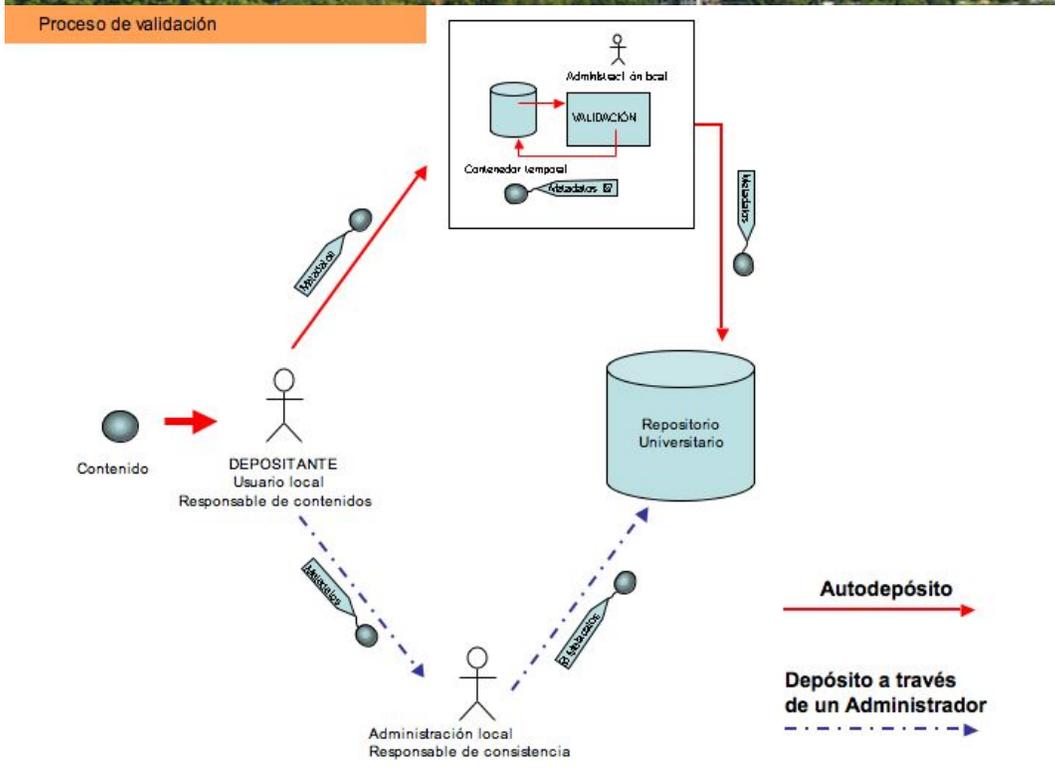


Figura 3. Proceso de validación de nuevos objetos.

4.3 Políticas

El éxito y estabilidad de los repositorios depende de manera decisiva de la implementación de políticas que contribuyan gradualmente a promover la adhesión de nuevas comunidades, incrementar el número de usuarios, estimular el autodepósito y aumentar el acervo de documentos.

Dado que esta es una red compleja en diversos aspectos, para la estandarización de contenidos, prácticas y tecnologías, es importante establecer políticas y tener una expectativa clara sobre los roles y responsabilidades de las partes involucradas.

En este sentido las políticas se están enfocando en tres aspectos generales: gestión, depósito y uso (figura 4).

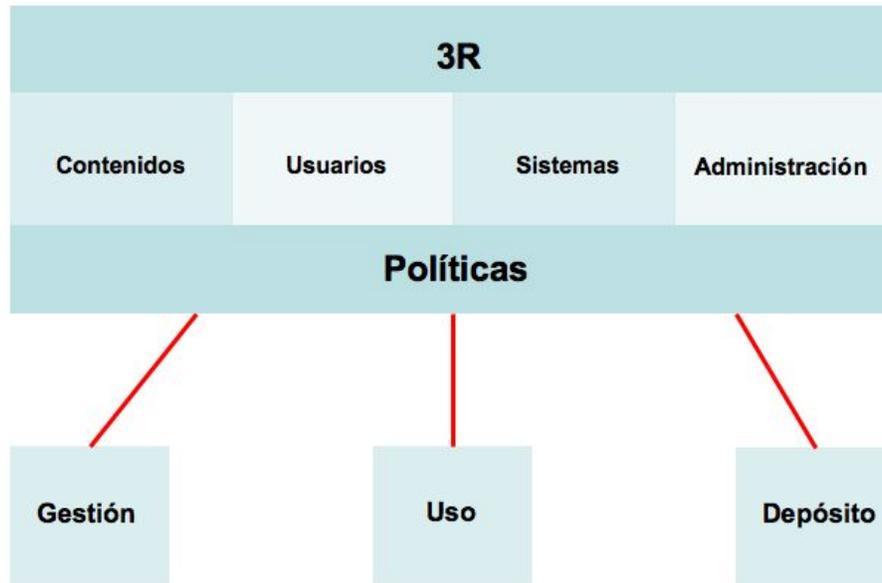


Figura 4. Políticas

Las políticas más relevantes están orientadas a la estructura y organización de los contenidos y a la conformación de colecciones, a la validación de metadatos, así como al depósito y uso de contenidos.

5. Conclusiones

Los repositorios digitales están siendo utilizados como una buena herramienta para la gestión y la preservación de los recursos producidos por las comunidades académicas y de investigación de las escuelas de educación superior, y se han permeado rápidamente en este ámbito, desplazando en cierta medida el papel que las bibliotecas digitales juegan en la gestión de recursos digitales.

La creación de un repositorio requiere del esfuerzo conjunto entre grupos de diversas especialidades, ya que se requiere contemplar organización de contenidos, políticas de operación y tecnologías a utilizar. Deseablemente debe ser una directiva institucional la que rija la implantación de un proyecto de ésta índole, especialmente en organizaciones de gran escala en donde las problemáticas son mayores y en las que



las actividades de estandarización requieren de la normalización de prácticas y de tecnologías de grupos heterogéneos.

Este proyecto se realiza con el apoyo de la Secretaría de Desarrollo Institucional de la UNAM, en el Programa Transdisciplinario de Investigación y Desarrollo para Facultades y Escuelas, Unidad de Apoyo a la Investigación en Facultades y Escuelas, a través de su Macroproyecto Universidad de la Información y la Computación, Unidad de Desarrollo de Software y Tecnologías de Información.

6. Referencias

- JISC. Joint Information System Committee (2005). Digital repositories: helping universities and colleges. Disponible en http://www.jisc.ac.uk/uploaded_documents/HE_repositories_briefing_paper_2005.pdf
- Lopez, C. (2000) Modelo para el desarrollo de bibliotecas digitales especializadas. Disponible en <http://www.bibliodgsca.unam.mx/tesis/tes7c1lg/tes7c1lg.htm>
- MIT Libraries (2007). Introducing DSpace. Disponible en <http://dspace.org/>
- OAI (2004). The Open Archives Initiative Protocol for Metadata Harvesting. Disponible en <http://www.openarchives.org/OAI/openarchivesprotocol.html>
- OpenDOAR (2007). Directory of Open Access Repositories. Disponible en <http://www.opendoar.org/>
- SDEU. Sistema Dinámico de Estadísticas Universitarias (2006) Población Escolar en la UNAM 2005-2006. Disponible en http://www.estadistica.unam.mx/2006/docencia/pob_escolar_2005-2006.html.