

**CARRERA DEL INVESTIGADOR CIENTÍFICO Y
TECNOLÓGICO**
Informe Científico

PERIODO: 2006-2007

Legajo N°:

1. DATOS PERSONALES

APELLIDO: De Guisti

NOMBRES: Marisa Raquel

Dirección Particular: Calle: N°:

Localidad: La Plata CP: 1900 Tel:

Dirección electrónica: marisa.degiusti@sedici.unlp.edu.ar

2. TEMA DE INVESTIGACION

Investigación en Ingeniería Ontológica y Web Semántica. Desarrollos para la biblioteca digital y modelación de sistemas.

3. DATOS RELATIVOS A INGRESO Y PROMOCIONES EN LA CARRERA

INGRESO: Categoría: Asistente Fecha: Octubre 1987

ACTUAL: Categoría: Adjunto sin director desde fecha: Diciembre 2005

4. INSTITUCION DONDE DESARROLLA LA TAREA

Universidad y/o Centro: Proyecto de Enlace de Bibliotecas (PrEBi)

Facultad: Facultad de Ingeniería - UNLP

Departamento: Ciencias Básicas

Cátedra:

Otros:

Dirección: Calle: 49 y 115 N°: S/N

Localidad: La Plata CP: 1900 Tel: 0221-4233086

Cargo que ocupa: Director

5. DIRECTOR DE TRABAJOS. (En el caso que corresponda)

Apellido y Nombres:

Dirección Particular: Calle: N°:

Localidad: CP: Tel:

Dirección electrónica:

.....
Firma del Director (si corresponde)

.....
Firma del Investigador

6. EXPOSICION SINTETICA DE LA LABOR DESARROLLADA EN EL PERIODO.

6.1 **Título:** Métodos formales en digitalización compleja, recuperación de información, desarrollos para la biblioteca digital y modelación de sistemas.

6.1.1. **Digitalización y compresión:** Convertir imágenes en entes discretos de forma eficiente. Guardar la imagen con la menor cantidad posible de información. Se realizaron numerosas pruebas combinando diferentes formatos tanto vectoriales como mapas de bits: PNG, JPEG, JPEG2000, GIF, TIFF, BMP y distintos algoritmos de compresión con y sin pérdida: RLE, LZW, CCITT, ZIP y LZ77. En el caso de JPEG2000 se evaluó el algoritmo de compresión basado en la transformada discreta de wavelets o DWT. Las pruebas permitieron encontrar las combinaciones más cercanas al óptimo de acuerdo a los diferentes tipos de materiales digitalizados, p.e. texto completo, paisajes, documentos degradados por el tiempo, entre otros y a las capacidades del hardware subyacente en relación a memoria, disco, procesador, velocidad de red, etc. Se incluye Informe técnico en Anexo 1.

6.1.2. **Restauración:** mejorar imágenes degradadas en lo relativo a rotación, traslación, escalación e interpolación de imágenes. Se continuó trabajando en base a componentes conexas mejorando la implementación algorítmica: mayor velocidad en la ejecución y más información relevante. Las mejoras permitieron nuevos usos para la herramienta: detección del ángulo de inclinación del documento digitalizado (componentes conexas con código de cadenas derivados de esta implementación usados en un algoritmo recursivo) y posterior corrección de la inclinación utilizando un algoritmo de rotación basado en una transformada afín y conservando el centro de la imagen. Se realizó asimismo la eliminación de puntos aleatorios y de bordes negros. Se realizó la automatización completa y controlada de todo el proceso de limpieza organizada en un "pipeline". **Los resultados se exponen en 7.3.4.**

6.1.3. **Recuperación:** Se investigaron las posibilidades que ofrece el vector de características en cuanto a clasificación de objetos en imágenes, los aportes que los códigos de cadenas podían generar en estas técnicas y finalmente las ventanas de muestreo basadas en filtros gaussianos, técnicas basadas en la curvatura normalizada y coeficientes de tangentes firmados. A partir de la implementación, las muestras permitieron determinar grupos de características dependientes de la rotación, traslación y escala como también invariantes respecto de las distorsiones geométricas mencionadas. La implementación sobre el caso de texto manuscrito digitalizado permitió separar renglones, palabras y realizar la preclasificación de dichas palabras. **Los resultados se exponen en 7.1.1 y en 13.5.**

6.1.4. **Biblioteca digital:** Almacenaje y organización de importantes volúmenes de datos. Se realizaron pruebas sobre distintos motores de bases de datos: XML nativas eXist, Xindice, la base de datos embebida Berkeley DBXML y el motor de indexación de texto Apache Lucene junto al servidor de búsquedas Apache Solr sobre el mismo. Las pruebas efectuadas sobre eXist indicaron la imposibilidad de manejar eficientemente grandes volúmenes de información, sumado a la inestabilidad causada por las permanentes corrupciones de la estructura de base de datos subyacente. Asimismo, se vio que Xindice definitivamente no está diseñada para volúmenes elevados de información, no respeta los estándares y no es performante. La base de datos Berkeley dbxml resulto ser una de las alternativas mas prometedoras, pero debido a la existencia de un bug en el programa no se pudieron continuar las pruebas, las que se continuarán cuando se libere la nueva versión. Por ultimo, las experiencias obtenidas del desarrollo del Portal de Acceso a Información realizado para la UNNOBA expuestas en 8.1.1 sobre el motor de indexación de texto Apache Lucene, con Apache Solr como servidor de búsquedas fueron muy satisfactorias por su escalabilidad.

6.1.5. **En relación a los compromisos internacionales:** se comenzó y culminó el desarrollo de Celsius Network, que provee nuevas funcionalidades y permite la interconexión de todas las instancias de Celsius en una red virtual entre las partes. posibilitando la automatización de tareas de solicitud y envío. La nueva versión ya está disponible en el sitio <http://celsius.prebi.unlp.edu.ar> y en el Anexo 5 se incluyen los cursos dictados en LOJA.

6.1.6. **Modelación de sistemas:** Se simuló el funcionamiento de las principales actividades del software CelsiusNT: Servicios de Búsqueda, Servicios de Provisión y Cálculo de Estadísticas Distribuidas, los resultados de la modelización se exponen en 8.1.3.

7. TRABAJOS DE INVESTIGACION REALIZADOS O PUBLICADOS EN ESTE PERIODO

7.1 PUBLICACIONES. *Debe hacer referencia exclusivamente a aquellas publicaciones en las que haya hecho explícita mención de su calidad de Investigador de la CIC (Ver instructivo para la publicación de trabajos, comunicaciones, tesis, etc.). Toda publicación donde no figure dicha mención no debe ser adjuntada porque no será tomada en consideración. A cada publicación, asignarle un número e indicar el nombre de los autores en el mismo orden que figuran en ella, lugar donde fue publicada, volumen, página y año. A continuación, transcribir el resumen (abstract) tal como aparece en la publicación. La copia en papel de cada publicación se presentará por separado. Para cada publicación, el investigador deberá, además, aclarar el tipo o grado de participación que le cupo en el desarrollo del trabajo y, para aquellas en las que considere que ha hecho una contribución de importancia, deberá escribir una breve justificación.*

7.1.1. Manuscript Character Recognition. Overview of features for the Feature Vector. **Marisa R. De Giusti, María Marta Vila, Gonzalo Luján Villarreal.** Journal of Computer Science and Technology. Vol. 6.No. 2. October 2006. Pag. 92-98. Abstract. The image recognizing process requires the identification of every logical object that compose every image, which first implies to recognize it as an object (segmentation) and then identify which object is, or at least which is the most likely one from the universe of objects that can be recognized (recognition). During the segmentation process, the aim is to identify as many objects that compose the images as possible. This process must be adapted to the universe of all objects that are looked for, which can vary from printed or manuscript characters to fruits or animals, or even fingerprints. Once all objects have been obtained, the system must carry on to the next step, which is the identification of the objects based on the called universe. In other words, if the system is looking for fruits, it must identify univocally fruits from apples and oranges; if they are characters, it must identify the character "a" from the rest of the alphabet, and so on. In this document, the character recognition step has been studied. More specifically, which methods to obtain characteristics exist (advantages and disadvantages, implementations, costs). There is also an overview about the feature vector, in which all features are stored and analyzed in order to perform the character recognition itself.

