

## **TÍTULO DEL PROYECTO**

Investigación y desarrollo en repositorios institucionales: aplicaciones y experiencias en universidades de la región bonaerense

## **UNIVERSIDADES PARTICIPANTES**

Universidad Nacional de La Plata  
Universidad Nacional del Sur  
Universidad Nacional de Luján  
Universidad Nacional del Centro de la Provincia de Buenos Aires  
Universidad Nacional de Mar del Plata  
Universidad Tecnológica Nacional – Regional Bahía Blanca

## **OBJETIVOS GENERALES**

En los últimos años, el movimiento Acceso Abierto (AA), que promueve el acceso libre y sin restricciones a los trabajos publicados por la comunidad científica, ha tomado mucha fuerza entre las instituciones académicas y científicas. El AA se materializa a partir de dos vías: publicar en revistas de acceso abierto (vía dorada), o bien depositar una copia de la producción intelectual en un repositorio institucional o temático de acceso abierto (vía verde).

El Repositorio Institucional (RI) se entiende como un sistema de información que reúne, preserva, divulga y da acceso a la producción intelectual y académica de las comunidades universitarias. En la actualidad el RI se constituye en una herramienta clave de la política científica y académica de la universidad. Por otro lado, el acceso al texto completo de los objetos de aprendizaje digitales hace que el repositorio se constituya en una pieza de apoyo fundamental para la enseñanza y la investigación, a la vez que multiplica la visibilidad institucional en la comunidad internacional (Bustos, A., 2007).

Numerosas universidades prestigiosas de todo el mundo vienen apoyando de manera comprometida la creación de repositorios institucionales, mientras desde la literatura se vienen ensayando diferentes propuestas de desarrollo de los mismos a partir de modelos de referencia para bibliotecas digitales, como el de las 5S elaborado por Fox 1999 (Witten, I. y Bainbridge, 2003), y otros proyectos de avanzada como DELOS (Network of Excellence on Digital Libraries) (Gonçalves, M. A., Fox, E. A., Watson, L. T., 2008). Estas iniciativas, junto con los nuevos conceptos y herramientas tecnológicas configuran un campo de estudio relevante para todas las actividades de aprendizaje y de investigación que se realizan en las universidades.

En nuestro país y en relación directa con el AA, Carolina De Volder (2008) realizó un relevamiento de los repositorios institucionales y determinó que existían 15, de los cuales sólo 4 pertenecen a Universidades Nacionales de la región bonaerense: SEDICI y Memoria Académica de la UNLP, Cartapaccio de UNICEN y Biblioteca Digital de Tesis, la edición de revistas científicas digitales y los objetos multimedia (principalmente videos y audios) de la UNS, a lo que se sumó recientemente Nulan de la UNMDP y otras iniciativas de digitalización de tesis, revistas científicas y libros de producción local que todavía son incipientes. Por otra parte, la creación del Sistema Nacional de Repositorios Digitales en Ciencia y Tecnología, iniciativa impulsada por el MINCYT está promoviendo activamente la generación de RIs en las UUNN, como forma de garantizar el libre acceso a esta producción. Paralelamente, es sabido que los docentes e investigadores tienen la obligación de presentar informes de su producción intelectual en formato digital,

junto con la documentación probatoria, a través de distintos sistemas de gestión internos (SIGEVA, Winsip, etc.) para las actividades de evaluación de la actividad científica; aunque estos documentos no se encuentran accesibles ni disponibles para su uso público. Al mismo tiempo, también existe una extensa producción de recursos y objetos de aprendizaje que se generan como parte de los materiales didácticos de los cursos académicos, a los que tampoco es posible acceder de forma pública y libre.

En este contexto, el proyecto que presentamos se propone **contribuir al desarrollo y mantenimiento de RIs para el archivo, preservación, distribución y acceso de la producción intelectual desarrollada por las universidades de la región bonaerense, con el objetivo de promover el acceso abierto al conocimiento y la implementación de políticas institucionales que favorezcan y sustenten su desarrollo**. De lo anterior se deduce que estas tareas implican, además, relevar la producción existente en los sistemas internos con miras a su posible reutilización y adaptación para formar parte de un RI, mejorar las herramientas de recuperación de información, así como como también estudiar y establecer indicadores de uso de los recursos allí disponibles para permitir la medición del impacto de los RI como forma de difusión de la producción intelectual de las instituciones.

Como resultados del proyecto se espera lograr el crecimiento y la consolidación de las iniciativas de RIs existentes en la región, así como el desarrollo de nuevos repositorios en las instituciones que aún no disponen de ello, con recursos humanos formados para su mantenimiento. Asimismo, se espera contribuir en el análisis y definición de metodologías que permitan importar de forma semiautomática a los RIs los objetos digitales y metadatos existentes en otros sistemas y contar con herramientas de búsqueda adaptadas al contexto de los RIs. Finalmente, se espera contribuir al estudio y aplicación de indicadores que permitan medir el impacto de los RIs en la evaluación de la actividad académico-científica.

