



SOLICITUD DE ACREDITACION DE PROYECTOS DE INVESTIGACION Y/O DESARROLLO PARA EL AÑO 2011 – PROGRAMA DE INCENTIVOS

1. IDENTIFICACION DEL PROYECTO

1.1 DIRECTOR:

Apellido y Nombres: DE GIUSTI, Marisa Raquel

Categoría de Docente Investigador: I II III X Sin Cat.

Cargo docente: Profesor Titular Dedicación: Simple

Unidad Académica: Facultad de Ingeniería, Universidad Nacional de La Plata

Cargo en la Carrera del Investigador (CIC): Investigador Adjunto Sin Director

Lugar de trabajo: Proyecto de Enlace de Bibliotecas (PrEBi), Facultad de Ingeniería, Universidad Nacional de La Plata

Documento Tipo: DNI

Títulos: Ingeniero en Telecomunicaciones

Domicilio Particular:

Tel.:

Localidad: La Plata Pcia: Buenos Aires C.P.1900

Email: marisa.degiusti@sedici.unlp.edu.ar

1. **UNIDAD ACADEMICA:** Facultad de Ingeniería-UNLP- **UNIDAD/ES EJECUTORA/S:** PrEBi (Proyecto de Enlace de Bibliotecas) y SeDiCI (Servicio de Difusión de la Creación Intelectual).

Dirección: 49 y 115 Tel: 0221- 423-6677/6696 int.141

E-mail: marisa.degiusti@sedici.unlp.edu.ar

1.3 DENOMINACION DEL PROYECTO:

“Aplicaciones para mejorar el acceso, la organización y la difusión del conocimiento”.

1.4 RESUMEN TÉCNICO:

El proyecto tiene como objetivo promover la socialización del conocimiento científico- académico con fines educativos y de integración, mediante la investigación y desarrollo en el área de las Tecnologías de la información y las comunicaciones (TICs) para potenciar y simplificar el acceso a la producción intelectual de la Universidad pública. Los resultados alcanzados serán aplicados en el Repositorio Institucional de la UNLP (SeDiCI) y estarán dedicados a: brindar un mecanismo de búsqueda y exploración de contenidos inteligente mediante la interpretación de consultas, la priorización de resultados y la incorporación de referencias secundarias; permitir el acceso a los contenidos a personas con deficiencias visuales; desarrollar nuevas técnicas de recolección desde repositorios externos; y evaluar y optimizar herramientas para metabúsquedas. Se utilizarán indicadores diversos como visitantes, búsquedas, resultados, cantidad de documentos y descargas, que darán cuenta del efecto logrado por cada nueva adición.

The main objective of this project is to support and promote the socialization of scientific and academic knowledge. The project aims to encourage both integration and education through the research and development in the ICT area in order to boost and simplify the access to the public University production. All achievements will be directly applied into UNLP Institutional Repository (SeDiCI) and will be devoted to: offer an intelligent search and browse mechanism including query interpretation, results sorting and similar references addition; allow content access to people with visual impairment; develop new techniques for external repositories data collection; and finally evaluate and optimize meta-search



Gracias a esta doble modalidad, SeDiCI permite a sus usuarios acceder a documentos locales (los actualmente casi 16000 registros), creados en la UNLP, y a documentos externos de Universidades e Instituciones académicas de todo el mundo. Así, SeDiCI cuenta con cerca de 18 millones de registros de documentos (provenientes de repositorios institucionales externos) de diferente naturaleza: tesis, artículos, imágenes, etc. Los usuarios pueden asimismo explorar y realizar búsquedas sobre estos *recursos externos* con un conjunto más pequeño de campos.

Del mismo modo que en SeDiCI, la información volcada en Internet es normalmente buscada por los usuarios utilizando motores de búsqueda comerciales —como Google, Bing, Yahoo!, etc.— que analizan las búsquedas de los usuarios por palabra clave en lugar de realizar una búsqueda por conceptos: de este modo, se pierden relaciones de importancia y con ello la información obtenida es diferente aún en el caso de que las palabras utilizadas para la búsqueda sean sinónimas.

Una de las alternativas actuales más interesantes y atinadas para superar este escollo es la denominada “web semántica”, una iniciativa del W3C liderada por Tim Berners-Lee. Los grandes buscadores —Google, Yahoo! y recientemente Microsoft Bing— han realizado algunos aportes en esta área, principalmente relacionados con la publicidad en línea. Por ejemplo, si un usuario busca el término “cansancio”, el buscador le ofrecerá adquirir complejos vitamínicos o visitar centros deportivos geográficamente cercanos. Existe también otro proyecto, relativamente nuevo, llamado Wolfram Alpha, el cual ofrece al usuario directamente la información requerida y no ya una lista de enlaces a páginas: esto lo hace mediante la utilización de técnicas para interpretar el contenido semántico de la búsqueda del usuario y generar la respuesta apropiada.

Más allá de los esfuerzos aislados realizados por grandes corporaciones internacionales como las citadas precedentemente, queda evidenciado que la gran masa de información, tanto la propia de la UNLP, como la recolectada y aún la nueva información temática disponible en la web que se quisiera incorporar a la Biblioteca, está sometida a la limitación de una búsqueda de naturaleza sintáctica.

La totalidad de los contenidos de SeDiCI, al ser presentados en su mayor parte de manera textual (a excepción de los documentos multimediales, con sonidos, imágenes y/o videos), requieren de determinadas capacidades por parte de los usuarios y no contemplan, hasta ahora, la posibilidad de capacidades diferentes. Por ejemplo, una persona con disminución visual no puede por sí misma tener acceso completo a la información del repositorio, lo que constituye una seria limitación, pues esto excluye a un conjunto numeroso de personas y contradice los principios que fueron fundamentales en el momento de la creación del portal, como la socialización del conocimiento o la igualdad de oportunidades que posibilita el aprendizaje sin limitaciones.

Otro problema constante de las bibliotecas y repositorios digitales está vinculado al manejo de volúmenes muy grandes de información, su disposición y su indexación, esto último, asociado, fundamentalmente, a la recolección en repositorios externos. Si bien esta problemática parecería alejada del portal en sí mismo y sus posibilidades, determina múltiples limitaciones: tanto a la cantidad de información que es posible ofrecer como a la posibilidad de buscar rápidamente esa información, y también otras cuestiones, como la frecuencia de recambio del material, la vigilancia de la calidad de los datos obtenidos y la constante observación de nuevos y mejores repositorios desde donde obtener el material externo para brindarlo a los usuarios. En este sentido, SeDiCI cuenta actualmente con un mecanismo de recolección que puede ser mejorado sustancialmente, gracias a las innovaciones que desean implementarse, para actualizar la información de manera automática, aumentar la cantidad de registros y lograr una adecuada vigilancia de los repositorios para que la información obtenida sea de la mejor calidad posible.

Uno de los objetivos centrales de SeDiCI es que los usuarios cuenten cada vez con más información de calidad, que pueda incluso lanzar una búsqueda en línea para incluir en los resultados



documentos provenientes de otros repositorios de similar naturaleza. La plataforma de SeDiCI, en su estado actual, no cuenta con ningún buscador de estas características disponible para los usuarios.

El estado actual de la plataforma tampoco permite realizar de manera sencilla y expeditiva cambios desde la administración interna, como incorporar nuevas secciones, noticias, novedades o bien la obtención de indicadores de acuerdo con las necesidades de información de la Universidad Nacional de La Plata. Todas estas características, que son susceptibles de ser mejoradas en el corto plazo, responden a la necesidad de que el portal se actualice y ofrezca a sus usuarios y visitantes una dinámica y fiabilidad cada vez mayor, acorde con las tendencias y desarrollos actuales. Aunque eran incipientes en el momento de la creación del proyecto, actualmente existen tecnologías y aplicaciones capaces de facilitar la interacción de los usuarios y administradores, con la facultad de hacer que el conocimiento y la información sean día a día más accesibles para todos y con la calidad que la comunidad universitaria requiere.