El trabajo completo se expone en Anexo 3.

7.1.2. Storage of Simulation and Entities History in discrete models. **De Giusti Marisa Raquel, Lira Ariel Jorge , Villarreal, Gonzalo Luján.** Journal of Computer Science and Technology. Vol. 8. No. 1. April 2008. Pag. 41-45. **Abstract.** Simulation is the process of executing a **model**, that is a representation of a system with enough detail to describe it but not too excessive. This model has a set of entities an internal state, a set of input variable that can be controlled and others that cannot, a list of process that bind these input variables with the entities and one or more output values, which result from the execution of the processes. Running a model is totally useless if it can not be analyzed, which means to study all interactions among input variables, model entities and their weight in the values of the output variables. In this work we consider Discrete Event Simulation, which means that the status of the system variables being simulated change in a countable set of instants, finite or countable infinite. Simulation programming languages provide a big range of tools for analysis of the results, for generation and execution of experiments and to perform complex analysis (such as Analysis of Variance). This is usually enough for common analysis, but many times more detailed information is required. In many circumstances it is desirable to have all run of simulations stored in order to make further analysis or comparisons between simulations, many time after they have been run. Most simulation environments provide reports and logs (journals) and permit to save them in text or formated files, which include all results and some aspects of the run itself. In this work, we propose to store not only all simulation results, but all

simulation history. This implies to store all permanent entities of the simulation, and all changes they have undergone along the simulation times. But it not only limits to permanent entities, since we also store the temporary ones, which are created and destroyed anytime in the simulation and which existence is subject to the execution of the simulation and the state of all model variables. This development takes the same syntax of GPSS language and the way it handles entities, to develop a simple interpreter and a tool that considers a subset of GPSS entities and permits to execute simulations with them. Besides, this tool permits to search and select entities for each simulation and displays their evolution along the simulation.

El trabajo completo se incluye en Anexo 6.

En las publicaciones precedentes fui responsable en el desarrollo del tema y dirección de recursos humanos que trabajaron en mismo.

7.2 TRABAJOS EN PRENSA Y/O ACEPTADOS PARA SU PUBLICACIÓN. *Debe hacer referencia exclusivamente a aquellos trabajos en los que haya hecho explícita mención de su calidad de Investigador de la CIC (Ver instructivo para la publicación de trabajos, comunicaciones, tesis, etc.). Todo trabajo donde no figure dicha mención no debe ser adjuntado porque no será tomado en consideración. A cada trabajo, asignarle un número e indicar el nombre de los autores en el mismo orden en que figurarán en la publicación y el lugar donde será publicado. A continuación, transcribir el resumen (abstract) tal como aparecerá en la publicación. La versión completa de cada trabajo se presentará en papel, por separado, juntamente con la constancia de aceptación. En cada trabajo, el investigador deberá aclarar el tipo o grado de participación que le cupo en el desarrollo del mismo y, para aquellos en los que considere que ha hecho una contribución de importancia, deberá escribir una breve justificación.*

7.3 TRABAJOS ENVIADOS Y AUN NO ACEPTADOS PARA SU PUBLICACION. *Incluir un resumen de no más de 200 palabras de cada trabajo, indicando el lugar al que han sido enviados. Adjuntar copia de los manuscritos.*

7.3.1. Simulation framework for teaching in modeling and simulation areas. **De Giusti, Marisa Raquel**; Lira, Ariel Jorge; Villarreal, Gonzalo Lujan. Enviado a: European Journal of Engineering Education EJEE.

Abstract. Simulation is the process of executing a model which describes a system with enough detail. This model has its entities, an internal state, its input and output variables and a list of processes binded to these variables. Teaching a simulation language like GPSS is always a challenge, because of the way it executes the models and the abstraction level it can achieve, quite different compared to most programming languages (well known by most students). This paper presents an open source simulation framework that implements a subset of entities of GPSS, which could help students to improve the understanding of this language. This tool also stores all entities of simulations in every single simulation time, which is very useful for debugging simulations, but also for getting a detailed history of all entities in the simulations, knowing exactly how they have behaved in every simulation time.

El trabajo completo se incluye en Anexo 7.

7.3.2. Servicio de Difusión de la Creación Intelectual. **Prof. Ing. Marisa R. De Giusti.** Enviado a: Revista Interamericana de Bibliotecología. El mismo trabajo bajo el título en inglés "The Intellectual Creation Difusion Service" fue enviado a D-Lib Magazine.

Resumen: El Servicio de Difusión de la Creación Intelectual (SeDiCI) de la Universidad Nacional de La Plata (UNLP) nació con el objetivo prioritario de socializar el conocimiento

generado en las diferentes áreas académicas de nuestra Universidad, con el fin de devolver a la comunidad los esfuerzos destinados a la Universidad Pública. SeDiCI realiza su trabajo en conjunto con la Dirección de Propiedad Intelectual de la UNLP en el caso que el o los autores deseen registrar su obra. La Dirección y SeDiCI se encargan de las tramitaciones pertinentes de modo de simplificar la tarea de quienes desean difundir su trabajo. La plataforma de software sobre la que se sustenta este servicio, denominada Celsius DL, fue íntegramente desarrollada por el equipo informático de PrEBi. En este trabajo se describe tanto el servicio como las posibilidades que ofrece esta plataforma desde el punto de vista de software: plataforma, arquitectura de la información, procesamiento de contenidos, servicios a usuarios, iniciativa OAI y servicios especiales. El trabajo completo en sus dos versiones (castellano e inglés) se incluye en Anexo 8.

7.3.3. Open modeling and simulation framework for evolutive analysis. **ENG. PROF.. De Giusti Marisa Raquel**; Lic. Lira Ariel Jorge; Lic. PROF. Villarreal, Gonzalo Luján. Enviado a: Transactions of the Society for modeling and Simulation International.

Abstract: This paper presents an open source simulation framework that implements a subset of entities of GPSS, and at the same time has the ability to store all entities of simulations in every simulation time, which is very useful for debugging, but also for having a detailed history of all entities in the simulations, knowing exactly how they have behaved in every simulation time. Programmers can both extend the model and create new analysis tools for studying the evolution of any entity inside the simulation, the whole simulation or even to compare different simulation runs.

El trabajo completo se incluye en Anexo 9.

7.3.4. Use of Connected Components for automatic digitalized documents restoration. **Marisa R. De Giusti**; María Marta Vila; Villarreal Gonzalo Luján. Enviado a: International Journal on Digital Libraries.

Abstract: The process of digitalizing documents has been simplified through the advance of technology and the development of devices for images capture (scanners, digital cameras), which has provoke the constant digitalization and on line movement of huge amounts for documents. This has generated great advantages in many environments, thanks to the ability of preserve very large document databases and transfer them almost immediately through digital mediums, allowing to a huge number of people the access to information that, in other way, would have been impossible. A clear example of this statement is given inside the Digital Library frame, where PrEBi and SeDiCI projects are strongly involved generating services, developing software tools and carrying out many research lines related to common subjects.

Despite all mentioned advances, the digitalization of documents keeps generating some problems originated from human or technical mistakes, and that result in images with redundant information, hard-to-read documents, extremely big files, etc. Among these problems we can stand out:

Documents with black or diffuse borders: It usually happens when the original documents are in poor conditions, or when the scanner has not been well closed. This implies bigger documents, heavier digital files, more ink consumption and some kind of difficulty for reading them.

Random dots along the document: It might happen because of original documents poor quality as well as problems with the scanner; this generates from dirty or untidy documents to hard-to-read texts due to the points among the text (sandy effect).

Images slant: due to distraction of the operator, automatic scanners malfunction or lack of space when digitalizing big documents (such as large books), many documents present some kind of slop compared to the base line. This slop might affect the reader, specially if it is grater than 5 degrees and under 85 degrees.