## **OBJETIVOS ESPECÍFICOS**

En las UUNN existen metadatos y objetos digitales de la producción intelectual de las comunidades académicas que podrían formar parte de los RIs, y cuya reutilización permitiría agilizar enormemente las tareas de procesamiento y registro de los datos, reducir la duplicación de esfuerzos y acelerar el desarrollo de los mismos, incorporando de manera automática o semiautomática información procedente de fuentes ya existentes, tales como el Winsip (aplicación de la SPU que se utiliza con fines de evaluación de resultados del Programa de Incentivos a la Investigación); el SIGEVA (Sistema Integral de Gestión Evaluación) desarrollado por el Consejo Nacional de Investigaciones Científicas y Técnicas –CONICET- y actualmente adoptado por algunas UUNN), así como otros gestores de contenido utilizados internamente en las distintas universidades. Por otro lado, distintos repositorios ya en funcionamiento proveen información muy rica para fortalecer, impulsar y facilitar dentro de cada institución el uso apropiado de las tecnologías para el aprendizaje. Un caso de ejemplo es el proyecto SciELO Argentina. En esta línea, uno de los objetivos específicos que nos proponemos con este proyecto es: **1- Estudiar la factibilidad de reusabilidad de datos y objetos digitales existentes en los diferentes sistemas de información de las universidades, para su integración automática o semiautomática a los RIs garantizando su accesibilidad e interoperabilidad.**

Por otra parte, si bien existen en las UUNN de la región bonaerense algunas iniciativas de RIs en marcha, algunas de las UUNN aún no cuentan con un repositorio, o bien poseen desarrollos incipientes de registro y digitalización de tipos diversos de documentos que aún no se han plasmado como RIs. Teniendo en cuenta además que en general las UUNN tienen una estructura administrativa descentralizada en Facultades, muchas veces las iniciativas se desarrollan solo en algunas de ellas, siendo necesario hacer extensivo los esfuerzos a las demás. En este sentido, la iniciativa Memoria Académica de la FaHCE-UNLP, que se ha posicionado en los últimos años como una experiencia exitosa de RI a nivel nacional e internacional, podría ser transferida a otras instituciones interesadas a través de la instalación del sistema informático, la capacitación de recursos humanos especializados y el apoyo técnico para su implementación. Por otra parte, la integración de los desarrollos parciales como los existentes en la UNS prevé resolver aspectos relacionados a la problemática de la recuperación de la información, experimentando el uso de herramientas modernas que ayuden a minimizar las inconsistencias propias de la mayoría de los sistemas de las bibliotecas. Se espera que el conjunto de universidades de la región pueden colaborar en esta iniciativa poniendo sus propios repositorios a disposición para realizar pruebas de integración y recibiendo una capacitación introductoria en el uso de las tecnologías involucradas, orientadas a la indización y recuperación de información en contexto, con resultados teóricos y empíricos que respalden la utilidad de los métodos propuestos y su integración como herramientas de búsqueda de los RIs. En esta dirección nos proponemos: **2- Impulsar el desarrollo de repositorios institucionales en las UUNN de la región y formar recursos humanos especializados para promover la cultura del acceso abierto en la comunidad académica y científica de las UUNN de la región.**

Los repositorios facilitan y amplían la difusión del conocimiento, ya que los documentos son accesibles de forma libre y gratuita. En este sentido, la medición del uso de los objetos digitales contenidos en los RiS aporta información sobre su visibilidad y contribuye a su vez a la generación de políticas a favor del AA. El apoyo telemático al proceso de enseñanza/aprendizaje plantea varios problemas para su implantación en situaciones reales, y en particular relacionados con la utilización y reutilización de objetos de aprendizaje disponibles. En esta línea se propone realizar una experiencia piloto de aprendizaje en gestión de la información a través de un entorno virtual que contemple el uso de contenidos de los RIs existentes en Argentina, y que aporte evidencia inicial sobre los desafíos encontrados. Para ello es importante contar con herramientas que generen indicadores de uso interoperables entre repositorios y compatibles con los utilizados por el sistema editorial de revistas científicas. En esta línea, otro de los objetivos que nos proponemos con este proyecto es: **3- Estudiar y analizar métricas de uso de objetos digitales en Ris basados en estándares internacionales para diseñar un prototipo o bien adaptar alguna herramienta preexistente que permita contar con estadísticas de uso de los OD disponibles en RIs.**

## **RELEVANCIA DEL PROBLEMA**

Las universidades son instituciones de generación y transferencia de conocimientos por excelencia. Producen grandes volúmenes de información académica y científica que se registra en diferentes medios y se difunde a través de distintos canales formales e informales. Las bibliotecas siempre han tenido un rol central en la preservación y difusión de los conocimientos; sin embargo, no siempre estas instituciones han logrado reunir en su acervo la producción intelectual desarrollada en la propia institución. Por otra parte, la diversidad de sistemas

informáticos y estructuras de datos en las que se registra la información dentro de las instituciones dificulta su acceso y en consecuencia su visibilidad.

Por otro lado, desde hace décadas ha habido un notable incremento en los precios de las revistas científicas, hecho que ha llevado a varias organizaciones a cancelar sus suscripciones al no poder hacer frente a los costes que implica mantenerlas. Así surge una paradoja: los investigadores generan conocimientos en instituciones académicas y científicas -muchas de las cuales son financiadas con fondos públicos- y publican en revistas científicas que luego son inaccesibles tanto para ellos como para las instituciones en las que ha tenido lugar la producción de esos conocimientos. Como reacción a esto, en distintos ámbitos fueron surgiendo iniciativas que hoy conocemos como movimiento de acceso abierto (AA) cuyo principio básico es permitir a los usuarios la lectura, descarga, copia, distribución, impresión, búsqueda o enlace a los textos completos de los artículos, sin otras barreras económicas, legales o técnicas que las que suponga Internet en sí misma (BOAI, 2001).

