

# Iniciativas del Acceso Abierto a la información científica y académica en el campo Ingeniería

---

Maria Vanesa Doria<sup>1</sup>, Claudia Inés Inchaurredo<sup>1</sup>

<sup>1</sup>Dpto. Sistemas. Facultad de Tecnología y Cs. Aplicadas. Universidad Nacional de Catamarca. Maximio Victoria 55. CP 4700. San Fernando del Valle de Catamarca. Argentina.

## Resumen

Las iniciativas que propone el Acceso Abierto (AA) propician el acceso a la información científico-técnica que se produce en todo el mundo y una mayor visibilidad de la producción científica de instituciones, países y regiones subdesarrolladas. La estrategia para llevar a la práctica dichas iniciativas se sustancian en la creación y mantenimiento de repositorios institucionales y/o temáticos. En nuestros días la clave es el acceso a información libre y gratuita, partiendo además del principio de compartir conocimientos como el principal recurso de cualquier organización. El presente trabajo recoge la importancia de la iniciativa para el AA a la información científica y académica. Se analizan la estructura y funcionamiento de los repositorios temáticos y se presenta una propuesta de repositorio en Ingeniería de Software para el Dpto de Sistemas de la Facultad de Tecnología y Ciencias Aplicadas de la Universidad Nacional de Catamarca.

*Palabras clave:* Movimiento para el Acceso Abierto a la información científica, iniciativas para el acceso abierto, repositorios temáticos, ingeniería de software.

## 1. Introducción

La comunicación científica es una de las instancias fundamentales de la actividad investigadora, siendo la publicación el mecanismo formal que regula el proceso por el cual los investigadores realizan sus contribuciones al acervo de conocimientos científicos. De este modo se genera la literatura científica que va configurando la dimensión bibliográfica de las disciplinas, entendidas como sistemas de producción y diseminación de conocimientos [1].

En el siglo XX las Tecnologías de la Información y las Comunicaciones (TIC) han generado profundos cambios en el modo de gestionar el almacenamiento, la organización y el acceso a la información, tanto impresa como digital. Estos cambios tienen que ver con un

nuevo modelo de comunicación, distribución y acceso a la literatura científica dando lugar a distintas iniciativas.

## 2. Movimiento Acceso Abierto

Es una filosofía en la que los usuarios pueden leer, descargar, copiar, distribuir, imprimir, buscar, o enlazar los textos completos de los artículos científicos, y, usarlos con cualquier otro propósito legítimo, sin otras barreras financieras, legales o técnicas más que las que suponga Internet en sí misma. La única restricción que se impone sobre la distribución y reproducción es la de dar a los autores control sobre la integridad de su trabajo y el derecho a ser citado y reconocido adecuadamente [2].

A partir de esta filosofía surgieron numerosas iniciativas con la premisa de facilitar el acceso a la información científico-técnica. Las principales son:

- Iniciativa para el AA firmada en Budapest [2]: plantea que las contribuciones de acceso abierto deben satisfacer fundamentalmente dos condiciones sobre la propiedad intelectual de los documentos. Una condición es el libre acceso a través de Internet a los textos completos, permitiendo su lectura, descarga, copia, impresión y distribución, respetando las leyes de copyright existentes. Esta declaración está a favor de que sean los autores o las instituciones quienes retengan los derechos de copyright, es decir que sean los que puedan permitir el acceso abierto. La otra condición declara los principales mecanismos de publicación, por un lado la publicación en revistas de acceso abierto, denominada vía dorada o gold road y por el otro el autoarchivo de documentos en repositorios institucionales o temáticos, denominada vía verde o green road.
- Declaración de Bethesda sobre Publicación de Acceso Abierto [3]: establece que el derecho de autor será garantizado bajo licencias y que los documentos deberán depositarse en repositorios con

un procedimiento normalizado con lo que se compromete con la interoperabilidad y la preservación a largo plazo.

- Declaración de Berlín sobre Acceso Abierto al Conocimiento en Ciencias y Humanidades firmada [4]: enfatiza en la necesidad de desarrollar una política de implicación de los investigadores con el acceso abierto, tanto a través de requerimientos para que depositen una copia de todo aquello que publiquen en un repositorio o en revistas de acceso abierto.

También se pueden destacar las siguientes iniciativas:

- Declaración del Salvador sobre el AA [5]: la perspectiva de un mundo en desarrollo expresa el punto de vista de aquellos países en vías de desarrollo, que son los que más dificultades tienen para el acceso a la investigación científica.
- Declaración sobre el Acceso a Datos de la Investigación Financiada con Fondos Públicos [6]: se solicita a los países firmantes a promover el libre acceso a la documentación científica generada de la investigación financiada con fondos públicos.
- Declaración sobre el Acceso Abierto a la literatura académica y a documentos de investigación [7]: declara que la red mundial de servicios de biblioteca e información proporciona acceso a la literatura académica pasada, presente y futura de la investigación y documentación, asegura su preservación; ayuda a los usuarios en el descubrimiento y uso entre otras cosas.
- La Resolución Welcome Trust [8] apoya el acceso sin restricciones a los resultados publicados de la investigación a través de su política de acceso abierto. La Fundación Welcome Trust respalda a las mentes más brillantes en la investigación biomédica y las humanidades médicas.

En este sentido, el Ministerio de Ciencia, Tecnología e Innovación Productiva (MINCYT) de Argentina en sintonía con los antecedentes y tendencias internacionales de acceso abierto a la producción científica presenta un anteproyecto de ley denominado "Creación de Repositorios Digitales Abiertos de Ciencia y Tecnología" [9] a la Comisión de Ciencia y Tecnología de la Honorable Cámara de Diputados de la Nación en Octubre de 2010. El objetivo de la presentación de este anteproyecto es promover la equidad en el acceso a la información y a los datos científicos que son resultado de la investigación financiada en diversas formas y momentos de la producción científica por parte del Estado Nacional.