In this article, we will present a optimized and automatized mechanism for improve digitalized documents that present some or even all problems described below. First we introduce briefly the concept of Connected Components, which are essential for the whole process. Then, we carry on with a technique for slant detection of document as well as a

basic algorithm for image rotation. Next, we explain an easy method for detection and elimination of black borders and random points. Last, we propose an automated mechanism that integrates all these tasks in a big and organized restoration process. El documento completo se incluye en Anexo 2.

7.4 TRABAJOS TERMINADOS Y AUN NO ENVIADOS PARA SU PUBLICACION. *Incluir un resumen de no más de 200 palabras de cada trabajo.*

7.5 COMUNICACIONES. *Incluir únicamente un listado y acompañar copia en papel de cada una. (No consignar los trabajos anotados en los subtítulos anteriores).*

7.6 INFORMES Y MEMORIAS TECNICAS. *Incluir un listado y acompañar copia en papel de cada uno o referencia de la labor y del lugar de consulta cuando corresponda.*

8. TRABAJOS DE DESARROLLO DE TECNOLOGÍAS.

8.1 DESARROLLOS TECNOLÓGICOS. *Describir la naturaleza de la innovación o mejora alcanzada, si se trata de una innovación a nivel regional, nacional o internacional, con qué financiamiento se ha realizado, su utilización potencial o actual por parte de empresas u otras entidades, incidencia en el mercado y niveles de facturación del respectivo producto o servicio y toda otra información conducente a demostrar la relevancia de la tecnología desarrollada.*

8.1.1. Pruebas con XML: Las experiencias obtenidas del desarrollo del Portal de Acceso a Información realizado para la UNNOBA fueron las más satisfactorias. La base de datos utilizada en este caso fue el motor de indexación de texto Apache Lucene, con Apache Solr como servidor de búsquedas. El volumen de registros almacenados no fue muy elevado, pero sin embargo suponemos que el uso de esta plataforma es extremadamente escalable. Los tiempos de búsqueda logrados para alrededor de 300.000 registros estuvo en el orden de los 10 milisegundos, por lo que es de esperar que se logre indexar millones de registros, y recuperarlos por medio de búsquedas en menos de un segundo. El proceso de prueba sobre Lucene-Solr es costoso debido a que los registros obtenidos previamente por el proceso de harvesting, deben ser transformados a una representación intermedia en XML propia de Solr, para luego ser agregados a la base de datos. Este problema se presenta también en el momento de recuperación de los datos, ya que la base de datos devuelve los resultados en un formato estático que debe ser transformado para que pueda ser procesado como si fuera un registro Dublin Core típico en XML.

8.1.2. Celsius Network: Se finalizó la primera etapa del desarrollo de este software de gestión de pedidos de usuarios del servicio de provisión y búsqueda que ofrece el PrEBi. Esta etapa abarcó el diseño y programación de este software, la instalación a nivel local, una etapa de pruebas y ajustes a partir de los resultados y opiniones de los operadores de Celsius de la UNLP, el desarrollo de un software de instalación y de la documentación que lo acompaña y finalmente la liberación oficial en el sitio web del software Celsius. A sólo 5 días de su liberación oficial, el sitio Celsius ya ha registrado más de 100 descargas tanto de este aplicativo como de la documentación adjunta. Es de destacar la relevancia internacional de este proyecto ya que Celsius es un software utilizado en diferentes países de América y aceptado como standard por el Consorcio Iberoamericano para Educación en Ciencia y Tecnología (ISTEC).

Simulación NT. El uso de transacciones como representación de los Pedidos en el caso de los Servicios de Búsqueda y Provisión permitió simular con toda su variabilidad el sistema global de pedidos. En relación a las tres funcionalidades bajo análisis, la utilización de entidades STORAGE, QUEUE y SAVEVALUE permitió obtener datos estadísticos referidos a cantidades y tiempos promedios, y la utilización de los

Operadores. En el caso del Cálculo de Estadísticas Distribuidas, esto es estadísticas disparadas desde un sitio celsius recabando datos de todos los sitios la simulación realizada nos permite medir y determinar el efecto producido por situaciones anómalas tales como la caída de servidores, enlentecimientos de las redes y timeouts que quedan almacenados en los tiempos de tránsito de las diferentes transacciones.

Se incluye informe técnico en Anexo 11.

8.1.3. Celsius Imago. Las líneas de investigación relacionadas con procesamiento de imágenes digitales, restauración de documentos y reconocimiento de patrones permitieron el desarrollo de una herramienta de software donde se pueden realizar pruebas aisladas sobre documentos, o combinaciones de pruebas complejas. Esta herramienta permite también importación de documentos PDF multipage en imágenes separadas, y procesamiento de directorios completos de manera totalmente automática, lo cual permitió realizar pruebas masivas sobre imágenes obtenidas a partir de documentos impresos con claros signos de deterioro o incorrectamente digitalizados.

Los resultados de este desarrollo fueron expuestos en 7.1.1, 7.3.4, 13.5.

8.1.4. Portal de Acceso a Información. Este desarrollo, solicitado por la Universidad del Noroeste de la Provincia de Buenos Aires (UNNOBA), ofrece a los alumnos de dicha Institución la posibilidad de realizar búsquedas en diversos repositorios de documentos. Para ello, el portal web interactúa de manera totalmente transparente con SeDiCI, del cual obtiene resultados de búsqueda; también realiza búsquedas de manera asíncrona en un amplio conjunto de registros, obtenidos mediante la Iniciativa Open Archives (OAI) de un listado preseleccionado de repositorios en línea. Adicionalmente, este portal provee a los visitantes de diversos enlaces de interés, agrupados y clasificados por área temática, cátedra o incluso carrera. Los docentes pertenecientes a cada área puede gestionar esta información como crean más conveniente, mediante una compleja herramienta de administración web provista dentro del mismo portal. Entre otras, se destacan las facilidades de gestión y edición de noticias en línea, la suscripción a feeds (RSS) y las opciones de búsqueda avanzada. Finalmente, cabe destacar que este desarrollo ha sido pensado de manera totalmente modular, haciendo especial énfasis en la posibilidad de realizar búsquedas sobre otros repositorios en línea como ser el catálogo de la biblioteca de la Universidad.

Este desarrollo puede verse en: <http://server-unnoba.prebi.unlp.edu.ar>.

8.2 PATENTES O EQUIVALENTES. Indicar los datos del registro, si han sido vendidos o licenciados los derechos y todo otro dato que permita evaluar su relevancia.

8.3 PROYECTOS POTENCIALMENTE TRANSFERIBLES, NO CONCLUIDOS Y QUE ESTAN EN DESARROLLO. Describir objetivos perseguidos, breve reseña de la labor realizada y grado de avance. Detallar instituciones, empresas y/o organismos solicitantes.

8.4 OTRAS ACTIVIDADES TECNOLÓGICAS CUYOS RESULTADOS NO SEAN PUBLICABLES (desarrollo de equipamientos, montajes de laboratorios, etc.).

8.5 Sugiera nombres (e informe las direcciones) de las personas de la actividad privada y/o pública que conocen su trabajo y que pueden opinar sobre la relevancia y el impacto económico y/o social de la/s tecnología/s desarrollada/s.

8.5.1. Dr. Ramiro Jordan Mealla. University of New Mexico, EEUU.
mrjordan@istec.org

8.5.2. Dr. Gregory Jason Randall. randall@iee.edu.uy

8.5.3. Dr. Mauro Sergio Miskulin, Universidad de Campinas, Brasil, y University of New Mexico, EEUU. mauromiskulin@terra.com.ar

8.5.4. Dra. Martina Trucco. Hewlett Packard. martina.trucco@hp.com

8.5.5. Dr. Wilfrido Moreno. University of South Florida. moreno@eng.usf.edu

8.5.6. Dr. Roberto Murphy Arteaga. Instituto Nacional de Astrofísica, Óptica y Electrónica. Puebla, Mexico. murphy@inaoep.mx

Hasta aquí todas las personas mencionadas conocen de mi trabajo en relación con las plataformas Celsius desarrolladas en PrEBi.