En la Declaración del BOAI (Budapest Open Access Initiative) se establecen dos rutas para alcanzar el AA: la ruta dorada, o publicación de artículos en revistas de acceso abierto, y la ruta verde que consiste en el autoarchivo o depósito de los documentos en repositorios institucionales o temáticos, ya sea antes (pre-print) o después (post-print) de su publicación. Este movimiento ha tenido gran repercusión a nivel mundial. Actualmente, el Directory of Open Access Journals, DOAJ (<http://www.doaj.org/>), repertorio de referencia internacional de la ruta dorada reúne 5.328 revistas. Por otra parte, algunas iniciativas como Sherpa/Romeo ([www.sherpa.ac.uk/romeo.php](http://www.sherpa.ac.uk/romeo.php)) tienen por objeto reunir información sobre las políticas editoriales de las revistas y los derechos de copyright de los editores con el fin de fomentar la vía verde.

Harnad (2004) argumenta que la ruta verde es la única opción para llegar al 100% del AA en un futuro próximo, ya que no requiere la reestructuración total del sistema de edición científica e invita a las instituciones a crear repositorios de acceso abierto. También propone incentivos para los autores que depositen sus resultados de investigación de cara a promociones, financiación, etc., en vista de los buenos resultados de impacto obtenidos por la ruta verde. Así, los autores pueden continuar publicando sus artículos en las revistas preferidas por ellos (aunque sean de suscripción), y simultáneamente ofrecer el acceso abierto a sus trabajos mediante la versión depositada en los repositorios.

En este sentido, un reciente estudio acerca de la influencia del movimiento AA en la industria editorial a nivel mundial revela que la ruta verde es la que más se ha desarrollado, especialmente en Europa y América del Norte, y que muchos de los editores de las revistas más prestigiosas del mundo han adoptado políticas a favor del autoarchivo en repositorios (Miguel y otros, 2010 en prensa). El estudio muestra también que en los países de América Latina y el Caribe la ruta verde tiene una muy escasa presencia, poniendo en clara evidencia la necesidad de crear repositorios y de impulsar políticas a favor del autoarchivo.

Por otro lado, como anticipamos precedentemente, además de la producción científica existe una amplia variedad de objetos digitales de diversos tipos documentales que se producen en las universidades, que no sólo tienen una difusión limitada sino que incluso, en muchos casos, no están debidamente registrados y almacenados con fines de preservación y acceso.

En este contexto, y teniendo en cuenta los antecedentes locales sobre la situación de los repositorios en las UUNN de la región bonaerense (DE VOLDER, 2008), y en concordancia con las políticas del MINCYT en torno a la creación del

Sistema Nacional de Repositorios Digitales en Ciencia y Tecnología, este proyecto contribuirá al desarrollo de RIs mediante cuatro principales acciones:

- la implementación de repositorios en las instituciones interesadas que no hayan podido concretar aún la creación de RIs;
- la formación de recursos humanos especializados para garantizar su desarrollo y mantenimiento;
- la búsqueda de estrategias y el desarrollo y puesta a punto de herramientas que favorezcan la reutilización de datos y de objetos de interés para los repositorios, y la recuperación de información en contexto;
- la propuesta de métricas de uso que aporten información sobre el impacto de los mismos en la comunidad académica y científica nacional e internacional.

### **Referencias bibliográficas**

Budapest Open Access Initiative. (2001). Budapest: Open Society Institute, 2001 [consultado 26-06-2010]. Disponible en:  
<http://www.soros.org/openaccess/read.shtml>

Bustos, A. (Coord.). (2007). Directrices para la creación de repositorios institucionales en universidades y organizaciones de educación superior. Valparaíso, Chile: Pontificia Universidad Católica de Valparaíso. Ediciones Universitarias de Valparaíso, 2007. [consultado 02-07-2010] Disponible en:  
[http://infolac.ucol.mx/observatorio/Directrices\\_RI\\_Spanish.pdf](http://infolac.ucol.mx/observatorio/Directrices_RI_Spanish.pdf)

DELOS (Network of Excellence in Digital Libraries) [Consultado el 7-9-2010]. Disponible en:  
[http://www.delos.info/files/pdf/ReferenceModel/DELOS\\_DLReferenceModel\\_0.98.pdf](http://www.delos.info/files/pdf/ReferenceModel/DELOS_DLReferenceModel_0.98.pdf)

De Volder, C. (2008). Los repositorios de acceso abierto en Argentina: situación actual. Información, cultura y sociedad, Nro. 19, p.79-98. ISSN 1851-1740. [Consultado 17/08/2009] Disponible en:  
[http://www.scielo.org.ar/scielo.php?script=sci\\_arttext&pid=S1851-17402008000200005&lng=es&nrm=iso](http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S1851-17402008000200005&lng=es&nrm=iso).

Fox, E. A. (1999) The 5S Framework for Digital Libraries and Two Case Studies: NDLTD and CSTC In Proceedings of NIT99, Taipei, Taiwan, Aug. 1999.

Gonçalves, M. A., Fox, E. A., Watson, L. T., (2008) Towards a digital library theory: a formal digital library ontology. Int. J. on Digital Libraries, 8(2): 91-114

Harnad, S. et al. (2004). The access/impact problem and the green and gold roads to open access. Serials Review, 30(4):.310-314.

Miguel, S. et al. (2010) Acceso abierto y Scopus: una nueva aproximación a la visibilidad científica desde el punto de vista del acceso. En prensa.