La premisa común de las iniciativas expuestas es aumentar la rapidez de la distribución de producción científico-técnica para lo cual se fomenta la construcción de repositorios institucionales o temáticos.

## 2.1 Organizaciones involucradas con el modelo AA

Siguiendo este movimiento se encuentran implicadas numerosas organizaciones y proyectos, entre las que se puede citar las más importantes de América latina y el Caribe por ser parte del contexto geográfico en el que se realiza este trabajo:

- Universidad Nacional Autónoma de México: desarrolló los siguientes repositorios: CLASE creado en 1975 se encuentran citas latinoamericanas en ciencias y humanidades, PERIÓDICA creado en 1978 es un índice de revistas latinoamericanas en ciencias (<http://www.unesco.org/new/en/communication-and-information/portals-and-platforms/goap/key-organizations/latin-america-and-the-caribbean/clase-periodica-biblat/>), Latindex creado en 1997 es un Sistema Regional de Información en Línea para la Cooperativa de Revistas Científicas de América Latina, el Caribe, España y Portugal. (<http://www.latindex.unam.mx/>).
- BIREME: Centro Latinoamericano y del Caribe de Información en Ciencias de la Salud, también conocido por su nombre original de Biblioteca Regional de Medicina (BIREME), es un centro especializado de la Organización Panamericana de la Salud / Organización Mundial de la Salud (OPS/OMS) orientado a la cooperación técnica en información científica en salud. La sede de BIREME se encuentra en Brasil, en el campus central de la Universidad Federal de Sao Paulo (UNIFESP) desde su creación en 1967 conforme acuerdo de la OPS con el Gobierno de Brasil. (<http://regional.bvsalud.org/local/Site/bireme/E/homepage.htm>.) Desarrolló los siguientes repositorios: 1.- SciELO creado en 1997 es una biblioteca electrónica que conforma una red iberoamericana de colecciones de revistas científicas en texto completo y con acceso abierto, libre y gratuito, cuenta con el financiamiento de la Fundación de Apoyo a la Investigación del Estado de Sao Paulo (FAPESP). A partir de 2002 se suma a este emprendimiento el CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico). Argentina se integra al proyecto SciELO en junio de 2005 y en octubre de 2006 obtiene la certificación para operar regularmente como sitio oficial. En Argentina este proyecto cooperativo regional forma parte de las políticas científicas del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) y se gestiona a través del Centro Argentino de Información Científica y Tecnológica (CAICYT), organismo dependiente del CONICET. Las revistas que integran la colección SciELO-Argentina tienen cobertura en todas las áreas del conocimiento y cuentan con la confiabilidad que les otorga el ser

parte del Núcleo Básico de Publicaciones Científicas Argentinas y con el rigor científico de sus artículos evaluados por pares; quienes son miembros del Comité Científico Asesor designado por el CONICET. La calidad editorial de las revistas es evaluada permanentemente en el Área de Publicaciones Científicas del CAICYT en base a los Criterios de Excelencia de LATINDEX: Nivel 1/CATALOGO Latindex. (<http://www.scielo.org.ar/scielo.php>)

2.- BVS creado en 1998 es una Biblioteca Virtual en Salud que representa una notable innovación que ha contribuido para que la Región se actualice sistemáticamente en metodologías, tecnologías, productos y servicios contemporáneos de información, conocimiento y evidencia científica en los sistemas de investigación, educación y atención a la salud, de acuerdo con las condiciones sociales, económicas y culturales de la Región. (<http://regional.bvsalud.org/>)

- Universidad Autónoma del Estado de México impulsa el proyecto Redalyc creado en 2002 es una Red de Revistas Científicas de América Latina y el Caribe, España y Portugal el objetivo es contribuir a la difusión de la actividad científica editorial que se produce en y sobre Iberoamérica. (<http://redalyc.uaemex.mx.>)
- FAPESP es una Fundación para el Apoyo a la Investigación del Estado de Sao Paulo, es una institución muy importante para la ciencia brasileña. Ha apoyado el fuerte uso de tecnologías de información digital en Sao Paulo, como la Red Académica de Sao Paulo (ANSP), que es el académico de Internet del proveedor en el estado del proyecto SciELO. (<http://www.fapesp.br/>)
- IICA es el Instituto Interamericano de Ciencias Agrícolas, este instituto en el año 1999 establece el Sistema de Información y Documentación Agrícola de América Latina y el Caribe (SIDALC) con el apoyo financiero de la Fundación Kellogg que busca satisfacer las necesidades de información de la comunidad científica, estudiantes, instituciones de desarrollo, empresarios y de los actores rurales en general. (<http://orton.catie.ac.cr/default.htm.>)
- CLACSO Consejo Latinoamericano de Ciencias Sociales es una institución internacional no-gubernamental mantiene relaciones formales de consulta con la UNESCO. Con el objeto de promover y facilitar el acceso a los resultados de las investigaciones de los Centros Miembros de CLACSO ha desarrollado en 1998 una Red de Bibliotecas Virtuales. (<http://www.biblioteca.clacso.edu.ar/somos/red-biblioteca-virtual/>)
- RedCLARA, Cooperación Latino Americana de Redes Avanzadas, es una Organización de Derecho Internacional sin fines de lucro. La visión es ser un

sistema latinoamericano de colaboración mediante redes avanzadas de telecomunicaciones para la investigación, la innovación y la educación. Formó la comunidad COLABORA, Comunidad Latinoamericana de Bibliotecas y Repositorios Digitales, responde a la sentida necesidad regional de consolidar los esfuerzos particulares existentes con el objetivo final y a largo plazo de contar con una única fuente de acceso en línea y sin restricciones a la producción académica, cultural e investigativa de América Latina, disponible para la consulta de la sociedad mundial; aumentando de esta manera la visibilidad y presencia de los autores y las instituciones latinoamericanos en la web. (<http://www.redclara.net.>)