8.5.7. Dra. Mónica Luque. Directora del Instituto de Estudios Avanzados para las Américas (INEAM) de la Organización de los Estados Americanos (OEA). monica_luque@ciudad.com.ar

Puede opinar sobre mis trabajos de desarrollo en el área de Bibliotecas Digitales.

8.5.8. Prof. Alejandra Rodriguez. Programa Nacional de "Educación en Contextos de Encierro". Ministerio de Educación, Ciencia y Tecnología. arodriguez@me.gov.ar

Puede opinar sobre el desarrollo ad hoc realizado para la encuesta sobre "Bibliotecas en prisiones" que se llevara adelante en PrEBi.

8.5.9. Dr. Eduardo Tavani. CIC-CETMIC. etavani@netverk.com.ar.

Puede opinar sobre el desarrollo del sitio WEB del IBS2009.

8.5.10. Dr. Fernando Tauber. Secretario General, Universidad Nacional de La Plata. ftauber@ciudad.com.ar

Puede opinar sobre los desarrollos que llevo adelante en la UNLP.

8.5.11. Dra. Silvia Gordillo. Miembro del Consejo Directivo del Laboratorio de Investigación y Formación en Informática Avanzada (LIFIA). silvia.gordillo@lifa.info.unlp.edu.ar

Puede opinar sobre los desarrollos que dirijo en el área de simulación.

9. SERVICIOS TECNOLÓGICOS. Indicar qué tipo de servicios ha realizado, el grado de complejidad de los mismos, qué porcentaje aproximado de su tiempo le demandan y los montos de facturación.

9.1. Provisión/Alojamiento de Documentación Científica-Académica.

En el marco del proyecto PrEBi se proveyó durante el período 2006-2008 de 6997 documentos equivalentes a 88.677 hojas impresas.

Actualmente, PrEBi tiene registrados mas de 1500 usuarios Investigadores y Académicos provenientes de 15 Unidades Académicas de la UNLP, CIC y CONICET.

Asimismo y en el marco de intercambio de documentación de la red DLL de ISTEAC se proveyeron al exterior cerca de 1200 documentos que equivalen a 11433 páginas enviadas. Por consultas acceda a <http://prebi.unlp.edu.ar/celsius> y visite estadísticas abiertas.

En el marco del Proyecto SeDiCI, se alojaron desde su puesta en marcha en Diciembre de 2003, un total de 1024 documentos, compuestos por 229 Tesis de Grado, 261 de Postgrado, 8 Disertaciones, 17 Series Documentales compuestas por 716 artículos, 2 Series Pictóricas, 6 EBook, 6 Libros, 22 Producciones y Presentaciones Académicas y 61 producciones artísticas independientes. Las investigaciones relacionadas con la obtención de referencias a texto completo han permitido obtener y dar acceso a más de 700.000 referencias a documentos externos los cuales son en un 40% Tesis de Postgrado en las áreas conceptuales mas variadas. En dicho servicio, han sido inscriptos mas de 2700 usuarios, cabe destacar en este punto que la contribución de documentación es completamente voluntaria. El portal web de SeDiCI ha recibido hasta el momento más de 214.946 visitas desde más de 90 países de todo el mundo, entre los que se incluyen Argentina, Estados Unidos, España, Israel, Japón, Italia, Méjico, Brasil, Taiwan, China, Canadá y muchos otros. A través de webservices (Ref.13.2 y 13.3), SeDiCI ha atendido hasta el momento más de 60.631 consultas de otras bibliotecas. Mas información en <http://sedici.unlp.edu.ar/>

La gestión de ambos proyectos en los cuales participan un total de 13 personas, insume el 50% de mi dedicación en CIC incluyendo actividades vinculadas a investigación para orientar los desarrollos de software (especialmente Celsius Network, Celsius DL , Celsius Imago, los sitios expuestos: Ministerio de Educación, IBS2009, etc. La Dirección y formación de recursos humanos, la coordinación de las distintas actividades, la gestión de los recursos y la creación de políticas de funcionamiento, reuniones de presentación de los proyectos y otras tareas afines.

9.2. Servicios tecnológicos a Terceros.

9.2.1. Portal de Acceso a Información UNNOBA. Desarrollo de una herramienta para la Universidad del Noroeste de la Provincia de Buenos Aires, que permite a los alumnos acceder a diversas fuentes de información, realizar búsquedas, obtener bibliografía, enlaces preseleccionados, entre otras.

9.2.2. Sistema Web Librería Lenzi. Desarrollo de un software de Gestión integral de la librería.

9.2.3. Celsius Network. Desarrollo de la nueva versión del software Celsius, con facilidades para interconectar las instancias de Celsius alrededor del mundo, compartiendo información y automatizando procesos.

9.2.4. Sitio web Libros Lenzi. Sitio web con posibilidades de búsquedas en catálogo. Este desarrollo se realizó en dos etapas: en la primera se hizo un sitio totalmente estático mientras que en la segunda se incorporaron las búsquedas. <http://www.libroslenzi.com.ar> .

9.2.5. Sitio web IBS2009. Sitio web para publicación de noticias, novedades, fechas límites y demás información del 18th International Biohydrometallurgy Symposium (IBS2009). <http://www.ibs2009.org.ar> .

9.2.6. Encuesta Ministerio de Educación y UNESCO. Encuesta en línea para proyecto de Bibliotecas en ámbitos de encierro, llevado adelante por el Ministerio de Educación de la Nación y la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO). <http://prebi.unlp.edu.ar/encuesta> .

Estos desarrollos insumieron un 10% de mi tiempo en tareas administrativas y de gestión de los proyectos. Se adjunta una planilla con el detalle sobre cada desarrollo en Anexo 12.

10. PUBLICACIONES Y DESARROLLOS EN:

10.1 DOCENCIA

10.2 DIVULGACIÓN

10.2.1. 21 al 31 de agosto de 2007. ExpoUniversidad.

Se expusieron 3 posters en la ExpoUniversidad, realizada en el Pasaje Dardo Rocha de la Ciudad de La Plata. En estos pósters se presentaron El Proyecto de Enlace de Bibliotecas (PrEBI), el Servicio de Difusión de la Creación Intelectual (SeDiCI) y la familia de software Celsius: Celsius Network, Celsius Imago y Celsius Bloodhound.

10.2.2. 7 de mayo de 2008. Feria del Libro.

Se realizó una conferencia en la Feria del Libro, titulada "Posibilidades que ofrece el Servicio de Difusión de la Creación Intelectual de la Universidad Nacional de La Plata". Se incluye presentación en Anexo 14.

11. DIRECCION DE BECARIOS Y/O INVESTIGADORES. *Indicar nombres de los dirigidos, Instituciones de dependencia, temas de investigación y periodos.*

11.1 Director del Becario de Postgrado del CONICET Lic. Gonzalo Luján Villarreal. Tema: “Desarrollo de herramientas para simulación y procesamiento de imágenes”, desde Abril de 2008 y hasta la fecha.

11.2 Director del Analista de Computación Ariel Sobrado, con un Cargo de Planta No Docente de la Universidad Nacional de La Plata. Desde Mayo de 2003 hasta la fecha. Dedicado a tareas de diseño de software y mantenimiento en el Proyecto de Enlace de Bibliotecas (PrEBi) y en el Servicio de Difusión de la Creación Intelectual (SeDiCI).

11.3 Director de la Analista de Computación María Marta Vila, con un Cargo de Planta No Docente de la Universidad Nacional de La Plata. Desde Octubre de 2003 hasta la fecha. Dedicado a tareas de diseño de software y mantenimiento en el Proyecto de Enlace de Bibliotecas (PrEBi) y en el Servicio de Difusión de la Creación Intelectual (SeDiCI).