Witten I. y Bainbridge, D. (2003). How to build a Digital Library, The Morgan Kaufmann Series in Multimedia Information and Systems, Fox, E. (Ed.) Morgan Kaufmann Publishers, Amsterdam.

## **RESULTADOS PRELIMINARES Y APORTES DEL GRUPO AL ESTUDIO DEL PROBLEMA EN CUESTIÓN**

Uno de los antecedentes de investigación de mayor relevancia para los objetivos de este proyecto es la participación del investigador responsable y de los integrantes de la FaHCE-UNLP en el proyecto "**El acceso abierto al conocimiento científico en Argentina. Estado de la cuestión y principales tendencias**", acreditado para el período 2010-2011 en el marco del Programa de Incentivos a docentes e investigadores. Las actividades y resultados de este proyecto son un valioso aporte de conocimiento que enriquecerá al presente.

Otro antecedente de pertinencia es la participación de algunos de los integrantes del grupo en el desarrollo del repositorio institucional "**Memoria Académica de la Facultad de Humanidades y Ciencias de la Educación de la Universidad Nacional de La Plata**" (<http://www.memoria.fahce.unlp.edu.ar>) creado en 2006. No sólo constituye una experiencia exitosa en RI, sino que además ha dado lugar al desarrollo de cursos de formación. Uno de los cursos actualmente en marcha se titula "La creación de repositorios institucionales en el contexto de las Bibliotecas de la UNLP", que se espera replicar en el marco del presente proyecto para otras bibliotecas de las universidades de la región interesadas.

Asimismo, los miembros de este grupo vienen produciendo publicaciones, presentaciones a congresos y documentos técnicos sobre el tema específico del acceso abierto y de los RIs. La siguiente lista es solo un ejemplo de algunos de los trabajos más recientes.

Miguel, S. (2010). Acceso abierto y Scopus: una nueva aproximación a la visibilidad científica desde el punto de vista del acceso. En prensa.

Pené, M. (2010). Nuevas estrategias para la gestión de colecciones en bibliotecas universitarias a partir del advenimiento de los repositorios institucionales. Ponencia presentada al Congreso Internacional de Información "INFO 2010", La Habana, Cuba, abril de 2010.

Fushimi, M. (2010) La biblioteca como espacio de construcción y difusión de la producción científica de las instituciones académicas. Ponencia a presentarse en el Encuentro Internacional de Editores Científicos a realizarse en Buenos Aires en noviembre 2010.

Fushimi, M., Banzato, G. (2010) Las políticas de acceso abierto en las universidades estatales argentinas: un análisis a través de la web. Trabajo en elaboración.

Pichinini, M. (2010) Los repositorios institucionales como vía de acceso abierto al conocimiento. En: Unesco. Greenstone: un software libre de código abierto para la construcción de bibliotecas virtuales. Disponible en: [http://portal.unesco.org/ci/en/ev.php-URL\\_ID=30700&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201.html](http://portal.unesco.org/ci/en/ev.php-URL_ID=30700&URL_DO=DO_TOPIC&URL_SECTION=201.html)

Fushimi, M. (2009). Memoria Académica, una experiencia de acceso abierto en el ámbito universitario. En 7a Jornada sobre la Biblioteca Digital Universitaria: La

biblioteca universitaria en la web. Disponible en:  
<http://www.uca.edu.ar/uca/common/grupo27/files/3.pdf>

Pichinini, M. (2009). Implementación del protocolo para cosecha de metadatos OAI-PMH 2.0. En el Panel Greenstone: software open source para biblioteca digitales. En 7a Jornada sobre la Biblioteca Digital Universitaria: La biblioteca universitaria en la web. Disponible en:  
<http://www.amicus.udes.edu.ar/documentos/7jornada/documentos/pdf/PONENCIA%20PICHININI.pdf>

Fushimi, M., Mallo, J., y Pichinini, M. (2006). Memoria académica y científica: el rol de la biblioteca universitaria en la preservación y difusión del conocimiento generado en las universidades. En Actas de las 4ª Jornadas de Sociología de la Universidad Nacional de La Plata: La Argentina de la crisis. (CD-ROM). La Plata: UNLP. FAHCE. Departamento de Sociología. ISSN 1850-2458. También publicado en E-Lis: <http://eprints.rclis.org/archive/00010464/>

Fushimi, M., Mallo, J., y Pichinini, M. (2006). Memoria académica y científica: el rol de la biblioteca universitaria en la preservación y difusión del conocimiento generado en las universidades. En Actas de las 4ª Jornadas de Sociología de la Universidad Nacional de La Plata: La Argentina de la crisis. (CD-ROM). La Plata: UNLP. FAHCE. Departamento de Sociología. ISSN 1850-2458. También publicado en E-Lis: <http://eprints.rclis.org/archive/00010464/>

Pichinini, M., Pené, M. (2008). Guía para la digitalización de documentos impresos y preservación de las imágenes digitales. La Plata: Biblioteca Prof. Guillermo Obiols, Facultad de Humanidades y Ciencias de la Educación, Universidad Nacional de La Plata. Disponible en:  
<http://www.memoria.fahce.unlp.edu.ar/memoria/descargables/guia-para-la-digitalizacion-de-documentos-impresos>

Por otra parte, como ya se ha mencionado en la sección primera, en otras universidades de la región se iniciaron proyectos de repositorios: Cartapaccio de UNICEN; Biblioteca Digital de Tesis y edición de revistas científicas digitales y de objetos multimedia de la UNS; Nülan de la UNMDP. Estas iniciativas muestran el interés y esfuerzo que vienen realizando las bibliotecas de las UUNN en la generación de RIs; esfuerzos que este proyecto pretende fortalecer a través de la generación de conocimiento, la transferencia de experiencias y la formación de recursos humanos.