- FLACSO, La Facultad Latinoamericana de Ciencias Sociales (FLACSO) fue creada en 1957 por iniciativa de la UNESCO. La FLACSO quedó constituida como un organismo internacional, intergubernamental, regional y autónomo integrado por los países latinoamericanos y del Caribe que adhieren al Acuerdo: Argentina, Bolivia, Brasil, Costa Rica, Cuba, Chile, Ecuador, Honduras, Guatemala, México, Nicaragua, Panamá, Paraguay, República Dominicana, Surinam y Uruguay. La Sede Académica Argentina de la FLACSO fue creada en 1974 y desde entonces, ha desarrollado una permanente labor científica en el campo de las Ciencias Sociales sustentada en las premisas del pluralismo ideológico y la excelencia, combinando rigor académico y compromiso con el entorno para contribuir al mejoramiento de la sociedad. Es una institución precursora en la formación virtual. Comenzó con la intención de distribuir de un modo más justo lo mejor de su saber académico. Hoy cuenta con una experiencia que le permite sostener que sólidas propuestas académicas y nuevas tecnologías pueden potenciarse y promover excelencia. (<http://www.flacso.org.ar/>)

Expuestas hasta aquí las iniciativas del movimiento y las organizaciones en América Latina y El Caribe que se encuentran trabajando dentro de él, este trabajo dirige ahora su foco hacia el análisis de la estructura y funcionamiento de los repositorios temáticos y se presenta una propuesta de repositorio en Ingeniería de Software para el Dpto. Sistemas de la Facultad de Tecnología y Ciencias Aplicadas de la Universidad Nacional de Catamarca.

### 3. Repositorios Temáticos

Un repositorio temático puede describirse como proveedores de servicios que recolectan datos estructurados relacionados con un tema, permitiendo posteriormente al usuario buscar a través de muchas fuentes distribuidas en todo el mundo desde un punto de acceso único [10].

Existe un consenso general de que los repositorios temáticos tienen mayor antigüedad que sus contrapartes institucionales, según Lynch [11] los repositorios temáticos abrieron camino en el ciclo de comunicación científica, esto permitió a las comunidades académicas identificar el potencial de esta alternativa de acceso a la literatura científica [10]. Algunos de los repositorios temáticos que se pueden citar son:

- arXiv: comenzó en 1991 como un repositorio para documentos de física y más tarde se amplió para incluir astronomía, matemáticas, informática, ciencia no lineal, biología cuantitativa y más recientemente estadística. (<http://arxiv.org/>)
- CiteSeer: fue desarrollado en 1997, se centra principalmente en la literatura de producciones referidas a la computadora y la ciencia de la información. Sobre la base de un análisis de los problemas encontrados por el sistema original y las necesidades de la comunidad de investigación, una nueva arquitectura y modelo de datos fue desarrollada para la "Next Generation CiteSeer," o CiteSeer<sup>x</sup>, con el fin de continuar con el legado CiteSeer en el futuro previsible. (<http://citeseerx.ist.psu.edu/>)
- RePEc: El proyecto se inició en 1997. Su precursor netec se remonta a 1993. Es un esfuerzo de colaboración de cientos de voluntarios en 57 países para mejorar la difusión de la investigación en economía. ([http://repec.org.](http://repec.org/))
- E-LIS: Establecida en 2003, es un archivo abierto internacional para Biblioteconomía y Documentación. E-LIS ha crecido para incluir un equipo de editores voluntarios de 44 países y el apoyo a 22 idiomas. El desarrollo de una red internacional LIS ha sido estimulado por la extensión del concepto de AA a las obras del LIS y facilitado por la difusión de material dentro de la comunidad LIS. Estas son algunas de las razones para el éxito de E-LIS. (<http://eprints.rclis.org>)

Las características principales de los repositorios temáticos, que los diferencian ostensiblemente de las bases de datos tradicionales, son las siguientes [12]:

- Los documentos almacenados pueden tener la forma de pre-prints (antes de pasar cualquier proceso de arbitraje) o post-prints (documentos revisados por pares y aceptados, publicados o en proceso de publicación) y pueden ser artículos de revistas, comunicaciones en congresos, capítulos para libros o cualquier otra forma de comunicación científica.
- Los documentos disponibles en los repositorios de información se encuentran a texto completo y el acceso es gratuito, libre de cualquier tipo de restricciones.

- Los propios autores son los responsables de la introducción de los documentos en el repositorio (autoarchivo).

### 3.1 El autoarchivo

Como se dijo en el apartado anterior el autoarchivo es el depósito de producciones científicas por parte del autor. El proceso consiste en que el autor se registre en un repositorio otorgándosele un espacio desde el cual incorporar los documentos. En este espacio introducirá los metadatos, descripciones estructuradas y opcionales que están disponibles de forma pública para ayudar a localizar objetos, en las áreas de telecomunicaciones e informática, es la información no relevante para el usuario final pero de suma importancia para el sistema que administra los documentos. Los metadatos son requeridos a través de formularios, y finalmente depositará el documento, que quedará en una zona no visible al usuario externo pendiente de la aceptación por parte de los gestores del repositorio, que una vez aceptado podrá visualizarse públicamente. Los metadatos son recolectados por diferentes proveedores de datos (administradores de contenido que soportan el protocolo Open Archives Initiative Protocol for Metadata Harvesting).