SeDiCI ha participado en la edición de libros electrónicos y su posterior catalogación y publicación en el portal web, de manera simultánea SeDiCI trabaja conjuntamente (desde el año 2008) con la Editorial de la UNLP (EduLP) para incluir en el repositorio institucional los materiales que son creados desde la editorial; este compromiso implica la definición de procedimientos y estándares para mantener un flujo continuo y evitar el re-trabajo de cualquiera de las partes. En Junio de 2010 SeDiCI comenzó a editar un libro electrónico multimedial, coordinado por Profesores e Investigadores de la Facultad de Ciencias Médicas de la UNLP y en la redacción de cuyas partes han participado Profesores de otras instituciones académicas de Argentina y el exterior. Este libro, que consta de más de dos mil páginas, contiene no sólo textos e imágenes, sino que incluye también videos de alta calidad con explicaciones muy detalladas sobre procesos e intervenciones quirúrgicas reales. Esto constituye un trabajo inédito en el país, y a la vez representa un material de estudio y aprendizaje con gran valor agregado. Durante esta experiencia, el equipo de SeDiCI se ha enfrentado a diversas situaciones desde el punto de vista tecnológico: formatos de publicación, tamaños de los archivos, estándares, calidad, combinación de contenidos textuales con multimediales y catalogación adecuada, fueron algunos de los nuevos desafíos que se presentaron. Todo este conocimiento adquirido puede y debe ser aprovechado para fomentar la creación de nuevos materiales de similares características por cátedras de todas las Unidades Académicas. Para ello, se deben definir procedimientos apropiados, redactar guías para usuarios, seleccionar y clasificar alternativas tecnológicas, y adaptar las herramientas existentes para difundir estos materiales de la mejor manera posible.

En consecuencia, todo lo procedente significará la incorporación de tecnologías y estándares web actualizados que permitan desarrollar aplicaciones complejas y completamente funcionales sin depender de una plataforma de escritorio específica. Un desarrollo adicional está vinculado a la posibilidad de ofertar a los usuarios una mayor personalización de contenidos y una más adecuada presentación de los mismos, utilizando las denominadas “Aplicaciones Enriquecidas de Internet” (Rich Internet applications-RIAs).

2.3 Aporte original al tema:

Desde su creación, en 1996, PrEBi ha participado en proyectos de cooperación científica con el Consorcio Iberoamericano para Educación en Ciencia y Tecnología (ISTEC), fundamentalmente en la Iniciativa Library Linkage (LibLink), dedicada a la constitución de una Red de Bibliotecas con el objeto de realizar el intercambio de información y la compartición de acervos bibliográficos usando TICs. Actualmente, PrEBi lidera el desarrollo de software para ISTEC y otras bibliotecas universitarias. Prueba de ello es el software Celsius NT, capaz de realizar todas las tareas de gestión bibliográfica, desarrollado en PrEBi y utilizado en más de 60 instituciones de América. En el año 2003 se creó, dentro del mismo PrEBi, el Servicio de Difusión de la Creación Intelectual (SeDiCI); la plataforma que sustenta esta Biblioteca Digital se denomina Celsius-DL y fue desarrollada completamente en el PrEBi. Desde el año 2004 se han realizado numerosas publicaciones sobre bibliotecas digitales y estándares, y asistido a foros nacionales e internacionales para presentar éstas y otras experiencias referidas al manejo de la información. En el año



2005, SeDiCI fue condecorado con el Premio INELAM de la Organización de Estados Americanos (OEA), galardón que le fue otorgado por su calidad como biblioteca digital.

El personal de PrEBi-SeDiCI ha desarrollado, además, el Portal de Información para la Universidad Nacional del Noroeste de la Provincia de Buenos Aires, ha puesto en funcionamiento el Portal de Revistas y el Portal de Congresos de la UNLP, para la gestión completa y en línea tanto de revistas académicas y científicas como de encuentros académicos. Mientras estos portales brindan sus servicios a los usuarios de la UNLP, el equipo de PrEBi-SeDiCI trabaja en conjunto con otros equipos de desarrollo a fin de mejorar las herramientas que sustentan estos portales, y para crear nuevos proyectos que sirvan de soporte para los usuarios de estas herramientas.

En síntesis, los proyectos PrEBi y SeDiCI lideran desarrollos vinculados a las TICs y los aplican en el manejo de la información, lo que ha significado un trabajo de investigación en paralelo a las actividades propias de cada servicio, que se ve reflejado en más de 30 publicaciones y presentaciones en congresos <http://sedici.unlp.edu.ar/difusion/difusion.php>, en las que se destacan temáticas como: digitalización y procesamiento de imágenes, preservación de patrimonio cultural, bibliotecas digitales, estándares y herramientas de búsqueda, manejo de grandes volúmenes de datos y simulación.

El acceso a información pertinente es un desafío permanente en cualquier biblioteca digital. Internet ofrece una cantidad de información casi infinita, pero el acceso a la información de calidad, apropiadamente clasificada y organizada es aún un grave problema a resolver. La capacidad de ofrecer una herramienta de búsqueda inteligente a los usuarios y a los visitantes del portal SeDiCI, que permita no sólo identificar registros a partir de la coincidencia de palabras sino también a partir de un contexto semántico más amplio, permitirá potenciar enormemente el servicio, ya que los usuarios de la biblioteca digital podrán ahorrar gran cantidad de tiempo al recibir información organizada en repositorios semánticos guiados concurrentemente por ontologías y tesauros.

El trabajo constante de recolección de nuevos materiales desde repositorios externos (a través del protocolo OAI-PMH), su almacenamiento e indexación significa en principio un trabajo de selección cuidadoso que se basa en parámetros como los siguientes:

- Cantidad de documentos: volumen de datos que el repositorio maneja.
- Actividad del repositorio: cuánta atención y trabajo se dedica al repositorio. Además, esto da una idea certera del nivel de actualización de la información que posee.
- Adecuación y corrección de la información que posee: contrastar datos similares en otros repositorios a fin de evitar aquellos que contengan errores o información no fiable.
- Tipos de material: que los documentos incluidos en los repositorios posean las temáticas buscadas y que se encuentren correctamente clasificados.

Se proyecta realizar en SeDiCI cosechas programadas, que permitirán dividir el trabajo de cosecha en partes individuales de modo de lograr una tarea de recolección incremental, lo que disminuirá el riesgo de una gran pérdida de datos por cortes en la conexión, cortes de luz o fallas en general, otorgando así también la posibilidad de interrumpir el proceso de cosecha para reanudarlo en otro momento.

Las normas de accesibilidad en la web existen desde hace ya muchos años, pero la realidad muestra que son raramente aplicadas en los portales en línea. En el caso particular de los repositorios y bibliotecas digitales, la falta de aplicación de estas normas implica que la tan valiosa información que estos ofrecen - principalmente textual- no puede ser fácilmente aprovechada por personas con dificultades visuales. En el marco de este proyecto, se planifica implementar dentro del portal de SeDiCi un conjunto de capacidades para mejorar la accesibilidad tanto al portal como a los contenidos que el mismo aloja.

2.4 Objetivos:

2.4.1. Objetivos Generales:



El objetivo general del proyecto es brindar un conjunto de herramientas/aplicaciones web más adecuadas y acordes a las necesidades de todos los visitantes y usuarios, y de este modo fomentar la cultura de compartición de recursos, vincular a personas en actividades e investigaciones académicas y científicas similares, contribuir a la generación de nuevos conocimientos y disminuir las brechas que dificultan el acceso de las personas a los materiales educativos. Para lograrlo, este proyecto se propone generar cambios constantes que inviten a las personas a consultar el portal y a agregar nuevo material, y que, de este modo, la Universidad Nacional de La Plata esté presente en los foros nacionales e internacionales contribuyendo con sus realizaciones a generar nuevas redes de conocimiento.

2.4.2. Objetivos Específicos:

El proyecto contempla tanto aspectos de investigación aplicada y de campo, como también una fase de transferencia. La línea de investigación aplicada tiene los siguientes objetivos específicos:

- a) la generación de un sistema constituido por dos grandes subsistemas: un módulo de extracción/población ontológica y un módulo de búsqueda accesible a los usuarios;
- b) la puesta en funcionamiento de una herramienta de código abierto para la recolección de recursos abiertos (a través del protocolo OAI-PMH) que sea confiable, eficiente, configurable y simple de utilizar, y que requiera una mínima atención por parte del administrador. Esto es, que cualquier persona, luego de una breve y adecuada capacitación, pueda administrar la herramienta sin que esta tarea le demande demasiado tiempo ni esfuerzo, teniendo la plena seguridad de que los documentos serán obtenidos y agregados al motor de búsqueda correctamente, y que hará un uso correcto de recursos del servidor en el que se encuentre funcionando.
- c) incorporar síntesis de voz a los documentos digitales presentes en SeDiCI, comenzando con el material propio de la UNLP, cuyos derechos ya han sido cedidos por los autores;
- d) poner en funcionamiento un herramienta personalizada de búsqueda múltiple, que permita agregar diferentes orígenes, independientemente de la tecnología utilizada por cada uno, posibilitando de este modo nuevas fuentes de consulta para los usuarios;
- e) materializar las experiencias y el conocimiento adquirido durante la creación del libro multimedial, para propiciar el desarrollo de estándares y procedimientos de trabajo y, a la vez, fomentar este tipo de iniciativas en las Unidades Académicas, tanto a nivel de grado como de posgrado.