11.4 Director del estudiante de Letras Carlos J. Nusch, con un Cargo de Planta No Docente de la Universidad Nacional de La Plata. Desde Mayo de 2004 hasta la fecha. Dedicado a las tareas de búsqueda bibliográfica en el Proyecto de Enlace de Bibliotecas (PrEBi).

11.5 Director de la Estudiante de Diseño Gráfico Gisele Jaquenod, con un Cargo de Planta No Docente de la Universidad Nacional de La Plata. Desde Marzo de 2004 hasta la fecha. Dedicada a las tareas de diseño web de los portales PrEBi y SeDiCI.

11.6 Director del Estudiante de Informática Esteban C. Fernández, con un Cargo de Planta No Docente de la Universidad Nacional de La Plata. Desde Mayo de 2004 hasta la fecha. Dedicado a las tareas de digitalización de tesis y documentos científicos en SeDiCI y tareas de búsqueda bibliográfica en el Proyecto de Enlace de Bibliotecas (PrEBi).

11.7 Director del Estudiante de Informática Ariel Jorge Lira con un Contrato de Servicio con la Universidad Nacional de La Plata. Desde Marzo de 2006 hasta la fecha. Dedicado a tareas de diseño de software y mantenimiento en el Proyecto de Enlace de Bibliotecas (PrEBi) y en el Servicio de Difusión de la Creación Intelectual (SeDiCI).

11.8 Director del Estudiante de Derecho Emanuel Sosa, con una Beca de Experiencia Laboral de la Universidad Nacional de La Plata. Desde Julio de 2003 hasta la fecha. Dedicado a tareas de búsqueda bibliográfica en el Proyecto de Enlace de Bibliotecas (PrEBi).

11.9 Director del Licenciado en Economía Manrique Lucio Altavista, con una Beca de Experiencia Laboral de la Universidad Nacional de La Plata. Desde Julio de 2003 hasta Febrero de 2007. Dedicado a tareas de búsqueda bibliográfica en el Proyecto de Enlace de Bibliotecas (PrEBi).

11.10 Director del estudiante de Administración de Empresas Gustavo Jaquenod, con una Beca de Experiencia Laboral de la Universidad Nacional de La Plata. Desde Mayo de 2004 hasta la fecha. Dedicado a las tareas de búsqueda bibliográfica en el Proyecto de Enlace de Bibliotecas (PrEBi).

11.11 Director del estudiante de Informática Fernando Gabriel Inafuku con una Beca de Experiencia Laboral de la Universidad Nacional de La Plata. Desde Mayo de

2006 hasta Julio de 2007. Dedicado a tareas de diseño de software y mantenimiento en el Proyecto de Enlace de Bibliotecas (PrEBi) y en el Servicio de Difusión de la Creación Intelectual (SeDiCI).

11.12 Director del estudiante de Informática Néstor Fabián Oviedo con una Beca de Experiencia Laboral de la Universidad Nacional de La Plata. Desde Noviembre de 2007 hasta la fecha. Dedicado a tareas de diseño de software y mantenimiento en el Proyecto de Enlace de Bibliotecas (PrEBi).

11.13 Director del Estudiante de Letras Analia Pinto, con una Beca de Experiencia Laboral de la Universidad Nacional de La Plata. Desde Mayo de 2008 hasta la fecha. Dedicado a las tareas de digitalización de tesis y documentos científicos en SeDiCI.

11.14 Director del estudiante de Informática Matías Reynoso con una Beca de Experiencia Laboral de la Universidad Nacional de La Plata. Desde Mayo de 2008 hasta la fecha. Dedicado a las tareas de búsqueda bibliográfica en el Proyecto de Enlace de Bibliotecas (PrEBi).

12. DIRECCION DE TESIS. *Indicar nombres de los dirigidos y temas desarrollados y aclarar si las tesis son de maestría o de doctorado y si están en ejecución o han sido defendidas; en este último caso citar fecha.*

12.1 Co-Director de la Tesis de Lic. del Lic. Gonzalo Luján Villarreal y Ariel Jorge Lira: "Entorno abierto de desarrollo y ejecución de simulaciones discretas para análisis evolutivo de entidades mediante persistencia". Presentada el 19 de Marzo de 2008 ante las autoridades del Jurado previsto en la Facultad de Informática de la Universidad Nacional de La Plata. Calificación obtenida Diez (10) puntos. La mencionada tesis se encuentra en: <http://sedici.unlp.edu.ar?id=ARG-UNLP-TDG-0000000032>

13. PARTICIPACION EN REUNIONES CIENTIFICAS. *Indicar la denominación, lugar y fecha de realización, tipo de participación que le cupo, títulos de los trabajos o comunicaciones presentadas y autores de los mismos.*

13.1 Evento: 4º Simposium Internacional de Bibliotecas Digitales: "Información digital al servicio de la sociedad".

Lugar: Málaga, España.

Organizado por Universidad de Málaga, ISTECS

Fecha: 21 al 23 de Junio 2006.

Carácter de la participación: Miembro del Comité Científico.

Se incluye el certificado en el Anexo 15.

13.2 Evento: 4º Simposium Internacional de Bibliotecas Digitales: "Información digital al servicio de la sociedad".

Lugar: Málaga, España.

Organizado por Universidad de Málaga, ISTECS

Fecha: 21 al 23 de Junio 2006.

Carácter de la participación: Expositor.

Se incluye el certificado en el Anexo 15.

13.3 Tecnologías para propagar los contenidos de una Biblioteca Digital. Experiencias desde SeDiCI. **Prof. Ing. De Giusti, Marisa Raquel**; Marmonti, Emiliano; A.C. Sobrado, Ariel; A.C. Vila, María Marta; A.C. Villarreal, Gonzalo Luján. Anales del 4to. Simposio Internacional de Bibliotecas Digitales. Málaga, España. 21 al 23 de Junio de 2006.

Resumen de la presentación :

En el panorama actual de las Bibliotecas Digitales puede observarse la creciente importancia de admitir protocolos que permitan la provisión de servicios a otras interfaces que accedan directamente, con las consideraciones apropiadas de seguridad, oportunidad y pertinencia, a los repositorios administrados por la Biblioteca Digital.

Dichas interfaces pueden ser, por ejemplo, clientes RSS para novedades en los contenidos o un medio para Diseminación Selectiva, metabuscadores, OAI Service Providers, portales web internos de las mismas unidades académicas que acceden a la biblioteca digital para ofertar a sus propios usuarios información residente en los repositorios, o incluso plataformas de e-learning que precisen integrar la información administrada como parte de los contenidos o como bibliografía a consultar.

Desde el año 2004, el Servicio de Difusión de la Creación Intelectual (SeDiCI-UNLP), brinda la posibilidad de devolver parte de sus acervos electrónicos a través de webservices basados en la tecnología SOAP. Sin embargo, dicha implementación no se halla basada en ningún estándar para acceder y consultar los recursos digitales administrados, otras plataformas tales como Fedora, por su misma arquitectura admiten naturalmente una variedad de “conexiones” al repositorio de objetos digitales administrados.

SeDiCI a su vez, es OAI Data y Service provider, siendo que buena parte del material colectado puede ser de utilidad a las bibliotecas al momento que los usuarios realizan consultas en sus respectivos catálogos locales en la UNLP. En un universo de usuarios que acceden a Google como fuente primaria de información, las Bibliotecas intentan día a día “competir” contra los metabuscadores y lograr que sus usuarios accedan al catálogo local como primer opción, SeDiCI tiene la responsabilidad de proveerles en forma transparente y sumamente performante ofertas de material accesible libremente y en línea, tanto proveniente del acervo local administrado como de la información referencial colectada a través del protocolo OAI. Una solución genérica para la atención tanto de los requerimientos de OPAC’s internos como de interfaces externas a la UNLP puede ser implementada mediante la adición de una interfaz de consulta a los acervos electrónicos de SeDiCI, a través del estándar SRU/W. Dicho estándar, que replica las características de avanzada del protocolo Z39.50 en cuanto a acceso y facilidades de búsqueda, está actualmente siendo adoptado por diversas plataformas de Bibliotecas Digitales (como es el caso de DSpace).