### 3.2. Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)

El Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) pertenece a un conjunto de acuerdos técnicos conocido como la Convención de Santa Fe [13]. OAI-PMH es un protocolo para la definición e intercambio de metadatos de un archivo. En la práctica esto significa que los archivos que cumplen con las iniciativas del AA se inscriben en un registro central y pueden ser tratados como un único y gran archivo virtual con propósitos de búsqueda. El mecanismo consiste en recolectar metadatos de dos clases de proveedores, proveedores de datos y proveedores de servicios. Los proveedores de datos son los archivos que depositan en repositorios y proporcionan los metadatos de los documentos almacenados y son recolectados por los proveedores de servicios [17]. Este protocolo ha permitido el establecimiento de proveedores de datos de diversa índole entre ellos se encuentran los repositorios temáticos. También cabe mencionar los repositorios institucionales, revistas electrónicas, bibliotecas digitales, hasta bases de datos de contenidos muy dispares, todos ellos exponiendo sus registros de metadatos conforme con el protocolo.

Los proveedores de servicio son los que recopilan los metadatos y proporcionan servicios de valor añadido, un valor puede ser una interfaz de búsqueda tipo google, filtrado, alertas temáticas, medición del uso e impacto de los documentos [17].

Los elementos del protocolo son los siguientes [24]:

- RECURSO: Es el documento o producción científica
- ITEM: Colección de propiedades sobre un RECURSO, es decir, metadatos que describen el recurso (título, autor, etc.)
- REGISTRO: Es un ITEM expresado en un formato específico formado por:
  - Encabezamiento: contiene un identificador único (identifica de forma inequívoca un artículo dentro de un repositorio), marca de fecha (fecha de creación, modificación o eliminación del registro con el fin de la recolección selectiva) cero o el conjunto de miembros de la partida con el propósito de la recolección selectiva, un opcional de estado de un atributo con un valor de borrado que indica el retiro de la disponibilidad del formato de metadatos especificado para el elemento.
  - Metadatos: son datos preliminares sobre el recurso, estos datos son título, palabras clave, resumen, fecha, entre otros que permiten la identificación y localización del recurso. El protocolo soporta múltiples formatos para expresar los metadatos. No obstante requiere que ofrezcan los registros utilizando Dublin Core no calificado, codificado en xml. El modelo de metadatos Dublin Core, es un sistema de 15 definiciones semánticas descriptivas que pretenden transmitir un significado semántico a las mismas. Este sistema fue diseñado para proporcionar información descriptiva básica sobre cualquier recurso.
  - Acerca de: es una parte opcional en la que se pueden definir declaraciones de derechos y declaraciones de procedencias.

A su vez, las características del protocolo son [17]:

- Simplicidad: los creadores buscaron una forma simple que se encuentre al alcance de cualquier usuario.
- Normalización: basados en estándares ampliamente utilizados como el protocolo http (Hiper Text Transport Protocol) para la transmisión de datos y órdenes y XML (Extended Markup Language) para la codificación de metadatos.
- Recolección: existe una entidad que pone a disposición de los interesados información bibliográfica sobre los documentos que almacena.

### 3.3. Marco Legal

Resulta importante a la hora de desarrollar un repositorio conocer los aspectos legales referidos a la propiedad intelectual, los derechos de autor, las licencias y su vinculación con la producción científica.

#### 3.3.1. Propiedad Intelectual

El concepto que brinda la Organización Mundial de la Propiedad Intelectual (OMPI) es [18]: La propiedad intelectual se refiere a las creaciones de la mente: invenciones, obras literarias y artísticas, así como símbolos, nombres e imágenes utilizadas en el comercio. La propiedad intelectual se divide en dos categorías:

- La propiedad industrial, que incluye las patentes de invenciones, las marcas, los diseños industriales y las indicaciones geográficas.
- El derecho de autor, que incluye obras literarias, tales como novelas, poemas y obras de teatro, películas, obras musicales, obras artísticas, tales como dibujos, pinturas, fotografías y esculturas, y diseños arquitectónicos. Los derechos conexos al derecho de autor incluyen los derechos de los artistas intérpretes o ejecutantes sobre sus interpretaciones o ejecuciones, los de los productores de fonogramas y los de los organismos de radiodifusión respecto de sus programas de radio y televisión.

#### 3.3.2. Derechos de autor

En lo referente a derecho de autor, la OMPI [18] expresa: El derecho de autor es el cuerpo de leyes que concede a los autores, artistas y demás creadores protección por sus creaciones literarias y artísticas, a las que generalmente se hace referencia como “obras”. El titular o titulares de una obra pueden prohibir u autorizar:

- su reproducción bajo distintas formas, incluida la publicación impresa y la grabación sonora;
- su interpretación o ejecución pública, así como su comunicación al público;
- su radiodifusión;
- su traducción a otros idiomas; y
- su adaptación.

Relacionado con el derecho de autor se tiene el Copyright que determina las modalidades de utilización de una obra.

#### 3.3.3. Licencias

Una licencia es un instrumento legal que indica que se puede y que no se puede hacer con una obra intelectual. Toda licencia presenta tres partes:

1. Condiciones a cumplir.
2. Derechos cedidos por el titular del Copyright.
3. Limitación de responsabilidad.

Feltre Oreja [19] propone el uso de licencias libres para la protección de los usos legítimos de la producción intelectual, “especialmente si se trata de

conocimiento científico o humanístico subvencionado públicamente”.