Con el fin de asegurar la interoperabilidad semántica, imprescindible para el punto a) de los objetivos específicos, se analizarán en detalle las Tipologías Documentales (Nivel Bibliográfico) que se utilizarán en el elemento TIPO (dc:type) del modelo de metadatos de Dublin Core para que éstos sean altamente interoperables, construyendo un vocabulario, basado en los contenidos actuales de SeDiCI, pero también atendiendo a otros servicios similares para así proveer una estandarización base en el proceso de carga de material.

La línea de desarrollo constituye una constante en los proyectos del PrEBi; tal es así que actualmente SeDiCI está siendo sometido a una reingeniería completa que permitirá una mayor compatibilidad con otros servicios y una modernización del software que facilite la introducción de contenidos y cambios. Si a esto se suman las novedades que se lograrán a través de estas nuevas propuestas, se podrá contar nuevamente con un desarrollo probado y funcionando en búsqueda de su mayor capacidad de rendimiento. En este punto es interesante tener en cuenta que el equipo de trabajo del PrEBi desarrolla software y cuenta ya con cinco patentes de desarrollo en el área de las bibliotecas y los sistemas de información.



2.5 Metodología:

La metodología propuesta implica la utilización de nuevas técnicas en el área de las bibliotecas digitales, a saber:

a) uso de agentes, “crawlers”, procesadores GRDDL, mecanismos de inferencia, creación, uso y reuso de ontologías y generación de una plataforma de búsqueda semántica;

b) estudio e investigación en el área de procesamiento de voz y de los mecanismos de integración de la misma a una biblioteca digital;

c) estudio, investigación, instalación y personalización de cosechadores (“harvesters”);

d) incorporación de un metabuscador al portal web, capaz de ser configurado bajo distintos parámetros, y que luego será una nueva fuente de contenidos, los cuales serán sometidos a marcación y utilizados en la población de las ontologías y que, por lo tanto, incrementarán los contenidos del repositorio semántico;

e) estudio y adecuación para la especificación de un modelo de metadatos común para SeDiCI, atendiendo a las colecciones documentales básicas actuales y la interoperabilidad con otros servicios.

2.6 Metas / Resultados esperados en el desarrollo del proyecto:

Desde el punto de vista de los aspectos de aplicación y desarrollo en el área de web semántica, los resultados de este proyecto ampliarán en mucho el relativamente poco conocimiento de que se dispone actualmente sobre el área en las bibliotecas digitales. El agregado de tecnología para la síntesis de voz permitirá a la Universidad Nacional de La Plata incursionar por vez primera en una iniciativa de este tipo a nivel nacional y pondrá así de relieve las propuestas de inclusión que forman parte de sus fundamentos. En el proceso de recolección, algunos de los resultados esperados incluyen la posibilidad de que desde la administración de SeDiCI se puedan configurar los “sets” de interés de cada repositorio para una solicitud más puntual a los respectivos “data providers”; también se espera contar con la posibilidad de eliminar una cosecha ya realizada, quitando del motor de búsqueda los documentos obtenidos por ésta, y acceder a la bitácora (logs) generada durante la ejecución del harvester, para analizar su funcionamiento y determinar posibles puntos de falla. El agregado de más y mejores contenidos, obtenidos de repositorios académicos verificados en cuanto a la calidad de sus datos y a las necesidades de los usuarios, ayudará a la generación de nuevos conocimientos derivados de otras experiencias y, a la vez, propiciará la inclusión de nuevas creaciones de la UNLP, una comunicación más dinámica y fluida entre creadores y, en definitiva, un beneficio palpable e importante para la sociedad en general. A su vez, la posibilidad de una consulta simultánea (una metabúsqueda) por parte de los usuarios, a través de una interfaz amigable que liste y organice apropiadamente los resultados, dará, por un lado, un mayor índice de satisfacción a quienes utilizan el portal de SeDiCI y por el otro, un mecanismo de inclusión de nuevos contenidos agregados por los propios interesados, siguiendo algunas sencillas pautas estipuladas por los administradores de la plataforma de SeDiCI, debidamente expuestas en su página web (autoarchivo). Por último, la definición de estándares y procedimientos para la generación y publicación de materiales multimediales, junto a la difusión apropiada de los mismos en el portal web, incentivará este tipo de iniciativas entre todos los actores de la Universidad.



3. **ANTECEDENTES:** Desarrollar en hojas aparte numeradas según los antecedentes de la Unidad Ejecutora sobre la temática del proyecto propuesto restringido a los últimos 5 (cinco) años. Especificar: publicaciones, presentaciones a congresos de la especialidad, convenios con otras instituciones, etc.

ANTECEDENTES (2006-2011)

PUBLICACIONES

1. De Giusti, M. R.; Oviedo, N.; Lira, A. “Arquitectura ETL para la recolección de metadatos”. VI Simposio Internacional de Bibliotecas Digitales, Porto Alegre, Brasil, mayo de 2011.
2. De Giusti, M. R.; Oviedo, N.; Lira, A.; Sobrado, A.; Martínez, J. P.; Pinto, A. “SeDiCI – Desafíos y experiencias en la vida de un repositorio digital”. BIREDIAL, Primer Conferencia sobre Bibliotecas y Repositorios Digitales, Bogotá, Colombia, mayo de 2011.
3. De Giusti, M. R.; Gonzalo L. Villarreal; Agustín Vosou and Juan P. Martínez. “An Ontology-based Context Aware System for Selective Dissemination of Information in a Digital Library”. *Journal of Computing*. Vol.2, Issue 5: pp. 6-13; May 2010. ISSN 2151-9617.
4. De Giusti, M. R.; Villarreal, G. L.; Sobrado, A.; Lira, A. J. “Estadísticas distribuidas y gestión de la calidad en el marco de la Iniciativa de Enlace de Bibliotecas”. V Simposio Internacional de Bibliotecas Digitales. 27 y 27 de Octubre 2009. Albuquerque, New Mexico USA.
5. De Giusti, M. R.; Villarreal, G. L.; Sobrado, A.; Vosou, A. “Recuperación y clasificación automática de información, resultados actuales y perspectivas futuras”. V Simposio Internacional de Bibliotecas Digitales. 27 y 27 de Octubre 2009. Albuquerque, New Mexico USA
6. De Giusti, M. R.; Sobrado, A.; Vosou, A.; Villarreal, G. L. “Platform for Collection from Heterogeneous Web Sources and its Application to a Semantic Repository Organization at SeDiCI: Preliminaries”. *Journal of Computer Science & Technology*. Vol. 9 No. 2 pp 89-92. Oct. 2009.
7. De Giusti, M. R., Villarreal, G. L. “Socializing and Disseminating the Academic and Intellectual Creation: Experiences from La Plata National University”. International PKP Scholarly Publishing Conference 2009. 8-10 July 2009. Simon Fraser University, Vancouver, BC, Canada.
8. De Giusti, M. R.; Lira, A.; Villarreal, G. L. “Simulation Framework for teaching in Modeling and Simulation Areas”. *European Journal of Engineering Education*, Volume 33, Issue 5 & 6, October 2008, pages 587-596. ISSN 1469-5898.
9. De Giusti, M. R.; Lira, A.; Villarreal, G. L. “Storage of Simulation and Entities History in discrete models”. *Journal of Computer Science & Technology*. Volumen 8, Número 1, April 2008. ISSN 1666-6038.
10. De Giusti, M. R.; Villarreal, G. L.; Sobrado, A.; Lira, A. y Vila, M. M. “SeDiCI: Intellectual Creativity Diffusion Service at the Universidad Nacional de La Plata (UNLP)”. *D-Lib Magazine*. September/October 2008, Volume 14, Number 9-10. ISSN 1082-9873.