El objetivo de esta ponencia será presentar un estudio de avance junto con los resultados, así como también el impacto de la implementación de SRU/W sobre el repositorio de documentos locales de SeDiCI. También se estudiará el grado de complejidad basado en la bibliografía existente acerca de la implementación de un gateway SRU/W que permita acceder a la información referencial obtenida a través de OAI y el grado de impacto esperado en los OPAC’s de la Universidad Nacional de La Plata.

La idea rectora de esta clase de servicios es la de fortalecer los “puntos simples de acceso” y reducir la cantidad de sitios a los que los usuarios deben acceder para hallar información académica generada en sus propias instituciones. En el marco de dicha tendencia, se debatirá si la estrategia en software de administración de Bibliotecas Digitales apunta a simplificar o incluso desaparecer la interfaz web autónoma y transformarse en un conjunto de servicios-web que son usados desde los diversos portales internos o externos en forma transparente.

Se incluye el trabajo completo en Anexo 16.

13.4 Celsius Network. Interconexión entre portales de Referencia a través de webservices. **Prof. Ing. De Giusti, Marisa Raquel**; Marmonti, Emiliano; A.C. Sobrado, Ariel; A.C. Vila, María Marta; A.C. Villarreal, Gonzalo Luján. Anales del 4to. Simposio Internacional de Bibliotecas Digitales. Málaga, España. 21 al 23 de Junio de 2006.

Resumen:

El software Celsius, creado en PrEBi UNLP, permite realizar la gestión de pedidos de documentación de los usuarios de dicho servicio de referencia. Esta gestión comprende

el registro de los datos por parte del usuario, la normalización de los mismos y del documento en sí por parte de los operadores, el registro de las distintas instancias por las cuales el pedido va atravesando y finalmente, la entrega del material solicitado, ya sea en papel o descargado desde el sitio personalizado de cada usuario. Los usuarios de dicho servicio pueden hacer uso de toda esta información, pudiendo realizar un seguimiento de cada uno de sus pedidos y accediendo a la gran cantidad de estadísticas que el software genera de manera totalmente automática y constantemente actualizadas. Estas estadísticas también pueden ser utilizadas como indicador de gestión del intercambio, gracias a la gran variedad de datos que las mismas ofrecen, como ser cantidad de documentos solicitados, tipo de documentación más requerida, destino de los pedidos de usuario (por país, institución o dependencia), y muchas otras.

Celsius ha sido adoptado por el consorcio ISTECA como la plataforma estándar para la iniciativa LibLink y se encuentra actualmente instalado en 8 países de América, contando con 24 instancias conocidas y funcionales en constante avance y difusión. El equipo de desarrollo de Celsius está trabajando actualmente en la estabilización de las nuevas instalaciones del software, así como también en la actualización de las instancias anteriores a la versión actual (1.6).

Actualmente cada operador de Celsius debe procesar el pedido de su usuario local, completar la información faltante, normalizarla, seleccionar la biblioteca destino y en caso que la misma posea Celsius funcional (como sucede en la mayor parte de los casos), debe repetir el registro de la información del pedido como un usuario externo. Esto representa una enorme duplicación de tareas y un factor adicional para la pérdida de calidad e incremento en el factor de error para la información generada.

El objetivo de esta comunicación es la realización de un estudio detallado de las distintas opciones para maximizar la eficiencia del servicio, obteniendo beneficios tanto para los usuarios finales como para el consorcio en general. Se analizarán las ventajas y desventajas de un único Celsius centralizado, de la exportación e importación de la información de los pedidos, y finalmente la posibilidad actualmente seleccionada, que implica el establecimiento de una red P2P en la cual las distintas instancias de Celsius se comunicarán entre sí por medio de un protocolo estándar de intercambio de información, como los webservices a través del protocolo SOAP. El estudio de esta última solución abarca el análisis de los distintos casos y los beneficios que esta comunicación transparente conllevan. Finalmente, se analizará la factibilidad de realizar un único portal a nivel global, cuyo objetivo será la generación de estadísticas centralizadas para todas las instancias de Celsius.

Se incluye póster en Anexo 17.

13.5 Preservación del Patrimonio Cultural – Digitalización y reconocimiento de documentos manuscritos: Extracción de Características. **Prof. Ing. De Giusti, Marisa Raquel**; A.C. Vila, María Marta; A.C. Villarreal, Gonzalo Luján. Anales del 4to. Simposio Internacional de Bibliotecas Digitales. Málaga, España. 21 al 23 de Junio de 2006.

Resumen de la presentación.

Las nuevas tecnologías han permitido la digitalización de una gran cantidad de piezas documentales que se encuentran en bibliotecas y archivos de todo el mundo, lo cual ha facilitado el acceso a un mayor número de personas a dichos documentos sin alterar o deteriorar los documentos originales.

Al ser digitalizados, los documentos son mejor preservados, pero la búsqueda y el acceso a la información allí contenida es un proceso lento, secuencial y por consiguiente, altamente ineficiente. Este problema de acceso a tales documentos presenta dos posibles soluciones:

- la catalogación de los documentos
- su indexación automática

La primera es una solución factible pero que demanda una gran cantidad de tiempo, ya que el proceso de catalogación requiere gran esfuerzo y dedicación. Sumado a esto,

es necesaria la intervención de seres humanos durante todo el proceso de catalogación, lo que se traduce en altos costos económicos.

La segunda solución propone el desarrollo de una aplicación capaz de interpretar esos documentos, procesándolos y reconociendo automáticamente los textos que cada documento contiene, que posteriormente pueden ser indexados utilizando base de datos, y permitiendo un rápido acceso a la información. Esto implica la utilización de sistemas de reconocimiento de patrones, así como también una gran cantidad de técnicas de procesamiento de imágenes digitales. Esta solución no solo aumentará los tiempos de acceso a los datos, sino que se traduce en una notoria reducción de costos y de tiempo.

Este proyecto, llevado a cabo en PrEBi – UNLP, está centrado en el reconocimiento de la escritura manuscrita, lo cual agrega un nivel de complejidad mucho mayor, debido a las características propias de los textos manuscritos (irregularidades, deformaciones, etcétera), lo que se le suma al deterioro que han sufrido las piezas documentales con las que se trabaja.

El reconocimiento de los textos manuscritos es un proceso que debe realizarse en etapas, que van desde la digitalización de los documentos, pasando por el preprocesamiento y limpieza de las imágenes, segmentación en objetos lógicos, extracción de características, reconocimiento, corrección o refinamiento mediante el uso de diccionarios, hasta llegar finalmente a la indexación de dichos documentos.

El presente trabajo abarca la etapa de extracción de características, en la cual se toman los objetos lógicos obtenidos en la etapa de segmentación y se realizan un gran número de operaciones, tanto morfológicas como espaciales, obteniendo así una gran cantidad de información que permitirá a la etapa posterior – etapa de reconocimiento – hacer coincidir los objetos lógicos con los objetos reales. Todas las características se almacenan en un vector, el cual relaciona directamente a cada una con un valor (peso), que denota su importancia, y que indica al programa “cuanta importancia debe darle a la misma”. Dentro del vector se encuentran características “simples” (alto y ancho, cantidad de agujeros, posición respecto de la línea base), y características complejas, (código de cadenas, características basadas en técnicas con ventana de muestreo)

Se incluye póster en Anexo 4.

13.6 SeDiCI Servicio de Difusión de la Creación Intelectual. UNLP – PrEBi. Prof. Ing. Marisa R. De Giusti, María Marta Vila, Ariel Sobrado, Gisele Jaquenod y Esteban C. Fernández.

Lugar de Presentación: Hotel Howard&Johnson. Ciudad de La Plata, Provincia de Buenos Aires. Argentina.

Evento: Congreso TE&ET.

Fecha Presentación: 2006-08-08.

Resumen.