"Copyleft" es el término usado para referirse a las licencias libres. La Fundación Copyleft define copyleft como “grupo de licencias cuyo objetivo es garantizar que cada persona que recibe una copia de una obra pueda a su vez usar, modificar y redistribuir el propio trabajo y las versiones derivadas del mismo. Unas veces se permite el uso comercial de dichos trabajos y en otras ocasiones no, dependiendo qué derechos quiera ceder el autor.” Se pueden encontrar tres tipos de licencias que cumplen con la definición de la Fundación Copyleft [20]:

1. Licencias Creative Commons o CC (en español: «bienes comunes creativos») están inspiradas en la licencia GPL (General Public License) de la Free Software Foundation. Es un sistema de licencias que presenta un conjunto de soluciones frente a la producción, circulación y regulación de obras intelectuales garantizando la autoría al mismo tiempo que contribuye a la libre circulación del conocimiento. Pertenece a la organización no gubernamental sin fines de lucro de Creative Commons ([www.creativecommons.org](http://www.creativecommons.org)). El autor podrá decidir que quiere reservarse algunos derechos en lugar de todos los derechos reservados como es el caso de la licencias Copyright. Para el caso de Argentina en base a estas decisiones, para cada trabajo se debe elegir entre una de estas licencias:
  - Attribution 2.5 Argentina
  - Attribution-Noncommercial 2.5 Argentina
  - Attribution-Noncommercial-No Derivative Works 2.5 Argentina
  - Attribution-Noncommercial-Share Alike 2.5 Argentina
  - Attribution-No Derivative Works 2.5 Argentina
  - Attribution-Share Alike 2.5 Argentina
2. Licencias de Documentación Libre de GNU (FLD GNU), estas licencias son una forma de copyleft para ser usada en un manual, libro de texto u otro documento que asegure que todo el mundo tiene la libertad de copiarlo y redistribuirlo, con o sin modificaciones, de modo comercial o no comercial. La última versión es la 1.3. El texto de la licencia libre de documentación de GNU puede estar en varios formatos: HTML, texto plano, Docbook, Texinfo y LaTeX. Estos documentos no están maquetados para publicarlos por sí solos, sino que están pensados para ser incluidos en otro documento.
3. Licencias ColorIuris es un sistema mixto de autogestión y cesión de derechos de autor en línea a partir del modelo continental; a partir del artículo 27 de la Declaración Universal de Derechos Humanos, y con respeto al Convenio de Berna, los Tratados Internet de la Organización Mundial de la

Propiedad Intelectual (OMPI) de 1996, la normativa de la Unión Europea (UE) y la legislación nacional de los respectivos Estados con efectos legales para creadores de los países soportados.

### 3.4. Software

Se han desarrollado diversos programas que permiten crear repositorios y hacerlos compatibles con el protocolo OAI-PMH. La mayoría de los programas se caracterizan por ser software libre y haber sido desarrollados por implementadores de la comunidad que pertenece al movimiento AA, principalmente estos programas nacieron de proyectos de investigación en universidades, pero también en centros de investigación, bibliotecas y consorcios bibliotecarios.

Las estadísticas del 24 de marzo de 2012 del Directorio de Repositorios de AA (OpenDOAR) destacan la utilización de los siguientes software a nivel mundial: DSpace, 39,2%; Eprints, 15,3%; Digital Commons, 4,3%; DLibra 2,6%; Opus, 2,5%; Otros, 17,7 %; Desconocido, 18,4%, estos datos se consideran teniendo en cuenta 2179 repositorios. En nuestro país los software más usados son: Greenstone, 31,6%; DSpace, 21,1%; Eprints, 21,1%; SciELO, 5,3%, Desconocido, 21,1%.. Estos datos se consideran teniendo en cuenta 19 repositorios (<http://www.opendoar.org>).

A continuación se describen brevemente los software más usados en nuestro país:

- Greenstone: es un paquete de software que permite la creación y utilización de una biblioteca digital, con sus correspondientes colecciones de documentos. Se distribuye bajo licencia GNU. Su desarrollo lo lleva a cabo un equipo de investigadores de la Universidad de Waikato (Nueva Zelanda) ([www.greenstone.org](http://www.greenstone.org)). Se encuentra disponible para plataformas Windows, Linux (REdHat/Fedora/Debian/Suse/Slackware), Sun Solaris y Macintosh OS/10. Es altamente configurable pues el usuario puede modificar el diseño de la apariencia, su comportamiento o la interfaz web, Greenstone soporta múltiples lenguajes: inglés, español, francés, árabe y chino entre otros.
- DSpace: es un software de código abierto diseñado por el Massachusetts Institute of Technology (MIT) y los laboratorios de HP para gestionar repositorios de ficheros (textuales, audio, vídeo, etc.), facilitando su depósito, organizándolos en comunidades, asignándoles metadatos y permitiendo su difusión a recolectores o agregadores. Es una aplicación cliente/servidor que se gestiona vía web. Puede instalarse tanto en entornos Linux como Windows. Dispone de traducciones para seis idiomas, entre ellos el catalán y el castellano. ([www.dspace.org](http://www.dspace.org))

- Eprints: Es un software desarrollado en la Universidad de Southampton en el marco del proyecto: Open Citation Project. Se distribuye bajo licencia GNU. (www.eprints.org) Se ha ejecutado correctamente en Linux, Solaris y Mac OS X. Una versión de Microsoft Windows fue lanzado 17 de mayo 2010. Traducciones existentes son búlgaro, francés, alemán, húngaro, italiano, japonés, ruso, español y ucraniano. Comparte muchas de las características que se observan en gestión de documentos de sistemas, pero se utiliza principalmente para repositorios institucionales y revistas científicas.
- SciELO: es una herramienta desarrollada por el proyecto SciELO Brasil basada en el software Open Journal System (OJS) Sistema de Administración y publicación de revistas y documentos periódicos (Seriadadas) en Internet. Utiliza CDS/ISIS software database format es un programa conjunto para el almacenamiento de información generalizada y sistemas de recuperación desarrollado, mantenido y difundido por la UNESCO. CDS/ISIS significa Servicio de Documentación informatizada / Set Integrado de Sistemas de Información. (<http://www.scielo.org>.)

### 3.5. Repositorios temáticos en el campo de la ingeniería y tecnología

El pasado 13 de enero se celebró en la Harnack-Haus de Berlín el Simposio SOAP (Study on Open Access Publishing). En el mismo se hicieron públicos los resultados finales de una gran encuesta sobre acceso abierto realizada a investigadores de todo el mundo. Este proyecto, que transcurrió desde marzo de 2009 hasta febrero de 2011, se propuso estudiar el panorama mundial de la publicación en acceso abierto, las actitudes y creencias de los investigadores con respecto a la publicación en acceso abierto y los estímulos y las barreras para el envío de artículos para su publicación en revistas de AA. Los resultados arrojaron que para el campo de la ingeniería y tecnología aproximadamente el 85% consideran que la publicación en AA es beneficioso para el campo [14].