Otra de las fortalezas del equipo de trabajo que integra este proyecto se halla en el grupo de investigación de la UNS, cuya línea de estudios en recuperación de información contribuirá al desarrollo de herramientas que faciliten la búsqueda en repositorios desarrollados en diferentes sistemas. Los importantes avances de este grupo giran en torno a:

**Búsqueda contextualizada.** En este sentido, los sistemas desarrollados intentan mejorar el vocabulario con el que se especifican los requisitos de información así como las descripciones de los recursos indexados. Los resultados obtenidos hasta el momento han sido publicados en congresos y revistas internacionales.

**Búsqueda distribuida:** En colaboración con la School of Informatics de Indiana University, ha desarrollado métodos y realizado simulaciones para el estudio de las comunidades semánticas emergentes en sistemas de recuperación de información

distribuidos. Una topología es considerada buena si favorece la eficiencia, haciendo posibles que las consultas alcancen el destino apropiado en pocos pasos, sin imponer una gran sobrecarga sobre el ancho de banda del sistema.

**Soporte inteligente para la captura, refinamiento y distribución de conocimiento:** En colaboración con Indiana University y el IHMC (Institute for Human and Machine Cognition, Pensacola, USA) han diseñado e implementado mecanismos incrementales de búsqueda capaces de identificar recursos electrónicos relevantes para el contexto del usuario. Su investigación fue desarrollada en el marco del proyecto "Integrated Intelligent Support for Knowledge Capture, Refinement and Sharing" financiado por la NASA (NCC 2-1216). En dicho proyecto se han desarrollado herramientas de asistencia al usuario capaces de identificar y sugerir recursos de potencial interés facilitando la exploración de espacios de información relevantes al contexto bajo análisis.

**Herramientas inteligentes para el mantenimiento de RI y para su acceso contextualizado.** Los investigadores Luis Herrera y Ana Maguitman han colaborado en la dirección de estudiantes avanzados de la carrera de Licenciatura en Ciencias de la Computación de la UNS, quienes han desarrollado herramientas inteligentes para la interacción con recursos almacenados en RI de la Biblioteca Central de dicha Institución. Cabe destacar la implementación de una herramienta de búsqueda basada en contexto para facilitar el acceso a tesis de posgrado. Dicha herramienta tiene además la capacidad de facilitar el mantenimiento de dicho repositorio, generando sugerencias relevantes (tales como código Dewey y descriptores) para cada nueva tesis que es almacenada en el RI.

A continuación incluimos una lista indicativa con algunas de las publicaciones más relevantes del grupo de investigación de la UNS en los temas señalados:

Akavipat, R., Wu, L.-S., Maguitman, A., Menczer, F. (2006) Emerging Semantic Communities in Peer Web Search. P2PIR 2006: International Workshop on Information Retrieval in Peer-to-Peer. ACM Press. Arlington, USA, November 2006.

Cañas, A., Leake, D., Maguitman, A. (2001). Combining Concept Mapping with CBR: Experience-Based Support for Knowledge Modeling. Proceedings of FLAIRS-01. Key West, Florida. AAAI Press, May, 2001

Cañas, A. J., Hill, G., Carff, R., Suri, N., Lott, J., Eskridge, T., Gómez, G., Arroyo, M., & Carvajal, R. (2004a). CmapTools: A Knowledge Modeling and Sharing Environment. Concept Maps: Theory, Methodology, Technology, Proceedings of the First International Conference on Concept Mapping. Pamplona, Spain: Universidad Pública de Navarra.

Cañas, A. J., Carvalho, M., Arguedas, M., Leake, D. B., Maguitman, A., Reichherzer, T., (2004b). Mining The Web to Suggest Concepts During Concept Map Construction. Concept Maps: Theory, Methodology, Technology, Proceedings of the First International Conference on Concept Mapping. Pamplona, Spain: Universidad Pública de Navarra.

Cecchini, R., Lorenzetti, C., Maguitman, A., Brignole, N. B. (2007a) Searching the Web in Context: Genetic Algorithms for Exploring Query Space. SSI 2007, 36º JAI-IO, Jornadas Argentinas de Informática e Investigación Operativa. Mar del Plata, 27-31 de Agosto de 2007.



Cecchini, R., Lorenzetti, C., Maguitman, A., Brignole, N. B. (2007b) Genetic Algorithms for Topical Web Search: A Study of Different Mutation Rates. VIII WASI: VIII Workshop de Agentes y Sistemas Inteligentes – CACIC 2007: XIII Congreso Argentino de Ciencias de la Computación. Resistencia (Argentina), 1 al 5 de Octubre de 2007.

Cecchini, R.L., Lorenzetti, C.M., Maguitman, A.G., & Brignole, N.B. (2008). Using genetic algorithms to evolve a population of topical queries. *Information Processing & Management*, 44(6), 1863–1878.

Cecchini, R.L., Lorenzetti, C.M., Maguitman, A.G., & Brignole, N.B. (2010). Multi-objective Evolutionary Algorithms for Context-based Search. *Journal of the American Society for Information Science and Technology*, 61(6):1258–1274.

Leake, D., Bauer, T., Maguitman, A., Wilson, D. (2000). Capture, Storage and Reuse of Lessons about Information Resources: Supporting Task-Based Information Search. Proceedings of the AAAI-00 Workshop on Intelligent Lessons Learned Systems. Austin, Texas. AAAI Press, July, 2000.