¿Qué es SeDiCI? El Servicio de Difusión de la Creación Intelectual (SeDiCI) de la Universidad Nacional de La Plata (UNLP) nació con el objetivo prioritario de socializar el conocimiento generado en las diferentes áreas académicas de nuestra Universidad, con el fin de devolver a la comunidad los esfuerzos destinados a la Universidad Pública.

SeDiCI inicialmente fue pensado como un reservorio exclusivo de Tesis y Disertaciones, pero casi de manera inmediata, y debido a las entrevistas con los futuros aportantes de las distintas unidades académicas de la UNLP, surgió la necesidad de ampliar el proyecto para incluir otros tipos de formatos de documentos y no tan sólo documentos textuales, por ejemplo: documentos musicales, cuadros, animaciones, etc.

SeDiCI realiza su trabajo en conjunto con la Dirección de Propiedad Intelectual de la UNLP en el caso que el o los autores deseen registrar su obra. La Dirección y SeDiCI se encargan de las tramitaciones pertinentes de modo de simplificar la tarea de quienes desean difundir su trabajo.

Luego de un análisis pormenorizado de las diversas plataformas que funcionaban en el mundo, el grupo de trabajo de SeDiCI decidió diseñar una plataforma de software

propia, la cual lleva el nombre de Celsius-DL y sirve de sustento al portal SeDiCI: <http://sedici.unlp.edu.ar>.

Celsius-DL soporta un esquema libre y configurable de tipos de material, dentro de los cuales se puede destacar: artículos de publicaciones periódicas, tesinas de grado, tesis de postgrado, producciones multimediales (inclusive al admitir documentos compuestos, una de las últimas incorporaciones consiste en un Libro de enseñanza de Música con ejecuciones sonoras incluidas), libros electrónicos, etc. Cada tipo de material tiene su propio conjunto de “metadatos”, que son datos altamente estructurados que describen la información, describen el contenido, la calidad, la condición y otras características de los datos y por lo tanto resultan aptos para la diversidad de información disponible y a .,disponer en el Portal de SeDiCI.

La arquitectura actual posibilita realizar la búsqueda sobre veinticinco (47) campos distintos pero es posible agregar un número mayor en el caso que procesos técnicos lo considere apropiado. Esta gran cantidad de metadatos permite una búsqueda con mayor pertinencia.

Todas las opciones de exploración del portal están visibles a través del “Recorrido Virtual” que presenta el sitio.

El portal SeDiCI permite la descarga completa o parcial de los documentos, atendiendo posibles problemas vinculados a la velocidad de enlace de los usuarios. Del mismo modo lleva registro de la cantidad de descargas de documento o parte de un documento; esta facilidad ha permitido corroborar, al tener las partes diferente cantidad de accesos, que la herramienta previamente descrita ha sido útil para los usuarios.

Desde el año 2004, el Servicio de Difusión de la Creación Intelectual (SeDiCI-UNLP), brinda la posibilidad de mostrar parte de sus acervos electrónicos a través de “webservices” basados en la tecnología SOAP.

SeDiCI es OAI Data y Service provider cumpliendo de este modo con un compromiso doble, brindar información constantemente actualizada a sus usuarios y difundir la propia en otros lugares del mundo.

Consulte demo en: <http://teyet.info.unlp.edu.ar/teyet2006> y anexo 13

13.7 Celsius Network – Celsius Imago – Celsius Bloodhound. **Marisa R. De Giusti**

Lugar: San Luis, California

Evento: Chameleon Federation Cal Poly Meeting

Fecha: 23-04-2007

Celsius Network

The Library Linkage Project (Proyecto de Enlace de Bibliotecas – PrEBi) of La Plata National University (UNLP) is part of Library Linkage initiative (LL) of Ibero American Science and Technology Education Consortium (ISTEC) which the UNLP belongs to. The main objective of PrEBi is to provide technical and scientific information required for Researchers, Teachers and Students from UNLP, as well as to provide to member institutions that are part of the project of documentations available in libraries in the dependencies of each University. In order to manage all bibliographic resources required by PrEBi users (journal articles, book chapters, congress proceedings, theses and patents) and to manage the requests of Universities participating in the project, it has been developed in UNLP the Celsius Software, which has evolved to Celsius Network platform. This platform permits interconnection among all instances of Celsius, installed in servers of each of more than 30 Institutions that participate in the project distributed in 9 countries of America and Spain. Through this interconnection of instances, many tasks are automated such as remote request creation, request administration and file sending between Institutions. In addition to this, counting with a network of Celsius instances, Institutions are allowed and encouraged to share historical information about searches between each other, constantly improving the service and minimizing response times. This network potentially permits to offer a great variety of services and facilities.

Celsius Network interacts with a unique module named Celsius Directory, also developed in PrEBi UNLP, in which it is centralized information referring to Celsius NT

instances network and to the hierarchical structure of entities participating in the project; in turn this directory regulates interconnection between Celsius Instances and generates global statistics centralizing distributed calculations on each instance.

Celsius Imago

For the digitalization of handwritten text

Research subjects related to digitalization tasks start from different needs: - In relation to documentation material sheltered in libraries and archives, the need of preserving and spreading information forces to work on the areas of digital topology and morphology, in order to optimize some image processing operations that will improve the digitalization result.

- The existence of documents already digitalized, but in really poor conditions: gray images, diffuse borders, spots; represents a similar challenge.

- The existence, also, of a large amount of handwritten documents prone to be subject of complex recognition processes in order to be stored in places with easy access from multiple users.

Celsius Imago will be a scalable software that has started improving problems from already existing documents: spots, borders, centralization, specially related with digital topology and it is expected to be added new functions linked with the research subjects above mentioned.

Huge Information Volumes: indexing, cataloging and storage The open Internet resources, the harvesting processes using OAI protocol that allows to obtain a big amount of information allocated in OAI remote servers (for example DSPACE or SeDiCI reservoirs) don't present special difficulties in some cases the efficient storage might be and would be necessary to analyze different compression methods, but specially the efficient management of this set of records in order to be able to take as much benefits as possible, organizing, making complex searches and minimize processing and access time is a task that presents new challenges.

Celsius Bloodhound

Heterogenous background search engine for Bibliographic Resources

The Celsius BloodHound search engine is an application that will allow to search bibliographic references around the world. This automatic searches will allow to speed up the tasks of bibliographic provision of the Library Linkage Project (PrEBi), integrated to the Library Linkage Net of the Iberoamerican Science and Technology Education Consortium; serving not only the users from UNLP, but also professors, researchers and students from over 60 universities that are part of the Net.

This application to be developed will contemplate the different technological variants in a diversity of available catalogs to allow to maximize the search universe and therefore the possible users that will be able to benefit from the software.

Some Celsius Bloodhound goals are:

1. It should be able to connect with different servers, catalogs, libraries, etc., (from now on just catalogs) taking into account that they can use Standard technologies (Z39.50, SRU, SRW) or their own technologies (web services, cgi, HTML only).

2. Search in the catalogs, recover the data they send and show them.

3. Allow consults from different origins: a web interface, an application, an existing software (like Celsius), some other server or catalog, etc.

4. Search in background.

5. Allow the existence of a dynamic list of catalogs: delete catalogs (offline servers) or add new catalogs with the URL or IP address in order to have access to them, to the technologies in use and the required information for each technology in special.

6. Scalable development: a starting version and new modules: new technologies of new catalogs, new searching criteria.

SeDiCI Intellectual Creation Diffusion Service. Chameleon Federation Cal Poly Meeting. Resumen

Objectives.

The Intellectual Creation Diffusion Service (SeDiCi) of the UNLP was created with the main objective of communicating to the society all the knowledge acquired in different academic fields of our university. This is how we can return all the efforts that were destined to our public institution.

SeDiCi was initially thought as an exclusive reservoir of Thesis and Dissertations, but immediately and after interviewing several of the participants from the different academic units, we had to enlarge the SEproject to include not only text formats but also other types such as musical documents, charts, animations, etc.

Se incluye póster en Anexo 13.