11. De Giusti, M. R.; Vila, M. M.; Sobrado, A.; Villarreal, G. L.; Lira, A. "SeDiCI | Servicio de Difusión de la Creación Intelectual – UNLP". *Revista Interamericana de Bibliotecología*. Volumen 31, Número 2, Julio-Diciembre de 2008, p.187-202. ISSN 0120-0976.
12. De Giusti, M. R.; Vila, M. M.; Sobrado, A.; Villarreal, G. L.; Lira, A. "Interconnection and sharing through webservices". Chameleon Federation Meeting. Setiembre de 2007, Bologna, Italia.
13. De Giusti, M. R.; Vila, M. M.; Sobrado, A.; Villarreal, G. L. ; Lira, A. "SeDiCI Intellectual Creation Diffusion Service". Chameleon Federation Cal Poly Meeting. Abril 2007. California Polytechnic Institute, San Luis Obispo, California, Estados Unidos.
14. De Giusti, M. R.; Lira, A.; Inafuku, F. G.; Sobrado, A. "Celsius Network". *Anales de la International Conference on Engineering Education ICEE 2007*. Coimbra, Portugal, 3-7 de Setiembre de 2007.
15. De Giusti, M. R.; Villarreal, G. L. "Celsius BloodHound". *Anales de la International Conference on Engineering Education ICEE 2007*. Coimbra, Portugal 3-7 de Setiembre de 2007.
16. De Giusti, M. R. "Celsius Network - Celsius Imago - Celsius Bloodhound". Chameleon Federation Cal Poly Meeting. Abril de 2007. California Polytechnic Institute, San Luis Obispo, California, Estados Unidos.
17. De Giusti, M. R.; Vila, M. M.; Villarreal, G. L. "Vector de características para el reconocimiento de texto manuscrito". 4º Simposium Internacional de Bibliotecas Digitales. Junio 2006. Málaga, España.
18. De Giusti, M. R.; Vila, M. M.; Villarreal, G. L.. "Manuscript character recognition. Overview of features for the Feature Vector". *Journal of Computer Science & Technology*, Volume 6, Number 2, pages 92-98, October 2006. ISSN:1666-6038.
19. De Giusti, M. R.; Marmonti, E. H.; Vila, M. M.; Sobrado, A.; Villarreal, G. L. "Tecnologías para propagar los contenidos de una Biblioteca Digital. Experiencias desde SeDiCI". 4º Simposio Internacional de Bibliotecas Digitales. Junio 2006. Málaga, España.
20. De Giusti, M. R.; Marmonti, E. H.; Vila, M. M.; Sobrado, A.; Villarreal, G. L.; Lira, A. "Celsius Network - Interconexión entre portales de Referencia a través de webservices". 4º Simposium Internacional de Bibliotecas Digitales. Junio 2006. Málaga, España.

8.

PATENTES

Programa de software. Derivación. Registro de la Dirección Nacional de Derechos de autor como Obra Publicada. Formulario número 07953, en fecha 20 de Abril de 2009. Expediente número 748735.

La mencionada obra registrada bajo el nombre de "Celsius NT software destinado al control de gestión de solicitudes bibliográficas en el marco del intercambio entre instituciones, y para el cálculo de estadísticas de la gestión. Permite realizar esta gestión mediante la interconexión de varias instancias Celsius".

Fecha de Edición de la obra: 1/12/2008



Autores registrados: De Giusti, Marisa R., Lira, Ariel, Sobrado Ariel y Villarreal Gonzalo. Titular: Universidad Nacional de La Plata.

Programa de software. Registro de la Dirección Nacional de Derechos de autor como Obra Publicada. Formulario número 07954, en fecha 22 de Junio de 2007.

Expediente número 597639.

La mencionada obra registrada bajo el nombre de “Celsius Network: Software destinado a la gestión de pedidos de material bibliográfico entre usuarios locales e instituciones, y a la interconexión con otras instalaciones de Celsius Network alrededor del mundo con el fin de automatizar parte de la gestión de los pedidos y el cálculo de estadísticas. Son autores del Software: De Giusti, Marisa R., Lira, Ariel, Sobrado Ariel e Inafuku, Fernando. Titular: Universidad Nacional de La Plata.

Página WEB. SeDiCI. Servicio de Difusión de la Creación Intelectual. Versión 2. Registrado ante la Dirección Nacional de Derechos de autor como Obra Publicada. Expediente número 471203, formulario 01999, en fecha 30 de Marzo de 2006. Son autores del Portal De Giusti, Marisa R., Marmonti, Emiliano y Jaquenod, Gisele. Titular: Universidad Nacional de La Plata.

Programa de software. Registro de la Dirección Nacional de Derechos de autor como Obra Publicada. Expediente número 356638, en fecha 12 de Octubre de 2004. La mencionada obra registrada bajo el nombre de “Celsius DL: Software para la gestión de biblioteca digital multimedial”. Son autores del Software: Vila, María Marta, Sobrado, Ariel, De Giusti, Marisa R. y Marmonti, Emiliano Horacio. Titular: Universidad Nacional de La Plata.

Página WEB. Registrado ante la Dirección Nacional de Derechos de autor como Obra Publicada. Expediente número 356128, en fecha 07 de Octubre de 2004. La mencionada obra registrada bajo el nombre de “Servicio de Difusión de la Creación Intelectual (SeDiCI)”. Son autores: Jaquenod De Giusti, Gisele, De Giusti, Marisa R. y Marmonti, Emiliano Horacio. Titular: Universidad Nacional de La Plata.

SUBSIDIOS RECIBIDOS

De Instituciones Nacionales

Institución otorgante: Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC). Año 2011. N° de resolución: 279/03. Monto: 4300\$. Duración: Un año.

Institución otorgante: Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC). Año 2010. N° de resolución: 279/03. Monto: 3420\$. Duración: Un año.

Institución otorgante: Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC). Año 2009. N° de resolución: 279/03. Monto: 3240\$. Duración: Un año.

Institución otorgante: Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC). Año 2008. N° de resolución: 279/03. Monto: 3000\$. Duración: Un año.

Institución otorgante: Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC). Año 2007. N° de resolución: 279/03. Monto: 3000\$. Duración: Un año.

Institución otorgante: Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC). Año 2006. N° de resolución: 279/03. Monto: 3500\$. Duración: Un año.



Institución otorgante: Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC). Año 2005. N° de resolución: 279/03. Monto: 3500\$. Duración: Un año.

De Instituciones Internacionales

Institución otorgante: Consorcio Iberoamericano para la Educación en Ciencia y Tecnología (ISTEC). Año 2008. Motivo: actualización del sitio web oficial de Celsius, el mantenimiento de los foros y el debugging de la nueva versión antes de la disponibilidad internacional. N° de resolución: no corresponde. Monto: 3000u\$s. Duración: Un año.

Institución otorgante: Consorcio Iberoamericano para la Educación en Ciencia y Tecnología (ISTEC). Año 2007. Motivo: desarrollo de Celsius Network. N° de resolución: no corresponde. Monto: 8000u\$s. Duración: dos años.

Institución otorgante: Consorcio Iberoamericano para la Educación en Ciencia y Tecnología (ISTEC). Año 2006. Motivo: desarrollo de la versión 1.6 de Celsius, el desarrollo del sitio web oficial de Celsius, el mantenimiento de los foros, la creación del actualizador y del instalador para la nueva versión, la actualización e instalación remota de todas las instituciones que así lo soliciten, la traducción al idioma inglés y al portugués de Celsius 1. N° de resolución: no corresponde. Monto: 2000u\$s. Duración: Un año.

Institución otorgante: Consorcio Iberoamericano para la Educación en Ciencia y Tecnología (ISTEC). Año 2005. Motivo: desarrollo del Software Celsius en su versión 1.5, su software instalador y la actualización y asistencia a las Universidades de América que lo instalaron, en el marco del Proyecto de Enlace de Bibliotecas (PrEBi). N° de resolución: no corresponde. Monto: 2000u\$s. Duración: Un año.

CURSOS DICTADOS:

“Gestión de la calidad según normas ISO. Aplicaciones a Software”. Curso de doctorado. Facultad de Informática de la Universidad Nacional de La Plata. Desde 2004 a la fecha.

“Modelos y simulación”. Curso de Grado. Facultad de Informática de la Universidad Nacional de La Plata. Desde 2004 a la fecha.

“Gestión y certificación de Bibliotecas Universitarias”. Curso de especialización. Facultad de Ciencias Exactas, Universidad Nacional de La Plata. Marzo de 2006.

CONFERENCIAS Y TALLERES:

“SeDiCI: desafíos y experiencias en la vida de un repositorio digital”. BIREDIAL 2011, Primer Conferencia sobre Bibliotecas y Repositorios Digitales, Bogotá, Colombia, mayo de 2011.