Leake, D., Maguitman, A., Cañas, A. (2002) Assessing Conceptual Similarity to Support Concept Mapping. Proceedings of FLAIRS-02. Pensacola, Florida. AAAI Press, May, 2002.

Leake, D., Maguitman, A., Reichherzer, T., Cañas, A., Carvalho, M., Arguedas, M., Brenes, S., Eskridge, T. (2003a). Aiding knowledge capture by searching for extensions of knowledge models. In Proceedings of KCAP-2003. St. Augustine, Florida: ACM Press.

Leake, D., Maguitman, A., Reichherzer, T., (2003b) Topic Extraction and Extension to Support Concept Mapping. Proceedings of FLAIRS-03. Saint Augustine, Florida. AAAI Press, May, 2003.

Leake, D., Maguitman, A., Reichherzer, T. (2004). Understanding Knowledge Models: Modeling Assessment of Concept Importance in Concept Maps. In Proceedings of CogSci 2004. Mahwah, NJ: Erlbaum.

Leake, D., Maguitman, A., Reichherzer, T., Cañas, A., Carvalho, M., Arguedas, M., and Eskridge, T.. "Googling" from a concept map: Towards Automatic Concept-Map-Based Query Formation. *Concepts Maps: Theory, Methodology, Technology*. Proceedings of the First Conference on Concept Mapping. A.J. Cañas, J.D. Novak, F. M. González Ed. Pamplona, Spain, September 2004

Leake, D., Maguitman, A., Reichherzer, T. (2005a). Exploiting Rich Context: An Incremental Approach to context-Based Web Search. To appear in Fifth International and Interdisciplinary Conference, CONTEXT 2005. Paris, France. *Lectures Notes in Artificial Intelligence*. Springer.

Leake, D., Maguitman, A., Reichherzer, T. (2005b). Integrating CBR with Knowledge Construction and Extension: Supporting Human Centered Knowledge Modeling. In *Case-Based Reasoning in Knowledge Discovery and Data Mining*, S. Pak, D. Aha and K. Gupta Ed. John Wiley & Sons, Inc, 2005.

Lorenzetti, C., Sagui, F., Maguitman, A., Simari, G., Chesñevar C. (2006) Incremental Methods for Context-Based Web Retrieval. XII Congreso Argentina de Ciencias de la Computación (CACIC), Argentina, Octubre 2006.

Lorenzetti, C., Cecchini, R., Maguitman, A. (2007). Intelligent Methods for Information Access in Context: The Role of Topic Descriptors and Discriminators. VIII WASI: VIII Workshop de Agentes y Sistemas Inteligentes – CACIC 2007: XIII Congreso Argentino de Ciencias de la Computación. Corrientes y Resistencia (Argentina), 1 al 5 de Octubre de 2007.

Lorenzetti, C., Maguitman A., (2008a). Learning Better Context Characterizations: An Intelligent Information Retrieval Approach. CLEI 2008. XXXIV Conferencia Latinoamericana de Informática. Santa Fe, Argentina, 2008.

Lorenzetti, C., Maguitman A., (2008b). Tuning Topical Queries through Context Vocabulary Enrichment. QSI 2008. Quantitative Semantic Methods for Internet. LNCS. Springer. Monterrey, México, 2008.

Lorenzetti, C., Maguitman, A., (2009). A Semi-supervised Incremental Algorithm to Automatically Formulate Topical Queries. Information Science - Special Issue on Web Search. 179(12):1881-1892. Elsevier, 2009.

Maguitman, A., Leake, D., Reichherzer, T., Menczer, F. (2004a). Dynamic Extraction of Topic Descriptors and Discriminators: Towards Automatic Context-Based Topic Search. Proceedings of the Thirteenth Conference on Information and Knowledge Management (CIKM' 04). ACM Press. Washington, DC, November 2004.

Maguitman, A. (2004b). Intelligent Support for Knowledge Capture and Construction. PhD thesis, Computer Science Department, Indiana University, December 2004.

Maguitman, A., Leake, D., Reichherzer, (2005a). Suggesting Novel but Related Topics: Towards Context-Based Support for Knowledge Model Extension. International Conference of Intelligent User Interfaces (IUI'05). ACM Press. San Diego, January 2005.

Maguitman, A., Menczer, F., Roinestad, H., Vespignani, A. (2005b). Algorithmic Detection of Semantic Similarity. Proceedings of WWW 2005, Chiba, Japan, May 2005.

Maguitman, A., Menczer, F., Ferdinc, F., Roinestad, H., Vespignani, A. (2006a) Algorithmic Computation and Approximation of Semantic Similarity. WWW Journal. Springer Science+Business Media B.V, 2006.

Maguitman, A. G., Rechtsteiner, A., Verspoor, K., Strauss, C.E., Rocha, L.M. (2006b). Large-Scale Testing Of Bibliome Informatics Using Pfam Protein Families. Pacific Symposium on Bioinformatics, Maui, USA, January 2006.

Maguitman, A. G., Cecchini, R. L., Lorenzetti, C. M. & Menczer, F. (2010) Using Topic Ontologies and Semantic Similarity Data to Evaluate Topical Search. To appear in CLEI 2010. Asunción, Paraguay.

Reichherzer, T. R., A. J. Cañas, K. M. Ford, P. J. Hayes (1998). The Giant: An Agent-based Approach to Knowledge Construction & Sharing, , The Eleventh International Florida Artificial Intelligence Research Symposium Conference, Sanibel Island, 1998.