En relación a este tema el Ministerio de Ciencia, Tecnología e Innovación Productiva dio a conocer los resultados del relevamiento realizado sobre Repositorios Digitales en Ciencia y Tecnología en Argentina en 8ª Jornada sobre la Biblioteca Digital Universitaria JBDU2010 que se realizaron en noviembre del 2010 en la UBA (Universidad de Buenos Aires) y los resultados revelan lo siguiente[15]:

- El 20% de los repositorios digitales que existen en nuestro país dan cobertura al campo de ciencias exactas e ingeniería.
- El autoarchivo no es la modalidad preponderante de trabajo, ya que en el 74% de los casos es el personal especializado del repositorio el encargado

de procesar y archivar los documentos que entregan los autores. Asimismo, en el 61% de las instituciones, el personal especializado también se ocupa de recopilar y archivar los documentos con independencia de los autores. Este resultado se relaciona con el hecho de que hayan sido las Bibliotecas quienes tomaron a su cargo la tarea de reunir y difundir la producción de sus instituciones gestionando repositorios, como parte de sus funciones inherentes. Por otra parte, se observó un bajo nivel de conocimiento y participación de los autores en los repositorios, ya que el 60% de los repositorios informó que menos del 25% de sus docentes e investigadores entregó o realizó el depósito de documentos para su inclusión en los repositorios.

- En lo concerniente a la interoperabilidad, un poco más de la mitad de los repositorios encuestados (56%) ha implementado el protocolo OAI-PMH y la mayoría que aún no lo ha hecho reportó como motivos de ello problemas con el software utilizado o bien falta de conocimientos técnicos para hacerlo.

## 4. Propuesta de repositorio en Ingeniería de Software

En el contexto del movimiento AA y de los repositorios temáticos descritos en los párrafos anteriores se presenta en este punto una propuesta de repositorio en Ingeniería de Software para el Departamento Sistemas de la Facultad de Tecnología y Ciencias Aplicadas (FTCyA) de la Universidad Nacional de Catamarca. Para la elección del campo de aplicación del repositorio, se han tenido en cuenta las palabras de Imma Subirats Coll [16]: “La mejor manera de promover el acceso abierto es ponerlo en práctica en nuestra propia disciplina”.

El área de Ingeniería de Software es un área en la que recién se está comenzando a investigar en la FTyCA por lo que resulta interesante incursionar en este nuevo campo de trabajo a través de esta propuesta que significaría tener un espacio común en lo referido a investigaciones, transferencias, innovaciones, docencia y temas aplicados a la Ingeniería de Software.

La propuesta incluye la construcción de un prototipo que deberá cumplir con las siguientes funcionalidades: depositar, describir, preservar, facilitar el acceso. El prototipo administrará un cuerpo acotado de documentos que servirán de disparador para que se agreguen más documentos de manera voluntaria por parte de los usuarios del repositorio.

Para su construcción se realizaran las siguientes actividades: análisis documental del cuerpo de documentos que formaran parte del repositorio, análisis de requerimientos, diseño, selección de tecnología e implementación.

#### 4.1 Análisis del cuerpo de documentos

El cuerpo de documentos seleccionado para la construcción del repositorio son las publicaciones de los siguientes eventos de los años 2009-2010: Workshop de Investigadores en Ciencias de la Computación (WICC), Congreso Argentino de Ciencias de la Computación (CACIC) y Jornadas Argentinas de Informática (JAIIO) que formarán el cuerpo acotado de documentos del repositorio.

#### 4.2. Análisis de Requerimientos

Los requerimientos que se analizarán están divididos en requerimientos de usuarios y requerimientos técnicos. Los requerimientos de usuario son: ingreso de descriptores, seguridad para el contenido, soportar un proceso estándar (depósito, revisión, publicación, preservación, búsqueda, navegación). Los requerimientos técnicos son: almacenamiento de los objetos digitales, administración de bases de datos, esquemas de metadatos, organización de la información, gestión de usuarios, obtención de estadísticas, identificación persistente, transformación de los archivos, formatos admitidos.

#### 4.3. Diseño

El diseño abarcará el diseño de la arquitectura del repositorio, la modalidad de publicación, el formato de los contenidos a archivar, la organización del contenido y la definición de quienes aportan los metadatos. Para el diseño se tomarán las recomendaciones de los autores Alexander Ochoa Agüero, Athos Sánchez Mansolo, Rafael Jorge Pozo Lauzán [21].

La arquitectura del repositorio se basará en el sistema centralizado, en el cual todos los contenidos se encuentran en una base de datos, es decir es un único repositorio para todas las unidades. Los autores proponen adoptar este tipo de arquitectura para la construcción de repositorios temáticos cuando el volumen, diversidad de formato y contenido no son de gran complejidad.

En cuanto a la modalidad de publicación, describen tres tipos fundamentales:

1. Depósito directo: Persona X o Profesor una vez ingresado al repositorio como usuario autorizado carga el archivo y lo publica en el repositorio.
2. Depósito mediado por bibliotecas: Persona X o Profesor envía por correo electrónico el archivo y la información, Persona Y (Administrador) confirma origen, una vez ingresado al repositorio como usuario autorizado carga el archivo y lo publica en el repositorio.
3. Depósito semi-mediado: Persona X o Profesor una vez ingresado al repositorio como usuario autorizado carga el archivo, lo remite a una carpeta, Persona Z (Revisor) revisa los depósitos

pendientes, confirma, mejora, completa los datos necesarios, acepta o rechaza la publicación.