“SeDiCI” (videoconferencia). Semana del Acceso Abierto, Jornada Virtual organizada por el Ministerio de Ciencia, Tecnología e Innovación Productiva de la Nación, Octubre de 2010.

“Rumbo a la biblioteca digital”. Universidad La Salle (Bolivia) en el marco del Congreso Internacional de Tecnologías, Información y Comunicación, Septiembre de 2009 .

“LibLink: plan estratégico 2008-2010” y “Celsius Network”. Universidad de La Sabana, Bogotá, Colombia e ISTEC. 20 de Mayo de 2009.

“Portal de Revistas”. Universidad Nacional de La Plata. Octubre de 2008, Noviembre de 2008 y Diciembre de 2008.

“Portal de Congresos”. Universidad Nacional de La Plata. Abril de 2009.

“Celsius 1.6” y “Celsius Network”. Universidad Técnica Particular de Loja (Ecuador). 15 y 16 de Noviembre de 2006.



“Digital Library Administration: Celsius Software”. University of South Florida, Florida USA. 1 y 2 de Diciembre de 2004.

“Servicio de Difusión de la Creación Intelectual (SeDiCI)”. Ministerio de Educación, Ciencia y Tecnología de la Nación. 8 y 9 de Marzo de 2004.



4. APORTES POTENCIALES:

4.1 Contribución al avance del conocimiento científico y/o tecnológico.

El PrEBi ha desarrollado una gran experiencia en las áreas vinculadas al manejo de información y las bibliotecas. Es el principal grupo de investigación y desarrollo de la Universidad Nacional de La Plata en esta temática y está constituido por un equipo multidisciplinario de distintas áreas del conocimiento: ingeniería, informática, bibliotecología, lingüística y diseño, adecuadamente balanceadas, lo que ha permitido obtener, en los últimos cinco años, cinco patentes de software, realizar veinte publicaciones referidas a las TICs y su aplicación en el manejo y disposición de la información, el procesamiento de imágenes para preservación de patrimonio, los sistemas de gestión de bibliotecas, los sistemas de difusión, los modelos computacionales para la optimización, etc. Ha participado constantemente en reuniones científicas en el país y especialmente en el exterior, donde se han dictado conferencias, expuesto trabajos y organizado talleres de capacitación. Su liderazgo se aprecia en su penetración a través de los desarrollos propios en las instituciones de América y España, integradas a ISTECA, con las cuales constituye una red siempre activa.

Estos resultados son producto de la constitución de un grupo de trabajo que en pocos años ha incrementado progresivamente el número de sus integrantes, que en este momento alcanza una masa crítica de veinte personas, un investigador, dos becarios y profesionales que integran la planta de la Universidad o son pasantes. Se trata de un grupo muy joven, con un crecimiento acelerado, fruto de acciones y desarrollos concretos y verificables en el área de los sistemas de información.

El valor de las bibliotecas digitales depende de la calidad de los contenidos y de su organización, así como de los sistemas de gestión que facilitan el acceso a los datos almacenados; en el mismo sentido también resulta importante un sistema de comprobación y seguimiento continuo que permita ver la evolución del diseño adoptado y recoja las respuestas de los usuarios.

Una innovación como la que se pretende poner en marcha en el marco del SeDiCI y de las bibliotecas digitales, podría también ser una fuente de nuevos datos y comportamientos a ser analizados. Como se ha dicho, se pretende que las herramientas informáticas a utilizarse sean altamente configurables, tanto por quien las administra como por quien sea su usuario final. Esto determina, necesariamente, el desarrollo de nuevas capacidades que el proyecto está en condiciones de ofrecer.

Como se mencionó anteriormente, el campo de las bibliotecas digitales, del manejo de sus recursos y de cómo mejorar la estructuración de los grandes volúmenes de información para hacerlos llegar efectivamente al público, es todavía un campo en el que escasean los estudios serios y fiables, debido a la naturaleza heterogénea y novedosa aún de estas bibliotecas y a la falta de intercomunicación entre sus diferentes actores. Por ejemplo, no hay, hasta el momento, un modelo de catalogación estandarizado e igual para todas. Más allá de sistemas como Dublin Core, que parece ser el más funcional pero también el más problemático, las diferentes bibliotecas adoptan, en muchas ocasiones, sus propios métodos de catalogación, lo que complica enormemente la tarea de cosecha, sin contar con que esto también confunde al usuario final, quien puede llegar a encontrar el mismo recurso catalogado de maneras muy disímiles en distintos catálogos o bibliotecas. En este sentido, las innovaciones que se buscan introducir en SeDiCI quieren ser también un campo de pruebas para determinar qué método de catalogación resulta más adecuado, atendiendo a dos principios fundamentales: la interoperabilidad con otros estándares y la usabilidad y confiabilidad para el usuario final.

Uno de los desafíos, tanto de los sistemas de información como las bibliotecas o la propia web, es la gestión eficaz del gran volumen de documentos que se almacena para facilitar a los usuarios un acceso sencillo y ágil. Este problema adquiere una mayor dimensión cuando el usuario dispone de un alto grado de especialización y requiere recursos muy específicos, como en el caso de investigadores que buscan documentos de su área científica.



En el ámbito de las bibliotecas y centros de documentación tradicionalmente se han propuesto diferentes soluciones, como la creación de servicios de difusión selectiva de información (DSI), sistemas de recomendación basados en contenido, en los que de acuerdo con el perfil de los usuarios suscritos se generan periódicamente (o a petición del propio interesado) una serie de alertas en las que se les notifica la existencia de recursos que se adecuan a sus intereses.

La introducción de las nuevas tecnologías de la información ha provocado que las bibliotecas se vieran en la necesidad de reorientar los servicios que prestan para adaptarlos a este nuevo escenario, en un intento de satisfacer demandas cada vez más complejas y específicas. Así, por ejemplo, servicios como los de DSI han sido objeto de estudio en el área de las bibliotecas digitales o de forma muy incipiente aún en el entorno de sistemas multiagente.

Una novedad que se pretende instaurar es un modelo de plataforma web en la que los procesos de comunicación entre agentes, y entre agentes y usuarios, no se vean entorpecidos por las diversas formas en que la información está representada; que, aunque exista heterogeneidad en la representación de la información, ésta pueda ser reutilizada en otros procesos o por otras aplicaciones; y que se dispongan mecanismos ágiles y sencillos para generar las alertas superando otros servicios DSI que se implementan sobre plataformas web con una arquitectura multiagente.

4.2 Contribución a la formación de recursos humanos.

La formación de recursos humanos especializados en el área de las Tecnologías de la Información aplicadas específicamente a los repositorios digitales, es una de las prioridades del PrEBi-SeDiCI, dependiente de la Presidencia de la Universidad Nacional de La Plata. Con el apoyo solicitado no sólo se podrán continuar las tareas que vienen desarrollándose en este sentido, sino que además podrán implementarse todas aquellas que resulten necesarias para cumplir con los objetivos propuestos:

a) Cursos:

A través de los mismos se proporciona formación teórica y entrenamiento en los temas vinculados a la gestión de bibliotecas y a modelos discretos, útiles para simular y optimizar el comportamiento de sistemas informáticos. En estos cursos participan alumnos y graduados de diferentes provincias de nuestro país.

b) Seminarios, conferencias y talleres:

A través de los mismos, PrEBi dicta talleres del uso de las plataformas propias como Celsius NT, para gestión de bibliotecas, en los que participan profesionales de 22 países de América y España. Esta tarea la realiza en el marco del Proyecto Library Linkage (LibLink) del Consorcio Iberoamericano para Educación en Ciencia y Tecnología (ISTEC). Del mismo modo dicta talleres sobre los portales que ha instalado y configurado para la Universidad Nacional de La Plata (Portal de Revistas y Portal de Congresos), a los que asisten profesionales de las distintas unidades académicas. En todos los casos el equipo de trabajo está dedicado a difundir toda la experiencia y conocimientos generados en su espacio y a dar a conocer las diversas herramientas para la difusión del conocimiento.

c) Tesis doctorales:

Actualmente se encuentran en ejecución tres tesis doctorales.

d) Pasantías, becas o concurrencias:

De este modo se brinda capacitación a jóvenes graduados que se incorporan al grupo de trabajo o a estudiantes avanzados de las carreras de Informática, Bibliotecología y Letras que configuran perfiles diferentes para analizar la problemática vinculada a las bibliotecas digitales. En



2011, PrEBi está formando cuatro becarios de la Universidad Nacional de La Plata dedicados a desarrollo y gestión en dicho proyecto.

e) Interacción con otros grupos de investigación:

El software Celsius Network, desarrollado y mantenido por PrEBi, ha servido de enlace entre su equipo de trabajo y los equipos técnicos y de investigación de las universidades de América que utilizan la plataforma. De este modo, se han establecido estrechas relaciones, ya sea mediante el trabajo colaborativo dentro de la iniciativa LibLink, como en la capacitación cruzada entre los distintos grupos, la participación en congresos y simposios internacionales y el trabajo conjunto en tareas de detección de mejoras y mantenimiento de Celsius Network. Desde el año 2004, el ISTECS ha asignado recursos al PrEBi para mantener y ampliar el soporte que se brinda desde la UNLP hacia todas las demás universidades que utilizan este sistema, o que desean formar parte de esta red de bibliotecas.