Wu, L.-S., Akavipat, R., Maguitman, A., Menczer, F. Adaptive Peer to Peer Social Networks for Distributed Content Based Web Search (2007). In Social Information Retrieval Systems: Emergent Technologies and Applications for Searching the Web Effectively, IGI

## **JUSTIFICACION GENERAL DE LA METODOLOGIA DE TRABAJO**

Las experiencias de repositorios y bibliotecas digitales existentes en la región, más algunos resultados preliminares del proyecto de investigación sobre la situación del acceso abierto en la Argentina que el grupo de investigadores de la UNLP y UNR están llevando a cabo, revelan que el desarrollo de repositorios institucionales es un proceso complejo, en el que intervienen múltiples factores, entre los que podemos mencionar: políticas institucionales, recursos humanos formados, infraestructura tecnológica, flujos de trabajo con intervención de distintas áreas institucionales (biblioteca, académica, postgrado, ciencia y técnica, administrativa, informática, etc.), y por sobre todas las cosas hábitos de comunicación y difusión de los documentos académicos y científicos producidos en el seno de las Universidades. Por otra parte, la utilización de herramientas modernas en un contexto de aprovechamiento de las nuevas tecnologías debe aportar mejoras en la consistencia y optimización de los resultados de las búsquedas en los repositorios para favorecer al incremento de la visibilidad de tales documentos. La labor del bibliotecario será apoyada por la tecnología para incorporar elementos de control y aseguramiento de la calidad en la ingesta de los datos. Por otro lado, el usuario final de estos sistemas puede recibir sugerencias sobre cómo mejorar los resultados obtenidos luego de una primera intervención en un proceso de interacción.

Este escenario nos llevó a concentrar la atención en metodologías de trabajo orientadas al avance en una línea de acción primaria que contribuya al desarrollo de repositorios en la región, mediante la generación de conocimientos y la transferencia de experiencias de proyectos exitosos; con una propuesta metodológica que permita su sustentabilidad en el tiempo y con mecanismos orientados a la optimización de los recursos evitando la duplicación de esfuerzos. En este sentido, la transferencia de la experiencia de repositorios y la adaptación de los flujos de trabajo en cada institución implica una apuesta fuerte de formación de recursos humanos al que este proyecto dedicará una parte importante del subsidio solicitado, porque consideramos que los proyectos son sustentables en la medida que se garantice la existencia de especialistas formados para su desarrollo y mantenimiento, junto con estrategias de trabajo consensuadas e integradas entre las distintas áreas de la institución, comprometidas en los flujos de generación y transferencia de información y objetos digitales factibles de ser reunidos, difundidos y accedidos desde los repositorios. Por otra parte, los análisis de factibilidad y el desarrollo de herramientas de reutilización de datos procedentes de diferentes sistemas existentes en las instituciones se orientan a reducir la duplicidad de tareas de registro de datos favoreciendo además a agilizar el desarrollo y mantenimiento de los repositorios en cada institución. Finalmente, las métricas para la evaluación de uso de los repositorios se orientan a favorecer la elaboración de políticas institucionales a favor del acceso abierto junto a acciones que promuevan cambios de hábitos en la comunicación y difusión de los documentos académicos científicos producidos y objetos de aprendizaje en el seno de las universidades.

Por otra parte, teniendo en cuenta que este proyecto se inscribe en el contexto de un conjunto de otros proyectos regionales y de otras acciones coordinadas que se vienen desarrollando a nivel nacional para contribuir a la creación del Sistema Nacional de Repositorios Digitales de Ciencia y Tecnología del MINCYT, es que el grupo de investigadores, docentes, bibliotecarios y técnicos de las Universidades Nacionales de La Plata (UNLP), Mar del Plata (UNMdP), Luján (UNLU), Centro de la Provincia de Buenos Aires (UNICEN), y Sur (UNS) que integramos este proyecto hemos consensuado acciones que contribuyan al desarrollo y mantenimiento de repositorios institucionales en la región bonaerense que puedan ser trasferibles al resto de las regiones del país procurando la optimización de los recursos y evitando la duplicación de esfuerzos.

## **TIPO DE DISEÑO DE INVESTIGACION Y MÉTODOS**

Se propone un diseño metodológico que combina métodos de investigación exploratorio-descriptivos con el de diseño e implementación de prototipos. Asimismo la metodología contempla actividades de transferencia de conocimientos y experiencias, a efectos de hacer extensivos proyectos de repositorios en funcionamiento a instituciones que aún no cuentan con ello. En una línea asociada, los trabajos en la UNS proponen integrar los repositorios ya existentes y compartir asimismo la experiencia. Por otra parte se realizarán intensas actividades de formación de recursos humanos que permitan sustentar las iniciativas de repositorios que se creen en las instituciones.

Para alcanzar el objetivo específico 1, se realizará en cada institución participante un relevamiento de los sistemas y fuentes de datos propias en uso (por ejemplo: WinSip, SIGEVA u otros), en los cuales exista información (referencias) y contenidos académicos y científicos (objetos digitales) que podrían eventualmente formar parte de los repositorios institucionales (por ejemplo: cv de los investigadores, proyectos de investigación, revistas digitales, tesis, objetos de aprendizaje, etc.). Se realizará un análisis de los mismos, de su tipología documental, de la estructura de datos, y de la calidad de los datos. Se estudiará la factibilidad de reutilización para su incorporación a los repositorios. Se elaborará un informe diagnóstico y en caso de ser factible la reutilización en algún caso, se diseñará una propuesta de flujo de trabajo para la inclusión automática de los metadatos al repositorio. Como resultado de los estudios se explorará la necesidad de desarrollar un prototipo.