Para determinar el formato de los contenidos a archivar se debe tener presente los más utilizados y difundidos por los usuarios, para facilitar su consulta. Ellos son: PDF, Word, RTF, HTML, ODT, TXT, ZIP, Excel, PPT.

Para la organización del contenido se considera que la organización por índice es la más adecuada.

En la definición de quienes aportan los metadatos los autores recomiendan que siempre que se trate de depósito directo o semimediado, es suficiente con declarar el rol de administrador, revisor y autor (usuario autorizado).

#### 4.4. Selección de la tecnología

La selección de la tecnología requiere de un proceso de selección de una herramienta de software que debe centrarse principalmente en los siguientes aspectos [22]:

- Interfaz: La forma de presentación al usuario final, así como la presentación a la persona que se ocupa del procesamiento.
- Flexibilidad: Adaptación de la herramienta, según las características institucionales.
- Lenguaje: Idiomas del ambiente de procesamiento y de la interfaz de recuperación.
- Contenidos: Formato de los documentos que acepta en sus colecciones.
- Procesamiento: Facilidades para procesar los documentos para una recuperación efectiva.
- Recuperación: Formas que tiene el usuario de acceder a los documentos.
- Requerimientos de sistema: Características de las computadoras que soportarán la herramienta y de las que harán uso de las colecciones.
- Servidor Web: Requerimientos de los servidores en los que se soportará la herramienta.
- Licencia: Si es libre o privada.
- Costo: Gratis o pago.

### Conclusiones

Según expresa Weitzel [23], la principal ventaja de los repositorios radica en la centralización de la producción científica internacional por disciplinas o áreas del conocimiento. El mundo académico, es el más beneficiado con esta ventaja, puesto que le permite mantener un mayor control de la producción científica dentro de su institución; también significa un aumento en el intercambio científico técnico con investigadores del mundo entero. Existen varios ámbitos académicos de nuestro país que mediante distintos proyectos se suman a estas iniciativas [25][26][27][28][29][30].

En este contexto, el prototipo de repositorio a construir pretende introducir una filosofía de trabajo cooperativo en el Dpto de Sistemas de la FTyCA con el fin de facilitar el acceso a la información proveniente de

producciones científicas, ampliando así las vías de acceso y divulgación a información actualizada. En el mundo actual, la clave es el acceso a información libre y gratuita, partiendo además del principio de compartir conocimientos como el principal recurso de cualquier organización.

En el otro sentido, el Dpto de Sistemas podrá apropiarse de una metodología y técnica de trabajo que le permitirá hacer transferencia de sus producciones científicas en el campo de la ingeniería de software.

Por ultimo, se pretende poner a disposición de la comunidad la experiencia de una propuesta en el entorno abierto, algo factible de ser aplicado en contextos similares.

## Referencias

- [1] S. Miguel; El acceso abierto a las publicaciones de la comunidad científica argentina en el campo de la medicina. *Renata*, Vol. 1 N° 2, ISSN 2027-7415, pp 50-65, (2011)
- [2] Budapest Open Access Initiative. (2002) [online]. Disponible: <<http://www.soros.org/openaccess/read.shtml>> [Consulta: 18 de marzo de 2011].
- [3] Bethesda Statement on Open Access Publishing. (2003) [online]. Disponible: <[http://ictlogy.net/articles/bethesda\\_es.html](http://ictlogy.net/articles/bethesda_es.html)>[Consulta:1 de Noviembre de 2011].
- [4] Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. (2003) [online] <<http://oa.mpg.de/openaccessberlin/berlindeclaration.html>> [Consulta: 18 de marzo de 2011].
- [5] Declaración de Salvador sobre acceso abierto: la perspectiva del mundo en desarrollo. [online] <<http://www.bvs.org.ar/pdf/Salvador-AccesoAbierto.pdf>> [Consulta: 1 de Noviembre de 2011].
- [6] OCDE. Declaration on access to research data from public funding. 2004. [online]<<http://www.oecd.org/dataoecd/9/61/38500813.pdf>> [Consulta: 1 de Noviembre de 2011].
- [7] Declaración sobre el Acceso Abierto a la literatura académica y a documentos de investigación <[http://archive.ifla.org/V/cdoc/open-access\\_04.html](http://archive.ifla.org/V/cdoc/open-access_04.html)> [online][Consulta: 1 de Noviembre de 2011].
- [8] Resolución Welcome Trust <<http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Open-access/index.htm>>[online][Consulta: 1 de Noviembre de 2011]
- [9] PROYECTO DE LEY. Creación de Repositorios Digitales Abiertos de Ciencia y Tecnología. CIENCIA ABIERTA ARGENTINA 2010. Versión 19/10/10
- [10] PROJECT NECOBELAC. Red de Colaboración Entre Europa y Countries de América Latina y El Caribe. [online] Disponible: < <http://62.204.194.27/necobelac/>>[Consulta: 23 de Marzo de 2012]
- [11] C Lynch. Institutional repositories: essential infrastructure for scholarship in the digital age. *ARL No. 226*: 1-7 (2003). [online] Disponible: <[http://scholarship.utm.edu/21/1/Lynch\\_IRs.pdf](http://scholarship.utm.edu/21/1/Lynch_IRs.pdf)> [Consulta: 23 de Marzo de 2012]
- [12] J. Arancibia, J. Santillán Aldana, I. Subirats Colls; Iniciativas de Acceso Abierto en Ciencias de la Información y Documentación: evolución y perspectivas de E-LIS, *Revista Española de Documentación Científica*, Vol. 28, N°2, pp 221-232. (2005)
- [13] Convención de Santa Fe, Historia y desarrollo del OAI-PMH [online] Disponible:<<http://travesia.mcu.es/portaln/jspui/html/10421/1823/page2.htm>> [Consulta: 1 de Noviembre de 2011]
- [14] S. Dallmeier-Tiessen & et all, Highlights from the SOAP project survey. What Scientists Think about Open Access Publishing. [online] Disponible: <<http://arxiv.org/abs/1101.5260>> [Consulta: 23 de Marzo de 2012]
- [15] Ministerio de Ciencia, Tecnología e Innovación Productiva, “Repositorios Digitales en Ciencia y Tecnología en Argentina: resultados del relevamiento realizado por el Ministerio de Ciencia, Tecnología e Innovación Productiva”, en: 8ª Jornada sobre la biblioteca Digital universitaria JDBU 2010, Buenos Aires, Argentina, 11 y 12 de noviembre de 2010. [Online]. Disponible: <http://www.amicus.udes.edu.ar/documentos/8jornada/8jornada.html> [Consulta: 1 de Noviembre de 2011].
- [16] I. Subirats Coll, Principales factores que pueden dificultar la evolución de un archivo abierto y cómo afrontarlos, *Boletín de la SEDIC N° 42*, ISSN: 1137-0904 [online] Disponible:<[http://www.sedic.es/p\\_boletinclip42\\_confirma\\_imprimir.htm](http://www.sedic.es/p_boletinclip42_confirma_imprimir.htm)>[Consulta: 1 de Noviembre de 2011]
- [17] S. Jaroszczuk, Construcción de repositorios institucionales open source con software Greenstone, Tesis de Grado, Facultad de Humanidades y Cs. Sociales; Universidad Nacional de La Plata; (2010).
- [18] Organización Mundial de la Propiedad Intelectual, ¿Qué es La propiedad intelectual?, OMPI N° 450(S). ISBN 92-805-1157-0.
- [19] R. Feltrero Oreja, El Software Libre y la Producción y Socialización del Conocimiento en la Red: El problema de la Propiedad Intelectual. (2003)