Dentro del marco del ISTECS, el equipo del PrEBi lidera también el desarrollo de un software dedicado a realizar una metabúsqueda sobre los catálogos de las bibliotecas cooperantes de ISTECS, del cual también participan de manera activa otros equipos de trabajo de Colombia, Venezuela, Brasil y Estados Unidos.

Por otra parte, en estos momentos, PrEBi colabora con el equipo de trabajo de PKP (del cual participan la Universidad de Stanford, la Universidad Simon Fraser, la Universidad de British Columbia, de Canadá, y la Universidad del Estado de Arizona, de Estados Unidos), en el mejoramiento de sus plataformas OJS y OCS, en las que PrEBi ha aportado numerosas soluciones para problemas puntuales que se derivaban de su uso y también como equipo de *testing* de versiones en desarrollo. Estas plataformas de software sustentan actualmente los portales de Congresos y de Revistas de la UNLP, los cuales son mantenidos por el personal de PrEBi.

4.3 Transferencia prevista de los resultados, aplicaciones o conocimientos derivados del proyecto

Los resultados que se esperan obtener como consecuencia de la ejecución y la puesta en práctica de estas innovaciones dentro de un sistema ya existente y en funcionamiento implican, ante todo, una mayor visibilidad del sitio web de SeDiCI y, consecuentemente, una mayor visibilidad de la Universidad Nacional de La Plata en su conjunto, en tanto y en cuanto el núcleo fundamental del acervo bibliográfico de la biblioteca digital lo constituyen los documentos generados en esta alta casa de estudios. Una mayor visibilidad implica no sólo más visitas a la página web sino una mayor y mejor utilización de todos los servicios brindados por ella; servicios pensados no sólo para el público especializado sino también para la comunidad toda. De ahí la preocupación por integrar estas nuevas tecnologías que permitan, entre otras cosas, a los disminuidos visuales la utilización de los recursos disponibles con la misma celeridad que cualquier otro usuario.

Es de notar que un buscador inteligente como el que se desea implementar supone también un incremento sustancial de las visitas, a la vez que supondrá un ahorro de tiempo notorio para los usuarios: el buscador semántico le ofrecerá la información que más se aproxime a su pedido, en lugar de arrojar resultados disímiles, desordenados y, en la mayoría de los casos, no pertinentes con los elementos o materias buscados. Un buscador de esta naturaleza puede volverse, como es dable ver, una herramienta de primer orden a la hora de realizar ya sea una investigación académica, una revisión de bibliografía existente sobre el tema requerido o bien una fuente de información adecuada frente a cualquier tipo de consulta (erudita o no).



Así, la Universidad Nacional de La Plata, a través de SeDiCI, que bien podría ser su órgano difusor o su canal de contacto con otras universidades, grupos de trabajo y equipo de investigación alrededor del mundo, podría ponerse a la cabeza de las innovaciones en materia de TICs y de difusión y compartición de sus materiales educativos y académicos, convirtiéndose en un ejemplo a seguir por otras universidades nacionales y de la región.

5. PLAN DE TRABAJO:

5.1. SeDiCI: Repositorio Semántico: El propósito central es construir una plataforma que pueda atender consultas de los usuarios realizadas por medio de conceptos y encadenamiento de conceptos. Para lograrlo es necesario, más allá de posibilitar un mecanismo de búsqueda distinto, tener un repositorio de material distinto; deberá almacenar los recursos y sus metadatos (o solamente estos últimos), posibilitando la realización de búsquedas semánticas. Las ontologías juegan un rol importante para el soporte de un modelo semántico y el segundo elemento de importancia son los tesauros. Un tesoro es un instrumento terminológico que dota de coherencia al proceso de indización y recuperación, al posibilitar que los significantes utilizados en la búsqueda coincidan con los que representan a ese concepto ya indizado y almacenado.

El proyecto contempla varias etapas: En la primera hay que dotar a la información en sí misma de marcas semánticas e integrarla en un esquema clasificatorio donde se realicen las pruebas de inferencias (comprobar si las instancias cumplen las clases y sus relaciones). En este punto es posible decir que se ha constituido un repositorio de naturaleza semántica a través de dos macro-operaciones: extracción y población. La segunda es posibilitar al usuario una pantalla de búsqueda guiada sobre la o las ontologías. Es interesante aclarar que la marcación de la información se realizará a priori en una copia del repositorio SeDiCI, se comprobarán las operaciones y recién allí se buscarán nuevas fuentes heterogéneas de información como la web. El proyecto contempla atender a dificultades originadas por diferencias terminológicas, de idiomas, etc. de modo de brindar información útil para los usuarios. A semejanza de un proyecto que está en sus inicios, como el auspiciado por el Cross Lingual European Forum (CLEF), en etapas futuras lo que se pretende es satisfacer a un usuario que requiera información útil, en lugar de páginas web: esa información consistente podría requerirla para distintos fines, sean éstos un artículo, un trabajo, etc. Para llegar a este objetivo se requieren los subsistemas que se describen en los dos siguientes puntos.

5.1.1 Módulo de extracción y población: El sistema en su primera versión, realizará las siguientes tareas: a través de un agente de software recorrerá una lista de URLs de páginas web almacenadas en una base de datos, enviará las mismas a un “web-crawler”, encargado de recolectar otros enlaces embebidos en cada página visitada, para luego incorporarlos a la lista previa. Un segundo agente se encargará de pasar cada URL a un procesador del tipo GRDDL (Gleaning Resource Descriptions from Dialects of Languages) para aplicar determinadas transformaciones, de modo de contar con un documento textual con un formato RDF adecuado para el trabajo con ontologías. Este documento textual sólo contendrá las tuplas de interés, marcadas, lo que permitirá a la aplicación extraer de forma automática información de páginas web estructuradas para integrarlas en un repositorio. El procesador GRDDL detectará en las páginas los microformatos de interés y con la indicación de la dirección de la hoja de transformación (XSL) usada para capturar los microformatos y la URL del lugar de extracción, devolverá un XML con la información extraída por la hoja de transformación, este documento XML será transformado a RDF para que los datos sean homogeneizados, clasificados y se pase a poblar la ontología. La homogeneización se realizará mediante conversiones (uso de sinónimos) y también traducciones para tener un idioma único. La clasificación de la información se hará mediante una ontología de dominio; durante esta etapa, a partir del RDF, se crearán las instancias con atributos y relaciones que poblarán la ontología. Se deberá verificar que



las instancias satisfagan las restricciones impuestas por su clase a través de un módulo de razonamiento. Finalmente se realiza el almacenamiento en un repositorio semántico accesible vía web, el primer caso será sobre el propio repositorio de SeDiCI.

5.1.2. Módulo de búsquedas

El buscador semántico permitirá a los usuarios realizar búsquedas siguiendo la estructura de la ontología en el repositorio semántico poblado previamente. Hará uso de tesauros para expandir la consulta y unificará terminologías (idiomas). Será independiente del dominio de la ontología para poder reusarlo. Las búsquedas serán realizadas vía protocolo HTTP sobre el repositorio ontológico ya que se utilizará SPARQL (Protocol and RDF Query Language), que es la recomendación del W3C para crear un lenguaje de consulta dentro de la Web semántica.

El usuario ingresará al repositorio y se le mostrará una pantalla que lo guiará para realizar la consulta. Las pantallas atenderán los términos de la ontología, el usuario podrá acotar los valores de los atributos de las instancias, o añadir restricciones más complejas. Las opciones de la consulta se determinarán en función de la ontología utilizada, realizada la consulta al servidor, éste devolverá un archivo XML, lo procesará y mostrará los datos de forma simple.