Para el objetivo específico 2, se prevé la creación de repositorios en aquellas UUNN (y/o Unidades Académicas) seleccionadas que aún no cuenten con RiS, o que tengan interés en integrar iniciativas parciales. Se transferirá la experiencia de Memoria Académica de la FAHCE-UNLP. En primera instancia se analizará la adecuación de la infraestructura informática existente para la instalación y puesta en funcionamiento del repositorio en las instituciones seleccionadas. Contando con la infraestructura requerida se procederá a la instalación de los programas informáticos. Por otra parte se realizarán actividades de capacitación en la modalidad de cursos y talleres para formar recursos humanos en el desarrollo y mantenimiento de los repositorios. Se propone extender la capacitación para dar lugar a la transferencia del uso y aplicación de las herramientas desarrolladas en la UNS según se ha descrito más arriba. Asimismo, se identificarán en cada institución las áreas involucradas en los flujos de información que aportarán los datos y objetos digitales al sector que administrará el repositorio, de manera de adaptar los flujos de trabajo del caso Memoria Académica al contexto de las otras instituciones.

Para el objetivo específico 3, se explorarán los estándares internacionales sobre métricas de uso de documentos electrónicos como los proyectos COUNTER (Counting Online Usage of Networked Electronic Resources) <http://www.projectcounter.org/> y MESUR (MEtrics from Scholarly Usage of Resources), desarrollado en Los Alamos National Laboratory Research (<http://www.mesur.org/MESUR.html>). Se estudiará también, y especialmente, el proyecto PIRUS (Publisher and Institutional Repository Usage Statistics), avalado por la JIST (Joint Information Systems Committee), <http://www.jisc.ac.uk/publications/reports/2009/pirusfinalreport.aspx>, que se propone desarrollar una serie de normas, protocolos y procesos que permitan a los editores y a los repositorios instituciones generar y compartir estadísticas fiables del uso de los documentos que contienen. Se evaluarán los softwares disponibles para la evaluación del uso de los documentos en repositorios. Se analizará la adopción de uno, atendiendo a las necesidades locales y/o desarrollará un prototipo basado en las normas y protocolos internacionales que pueda ser utilizado por cualquiera de las instituciones participantes con independencia del software que administra el repositorio.

Se realizarán consultorías y asesorías técnicas para la resolución de problemas específicos de cada institución.

Se contratarán servicios especializados para estudios, traducciones y desarrollos informáticos. Se contará con profesores investigadores invitados con formación especializada que contribuyan a la generación de conocimientos en materia de desarrollo de repositorios. El conjunto de actividades de capacitación, asesorías, reuniones de trabajo y transferencias previstas requerirán la movilidad por parte de los integrantes del equipo de trabajo en el ámbito de la región bonaerense.

#### **ACTIVIDADES A DESARROLLAR**

1. Realización de 5 reuniones de trabajo de 3 días cada una en las que se movilizarán 15 personas (3 x UUNN) para el desarrollo y avance de los estudios y transferencias previstas. En cada reunión se dedicará un día al estudio y relevamiento de objetos digitales existentes en los sistemas de información vigentes, otro día al estudio de las métricas de indicadores de uso de repositorios, y el tercero a la demostración de prototipos desarrollados en la UNS y capacitación para su uso.
2. Realización de 1 seminario de posgrado de 5 días de duración a cargo del Dr. Pablo de la Fuente Redondo (Universidad de Valladolid, España) en calidad de profesor visitante del exterior, para el que se movilizarán 10 personas (2 x UUNN).
3. Edición y publicación de documentos con los resultados obtenidos por el equipo de trabajo en revistas, congresos y otros medios.
4. Adquisición de 1 servidor para cada UUNN participante.
5. Asistencia técnica especializada a cargo de la Lic. Mariana Pichinini (FaHCE-UNLP) para realizar la transferencia de la experiencia de Memoria Académica a las UUNN de la región. Esto incluye el diseño de materiales didácticos y armado de una plataforma virtual para el dictado de cursos semi-

presenciales y seguimiento del proceso de transferencia, la realización de trabajos específicos de instalación de programas y esqueleto de las colecciones, migración de objetos digitales preexistentes, y el dictado de cursos específicos y visitas técnicas periódicas a cada una de las sedes a lo largo del año.

6. Asistencia técnica especializada para complementar el diseño e implementación de prototipos de aplicación para los desarrollos que se propongan como resultado del estudio de factibilidad sobre la posible reutilización e interoperabilidad de los objetos digitales existentes en otros sistemas
  
7. Asistencia técnica especializada para complementar el diseño e implementación de un prototipo que permita registrar estadísticas de uso de los objetos digitales de los repositorios y la generación de indicadores de impacto según estándares internacionales vigentes.

### **CRONOGRAMA DE TRABAJO**

En la tabla que sigue hemos realizado un cronograma tentativo de actividades a desarrollar en el proyecto a partir de 2011. Las filas llevan la numeración de las actividades enumeradas más arriba y las columnas indican los meses del año.

ACT	1	2	3	4	5	6	7	8	9	10	11	12
1		x		x			x		x		x	
2				x								
3												x
4			x									
5				x	x	x	x	x	x	x	x	
6									x	x		
7									x	x		