- [20] C. Inchaurredo, M. Doria, *Licencias Libres y su relación con la Producción Científica*, Jornadas de Software Libre 2011, ISSN: 1850-2857, pp. 90 - 100
- [21] A. Ochoa Agüero, A. Sanchez Manzolo, R. Pozo Lauzán, *Repositorios De Objetos De Aprendizaje De Acceso Abierto Para La Educación De Postgrado*. República Bolivariana De Venezuela. 2010 [online] Disponible: <[http://www.rcim.sld.cu/revista\\_21/articulo\\_pdf/repositorioaprendizaje.pdf](http://www.rcim.sld.cu/revista_21/articulo_pdf/repositorioaprendizaje.pdf)> [Consulta: 23 de Marzo de 2012]
- [22] Y. Sarduy Domínguez, P. Urra Gonzalez, *Herramientas para la creación de colecciones digitales*, Revista Acimed, Vol. 14 N° 5, (2006) [online] Disponible: <[http://bvs.sld.cu/revistas/aci/vol14\\_5\\_06/aci19506.htm](http://bvs.sld.cu/revistas/aci/vol14_5_06/aci19506.htm)> [Consulta: 1 de Noviembre de 2011]
- [23] S.Weitzel, et all. "E-LIS: um repositório digital para a biblioteconomia e ciência da informação no Brasil." XV Seminário Nacional de Bibliotecas Universitárias, São Paulo, Brasil, 10-14 de noviembre, 2008).
- [24] The Open Archives Initiative Protocol for Metadata Harvesting, [online] Disponible: < <http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm>> [Consulta: 23 de Marzo de 2012]
- [25] Biblioteca Digital UNCU. [online] Disponible: <<http://bdigital.uncu.edu.ar/>> [Consulta: 18 de marzo de 2011].
- [26] Biblioteca Digital FCEN – UBA. [online] Disponible:<<http://digital.bl.fcen.uba.ar/gsd1-282/cgi-bin/library.cgi>> [Consulta: 18 de marzo de 2011].
- [27] Portal de Tesis y Disertaciones Electrónicas del Sistema de Bibliotecas y de Información (SISBI). [online] Disponible:<<http://www.sisbi.uba.ar/consultas/tesis.html>> [Consulta: 1 de Noviembre de 2011]
- [28] Repositorio Hipermedial UNR. [online] Disponible:<<http://www.unr.edu.ar/noticia/1744/repositorio-hipermedial-unr->> [Consulta: 18 de marzo de 2011].
- [29] Servicio de Difusión de la Creación Intelectual (SeDiCI) [online] Disponible:<<http://sedici.unlp.edu.ar/>> [Consulta: 18 de marzo de 2011].
- [30] Memoria Académica de La Facultad de Humanidades y Ciencias de la Educación. [online] Disponible:<<http://www.memoria.fahce.unlp.edu.ar/>> [Consulta: 1 de Noviembre de 2011]
- [31] Biblioteca Digital de Tesis y Disertaciones. [online] Disponible: <<http://tesis.uns.edu.ar/>> [Consulta: 1 de Noviembre de 2011]

*Dirección de Contacto del Autor/es:*

**Maria Vanesa Doria**  
Maximio Victoria 55 - C.P: 4700  
San Fernando del Valle de Catamarca  
Argentina  
e-mail: vanesadoria@gmail.com

**Claudia Inés Inchaurredo**  
Maximio Victoria 55 - C.P: 4700  
San Fernando del Valle de Catamarca  
Argentina  
e-mail: cinchaurredo@sistemas.frc.utn.edu.ar

---

**Maria Vanesa Doria.** Lic. en Sistemas de Información. Jefe de Trabajos Regular Dedicación Semiexclusiva. Cátedras: Ingeniería de Software III y Sistemas de Gestión Fac. de Tecnología y Cs. Aplicadas de la UNCa.

---

---

**Claudia Inés Inchaurredo.** Mg. en Ingeniería de Software. Ing. en Sistemas de Información. Profesor Titular Dedicación Semiexclusiva. Cátedra: Sistemas de Gestión. Fac. de Tecnología y Cs. Aplicadas de la UNCa.

---