5.2. Incorporación continua de materiales en SeDiCI: Proceso de cosecha

5.2.1 Incorporación continua de materiales del exterior: Completar las herramientas de harvesting actualmente utilizadas para mejorar el proceso de cosecha de SeDiCI. El objetivo es entonces implementar un mecanismo de cosechas semi-automático que permita una actualización permanente de la colección de recursos adquiridos vía OAI desde el exterior y, en particular, que sea posible determinar entre otras variables: los repositorios sobre los cuales se realizará el harvesting, soportando los diferentes formatos de metadatos; eligiendo los sets de entre los que provee el repositorio; posibilitar cosechas individuales por rangos de fechas; admitir pausa, reanudación e incluso retomar cosechas fallidas. Los objetivos hacen necesaria la creación de dos herramientas de software: una de administración del harvester y otra de ejecución del proceso de harvesting propiamente dicho. La primera herramienta consiste en un sistema de administración que permita gestionar de manera sencilla e intuitiva todos los aspectos relacionados con la configuración, programación y puesta en marcha del proceso de cosecha. Será posible desde la misma acceder al estado actual del harvester, pudiendo pausar alguna o posiblemente todas las cosechas en curso, retomarlas o reiniciarlas en caso de haberse producido algún error. Esta herramienta deberá asimismo permitir la visualización del registro histórico sobre los intentos de cosecha anteriores, además del estado del proceso actual de harvesting, para así realizar análisis de performance y robustez de la aplicación. Los recursos recolectados deben ser normalizados y luego transformados a un formato interno común para su almacenamiento; por ello es necesario que la herramienta permita definir reglas. La segunda herramienta es la encargada de obtener, transformar, guardar e indexar los datos provenientes del exterior según las configuraciones definidas. Por los volúmenes de información hay que tener en cuenta principios de concurrencia y asincronía de eventos con el fin de realizar el proceso en tiempos y formas medianamente aceptables, sin colapsar servidores ni conexión.

5.2.2 Incorporación continua de materiales de la UNLP: existe un volumen de documentos alojados en las bibliotecas de las unidades académicas de la UNLP, que podrían hacer uso de las herramientas disponibles en SeDiCI para organizar, recuperar y difundir el conocimiento. El mecanismo actual de incorporación a SeDiCI se basa en gran parte en trabajo manual: adquisición del material, y catalogación. Un ingreso manual de todo el acervo de producción intelectual de la UNLP no es viable, se plantea la definición de un mecanismo automático de adquisición de materiales por parte de las unidades académicas y de recolección desde SeDiCI. La comunicación entre las bibliotecas de las unidades académicas y SeDiCI se realizaría bajo el protocolo OAI-PMH de intercambio de metadatos.



5.3. Accesibilidad de SeDiCI: Estándares internacionales y Síntesis de voz: Se considera de importancia que los contenidos del portal SeDiCI sean accesibles para cualquier usuario, proporcionando así igualdad de oportunidades a las personas con capacidades diferentes. Al hablar de accesibilidad web se hace referencia a un diseño web tal que permita a estas personas percibir, entender, navegar e interactuar con la Web, aportando a su vez contenidos. La accesibilidad web debe beneficiar también a personas de edad o con algún tipo de discapacidad transitoria. La W3C ha creado la “Iniciativa de Accesibilidad Web” (WAI) que brinda normas, técnicas y pautas consideradas estándares, que proporcionan soluciones para el software y adaptaciones para los desarrolladores para permitir el uso de los sitios web a personas con capacidades diferentes. Una pauta puede ser: “a toda imagen debe adjuntársele un texto que la describe”; de este modo, los programas que aportan accesibilidad web en las máquinas de cada usuario, interpretarán estos textos asociados a imágenes y le informarán al usuario. Además de la adopción de pautas y normas de accesibilidad web, este proyecto tiene un objetivo aún más ambicioso: ofrecer a los usuarios del portal todos los documentos generados por miembros de la UNLP en un formato audible y reproducible dentro del mismo sitio web. Para llevar a cabo este objetivo, será requerido analizar y estudiar: 1) los mecanismos de reproducción de audio en páginas web y 2) la generación de documentos audibles a partir de otros documentos. Desde hace ya varios años, los sitios web son capaces de reproducir sonido y video embebido. Pero a diferencia de los sitios web con capacidad de reproducción de este tipo de contenidos (como por ejemplo YouTube, LastFM o las cadenas de TV en línea), que requieren el uso de software específico para reproducir cada documento multimedial (reproductores como Adobe Flash Player, Windows Media Player, Real Player, MPlayer, entre otros), SeDiCI busca mantenerse dentro los estándares internacionales y aprovechar los avances en estos estándares (como, por ejemplo, la incorporación de nuevas etiquetas para componentes multimediales en el estándar HTML versión 5), así como también la optimización de recursos tecnológicos como el ancho de banda requerido para descargar y reproducir el audio. Se estudiarán los procesos de síntesis concatenativa y de síntesis de formantes. Se analizarán algunos proyectos como MBROLA. Se estudiará también la “síntesis específica para un dominio” (concatenativa). Se analizarán en profundidad, entre otras, las siguientes plataformas que incluyen sintetizadores de voz abiertos: Festival, Flite, FreeTTS y MBROLA. Del mismo modo se analizarán los agentes animados con interfaces conversacionales que proporcionan un paradigma intuitivo de interacción que tienen como resultado que los usuarios encuentren estos sistemas más amigables y cooperativos.

5.4 SeDiCI: estandarización de datos de carga

Se desea destacar que hasta hace algunos años, la lectura de material gráfico para una persona con discapacidad visual implicaba, entre otros, la asistencia de un lector, la transcripción al sistema Braille o la grabación y posterior reproducción del texto. Las nuevas tecnologías han permitido el desarrollo de programas de lectura de pantalla y síntesis de voz, han mejorado marcadamente las posibilidades de acceso a la información, generando un nuevo horizonte de expectativas.

Algunas experiencias en el área de Bibliotecas Digitales para disminuidos visuales, como el proyecto Tiflos, marcan un nuevo compromiso para la biblioteca digital: consolidarse como un espacio de participación que aumenta su caudal de materiales y servicios a los usuarios de todo tipo en forma permanente. Esto significa que una biblioteca digital debe contar con contenidos digitalizados adecuadamente, esto es, accesibles para los programas de lectura y/o con mecanismos propios de asistencia a usuarios con problemas visuales y que, además, debe ofrecer acceso gratuito e ilimitado a los textos completos. Al mismo tiempo, debe llevar adelante una política de renovación permanente, teniendo en cuenta que la igualdad de posibilidades de acceso a la información y la cultura requiere de un trabajo constante y de objetivos sustentables y concretos. Es en este sentido que se desea orientar la tarea de explorar, generar y potenciar el uso de las nuevas tecnologías y metodologías para maximizar la accesibilidad a los contenidos y expandir la cultura de inclusión de todos los individuos por igual.



6. RECURSOS HUMANOS INTERVINIENTES:

6.1 DIRECTOR

- **Apellido y Nombres: De Giusti Marisa Raquel**

- Cargo/s: Profesor titular

- Dedicación/es: Simple

- Facultad o Escuela Superior a la que pertenece: Facultad de Ingeniería

- Universidad: Universidad Nacional de La Plata

- Categoría docente - Investigador : III .

Cargo en la Carrera del Investigador: Investigador Adjunto Sin Director de la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires. Cambio de categoría en trámite.

- Título: Ingeniero en Telecomunicaciones

- **Dedicación en este proyecto en horas por semana: 33**

6.5 TESISISTAS, BECARIOS

- **Apellido y Nombres: Sobrado, Ariel**

- Cargo/s: Jefe de Trabajos Prácticos

- Dedicación/es: Semi-Exclusiva

- Facultad o Escuela Superior a la que pertenece: Facultad de Informática

- Universidad: Universidad Nacional de La Plata

- Lugar de Trabajo. Proyecto de Enlace de Bibliotecas (PrEBi) y Servicio de Difusión de la Creación Intelectual (SeDiCI).

- Categoría docente - investigador : No posee

- Título: Analista en Computación.

- **Dedicación en este proyecto en horas por semana: 15**

- **Apellido y Nombres: Villarreal, Gonzalo Luján**

- Cargo/s: Ayudante Diplomado

- Dedicación/es: Simple

- Facultad o Escuela Superior a la que pertenece: Facultad de Informática

- Universidad: Universidad Nacional de La Plata

- Lugar de Trabajo. Proyecto de Enlace de Bibliotecas (PrEBi) y Servicio de Difusión de la Creación Intelectual (SeDiCI).

- Categoría docente - investigador: V

- **Becario de Posgrado Tipo II. Conicet.**

- **Dedicación: Exclusiva.**

- **Facultad de Ingeniería.**

- Título: Licenciado en Sistemas.

- **Dedicación en este proyecto en horas por semana: 33**

- **Apellido y Nombres: Texier, José Daniel**

- Cargo/s: Becario externo

- Dedicación/es: Exclusiva

- Facultad o Escuela Superior a la que pertenece: Departamento de Informática

- Universidad: Universidad Nacional del Táchira (UNET), Venezuela.



- Lugar de Trabajo: Proyecto de Enlace de Bibliotecas (PrEBi) y Servicio de Difusión de la Creación Intelectual (SeDiCI).
- Categoría docente - investigador:
- Título: Ingeniero en Informática. Master en Computación
- **Dedicación en este proyecto en horas por semana: 33**

6.6 COLABORADORES

- Apellido y Nombres: Oviedo, Néstor

- Cargo/s: Cargo No Docente, de Planta, categoría VII. Universidad Nacional de La Plata.
- Lugar de trabajo: Proyecto de Enlace de Bibliotecas (PrEBi) y Servicio de Difusión de la Creación Intelectual (SeDiCI).
- Facultad a la que pertenece: Facultad de Ingeniería
- Universidad: Universidad Nacional de La Plata
- Título: Estudiante de Informática.

- Apellido y Nombres: Lira, Ariel Jorge

- Cargo/s: Cargo No Docente, de Planta, categoría VII. Universidad Nacional de La Plata.
- Lugar de trabajo: Proyecto de Enlace de Bibliotecas (PrEBi) y Servicio de Difusión de la Creación Intelectual (SeDiCI).
- Facultad a la que pertenece: Facultad de Ingeniería
- Universidad: Universidad Nacional de La Plata
- Título: Licenciado en Informática

Apellido y Nombres: Islas, Joaquín Ezequiel

- Cargo/s: Beca de Entrenamiento Laboral. Universidad Nacional de La Plata.
- Lugar de trabajo: Proyecto de Enlace de Bibliotecas (PrEBi) y Servicio de Difusión de la Creación Intelectual (SeDiCI).
- Facultad a la que pertenece: Facultad de Ingeniería
- Universidad: Universidad Nacional de La Plata
- Título: Estudiante de Informática.

Apellido y Nombres: Pinto, Analia

- Cargo/s: Cargo No Docente de Planta, categoría VII. Universidad Nacional de La Plata.
- Lugar de trabajo: Proyecto de Enlace de Bibliotecas (PrEBi) y Servicio de Difusión de la Creación Intelectual (SeDiCI).
- Facultad a la que pertenece: Facultad de Ingeniería
- Universidad: Universidad Nacional de La Plata
- Título: Estudiante de Letras.

- Apellido y Nombres: Peloche, Silvia

- Cargo/s: Contrato de Obra. Universidad Nacional de La Plata.
- Lugar de trabajo: Proyecto de Enlace de Bibliotecas (PrEBi) y Servicio de Difusión de la Creación Intelectual (SeDiCI).
- Facultad a la que pertenece: Facultad de Ingeniería
- Universidad: Universidad Nacional de La Plata
- Título: Bibliotecólogo.



7. INFRAESTRUCTURA, EQUIPAMIENTO Y/O BIBLIOGRAFIA: Desarrollar según el esquema siguiente:

7.1 Disponible: Enumerar. Otros datos que estime convenientes.

PrEBi cuenta con dos espacios de trabajo, físicamente localizados en el Departamento de Ciencias Básicas de la Facultad de Ingeniería de la Universidad Nacional de La Plata, y distanciados uno de otro por menos de 40 metros. Estos espacios cuentan con los servicios básicos como teléfono, electricidad, calefacción, etc.

La infraestructura para acceso a información en línea está actualmente cubierta por una conexión a Internet dedicada con la Facultad de Informática de la UNLP, y otra conexión dedicada con la Facultad de Ingeniería de la UNLP.

En estos espacios, el PrEBi posee actualmente 8 equipos (PCs) de trabajo relativamente modernos, y 7 equipos servidores de distinta envergadura. Entre otros dispositivos, PrEBi posee actualmente 4 impresoras láser (2 de ellas adquiridas en los últimos meses), un escaner semi automático y dos puntos de acceso inalámbrico que ofrecen conectividad WiFi en los dos espacios de trabajo.

Finalmente, en PrEBi trabaja un equipo heterogéneo de personas compuesto por:

- desarrolladores de software, la mayoría de ellos han participado ya en otros proyectos de investigación y desarrollo en diversas áreas, y poseen cierta experiencia en desarrollos de sistemas basados en la web y tecnologías afines;
- referencistas y administradores de sistemas de áreas tan diversas como Humanidades, Bibliotecología e Informática, que utilizan constantemente sistemas en línea propios y externos

7.2 Necesario: Enumerar el equipamiento indispensable para la realización del proyecto que no se posee y costo aproximado.

Será necesario un equipo con buena capacidad de procesamiento, y especialmente de almacenamiento, para generar documentos sonoros, almacenarlos, recuperarlos y reproducirlos en línea. Este equipo deberá contar minimamente de un procesador multi-núcleo de altas prestaciones (Intel Xeon, Intel iCore, Intel AMD Opteron, AMD Phenom, o similares), una memoria RAM de al menos 4 GB, una capacidad de almacenamiento de 300 GB sobre uno o más discos de alta velocidad (10000 RPM), una interfaz de red de alta velocidad (100/1000 mbps) y hardware de sonido profesional (optimizado para tiempo real, múltiples canales de sonido y alta fidelidad).

7.3 Fuentes de información disponible y/o necesaria.

PrEBi es una fuente de información muy extensa, que permite a investigadores y docentes acceder a documentos científicos, libros, tesis, actas de congresos y patentes que se encuentran en bibliotecas de todo el mundo. Es el servicio bibliográfico de la UNLP al que acuden todos sus Docentes e Investigadores en caso de que la bibliografía no se encuentre en la unidad académica a la que pertenecen.

8. PRESUPUESTO DEL PROYECTO

8.1 Costo mínimo global necesario para llevar a cabo el proyecto

- Primer año: \$ 30.000
- Segundo año: \$ 30.000
- Tercer año: \$ 30.000
- Cuarto año: \$ 30.000

8.2 Fondos/Recursos disponibles

Monto	Fuente	Resolución
\$ 4300	CIC	279/03



\$ 6000	ISTEC	No corresponde
\$ 8000	Trabajos a terceros	No corresponde
\$ 10000	Recursos Propios	No corresponde

8.3 Fondos/Recursos en trámite

\$ 3000 (CIC)

US\$ 5000 (ISTEC -Metabuscador)

8.4 Explicitar la factibilidad del plan de trabajo propuesto con los recursos disponibles, en caso de no recibir financiamiento.

La realización del Proyecto es factible en un 50% por financiamiento de otras fuentes que posee el PrEBi.

9. PRESUPUESTO ESTIMADO PRELIMINAR (SUBSIDIO EROGACIONES CORRIENTES - UNLP)

Los fondos que puedan asignarse al presente proyecto serán exclusivamente utilizados para su realización de acuerdo con los objetivos y el plan de trabajo que constan en la presente solicitud.

DESCRIPCIÓN / CONCEPTO	IMPORTE 2010	IMPORTE 2011	IMPORTE 2012	IMPORTE 2013
BIENES DE CONSUMO	1000	1000	1000	1000
SERVICIOS NO PERSONALES (Viáticos, Pasajes, etc.)	3000	3000	3000	3000
EQUIPAMIENTO Y BIBLIOGRAFIA	5000	1000	5000	1000
EQUIPAMIENTO CIENTIFICO ESPECIFICOS				
EQUIPO DE COMPUTACIÓN	6000	10000	6000	8500
OTROS				1500
TOTAL	15000	15000	15000	15000

La información que detallo en esta solicitud es exacta y tiene carácter de **DELARACION JURADA**

FECHA: / /

.....
Firma del Director del Proyecto

10. AVAL DE LA UNIDAD ACADEMICA: De ser acreditado el presente proyecto se deja constancia que esta Unidad Académica otorga su conformidad para su realización en el ámbito de la misma.

FECHA: / /

Firma y Sello