13.8 Celsius Bloodhound. **Prof. Ing. Marisa R. De Giusti**, Gonzalo L. Villarreal. International Conference on Engineering Education – ICEE 2007. September 3-7 2007. Coimbra, Portugal.

Abstract.

The Library Linkage Project (Proyecto de Enlace de Bibliotecas – PrEBi) of La Plata National University (UNLP) is part of LL initiative of Ibero American Science and Technology Education Consortium (ISTEC). The main objective of LL is to share the bibliographic acervus resident in ISTEC Institutions, in order to satisfy the demand of Researchers, Teachers and Students. In this model, PrEBi serves both local and international requests.

To manage the requirements of bibliography (journal articles, book chapters, congress proceedings, theses and patents) made by PrEBi users and by foreign Universities, it has been developed in UNLP the Celsius Software. Nowadays the main job of Celsius Administrators is to look up users requests in every catalog of every library of ISTEC Institutions and many non-ISTEC member Institutions catalog. This task is made by hand and it takes many hours to visit and check every catalog until the resource is found, if so.

In order to perform this task more efficiently, it has been created in PrEBi the Celsius Bloodhound Project. This project, currently in analysis and design stage, intends to automatize this tedious job by searching and fetching records from libraries servers that expose some sort of interface for connecting, searching and retrieving information from it.

The modular design of Bloodhound intends to be flexible enough to allow connections from almost any application, using web services technology to expose the main functions of the application with a clear and simple interface, and to connect to many catalogues with different technologies to expose metadata, such as Z39.50, SRU/SRW, and so on. This way Bloodhound will be able to interact transparently with many servers, taking advantage of existing technologies and protocols, in order to improve quality service in time response in the search of resources in PrEBi and other Institutions as well.

Se incluye el trabajo completo en Anexo 18.

13.9 Celsius Network. **Prof. Ing. Marisa R. De Giusti**, Ariel J. Lira, Fernando G. Inafuku, Ariel Sobrado. International Conference on Engineering Education – ICEE 2007. September 3-7 2007. Coimbra, Portugal.

Abstract

The cooperative and collaborative work between Universities and Libraries are key points for teachers, researchers and students to make progress in their studies and research. The Library Linkage Initiative has pursued this purpose since its creation, and it has been very useful for many Institutions, Universities and Libraries in America and Spain, specially for those with scant resources. In PrEBi[2] UNLP[3] it has been created the Celsius software, which has become the standard application for management of users requests in the Initiative. The result of many years of use, improvement and experience acquisition has result in Celsius Network which has made a breakthrough in the Celsius development history. This last version intends to connect all instances of Celsius in a network, meeting the needs of users and operators that belong to the

participants Institutions members of ISTE[1]. In the following article we have included a brief description of what ISTE, LL Initiative and Celsius are - in order to present global view where Celsius Software fits - and a detailed description of users needs, new features and advantages in Celsius Network as well.

Se incluye el trabajo completo en Anexo 18.

13.10 INTERCONNECTION AND SHARING THROUGH WEBSERVICES. **De Giusti, Marisa R.**; Villarreal, Gonzalo L.; Sobrado Ariel; Lira, Ariel J.; Vila Maria M. Cameleon Federation Bologna.

Introduction.

Current communication technologies permit to implement new cooperative work models in which most complex task may be achieved by dividing work among several participants, where communication gets much easier, process are automatized the huge amounts of information are freely shared. These work models offer important advantages for an increasing number of users and Institutions, standing out a higher efficiency, lower costs and higher quality services.

Webservices technology is one of the stilts on which the concept WEB 2.0 is sustained. This technology allows to send information between computers connected to the network, providing a high portability level communication mechanism that permits this interconnection to be made between different hardware platforms and internal data representation, many operative systems, applications developed in several languages or databases and mainly, a very simple language, standardized and efficient for information transport such as XML.

The Library Linkage Project (PrEBi – Proyecto de Enlace de Bibliotecas; www.prebi.unlp.edu.ar) of La Plata National University (UNLP) develops and supports software of internal use and systems widely used in Institutions and Universities of America and Spain. These developments take part in the different projects that PrEBi carries out or is involved in, which are:

- **PrEBi:** LibLink is an initiative generated in the Iberoamerican Science and Technology Education Consortium (ISTEC; www.istec.org) in which a huge number of libraries from ISTE member Universities share their bibliographic acervus, which permits to researchers, students and teachers that belong to involved Institions to get bibliographic resources not available in its local libraries, in a relatively short time and with very low costs. PrEBi follows this project model and in an active member since 1996 (LibLink initiative operates since 1994). PrEBi is the center of Celsius software development, which is used for many partitipants for user requests managment and automated statistics obtaining.
- **SeDiCI:** The Intelctual Creation Difussion Service (Servicio de Difusion de la creacion intelectual; <http://sedici.unlp.edu.ar>) in a project generated inside UNLP and carried out by PrEBi through which teachers, stutents and researchers of this University can publish their grade work, theses and research results, and spread them among the whole comunity. On one hand, the project pretends to return to the comunity part of the efforts invested in the University; on the other hand the aim is to set out a centralized site where anyone can access freely to all the intelectual production generated in UNLP, where creators may contact with researchers in related areas in all the world and where contents are openly spread using very diverse ways.

As can be seen based on the brieft descriptions, project carried out by PrEBi are strongly related with cooperative work, document sharing and transmtion and contents diffusion. With the purpose of improving the quality of the services offered, all systems developed by PrEBi staff are constantly improved adding to them new functions and updated technolgy, distinguishing webservices addition in both projects for very different tasks and purposes, which shows the flexibility that this technology contributes with to web systems. In this article it will be explained how webservices technology has been

adopted for PrEBi developments, and how each project is benefit with the use of these tools.

Se incluye trabajo completo en Anexo 13.

14. CURSOS DE PERFECCIONAMIENTO, VIAJES DE ESTUDIO, ETC. *Señalar características del curso o motivo del viaje, período, instituciones visitadas, etc*

14.1 Nombre: "HP DPUC Seminar". 23 y 24 de Abril de 2007. Duración: 16 horas. Institución: Cal Poly State University. San Luis Obispo, California, Estados Unidos. Se incluye certificado en Anexo 15.

15. SUBSIDIOS RECIBIDOS EN EL PERIODO. *Indicar institución otorgante, fines de los mismos y montos recibidos.*

15.1 Institución otorgante: Comisión de investigaciones científicas. Período 2007-2008. Monto: \$ 3.000.

16. OTRAS FUENTES DE FINANCIAMIENTO. *Describir la naturaleza de los contratos con empresas y/o organismos públicos.*

16.1 En el año 2007, ISTECS otorgó US\$ 4000 para culminar el desarrollo de la versión Celsius Network del Software Celsius, junto con un software instalador y con la actualización y asistencia a las Universidades de América que lo están instalando, en el marco del Proyecto de Enlace de Bibliotecas (PrEBi).

16.2 El Proyecto del Proyecto de Enlace de Bibliotecas (PrEBi) cuenta con un presupuesto anual de 50.000\$ otorgado por la UNLP y aprobado en su presupuesto.

17. DISTINCIONES O PREMIOS OBTENIDOS EN EL PERIODO.

17.1 Concurso Internacional: a través de un Concurso Internacional con las características habituales de oposición y antecedentes fui elegida Directora Asociada de la Iniciativa Library Linkage (LibLink) del Consorcio Iberoamericano para Educación en Ciencia y Tecnología (ISTEC). En el Anexo 10 se incluyen: las notas de presentación de la UNLP, el programa presentado para la Iniciativa en las tres lenguas y la comunicación oficial de la decisión por parte del Consorcio ISTECS.

18. ACTUACION EN ORGANISMOS DE PLANEAMIENTO, PROMOCION O EJECUCION CIENTIFICA Y TECNOLÓGICA. *Indicar las principales gestiones realizadas durante el período y porcentaje aproximado de su tiempo que ha utilizado.*

18.1 En la Universidad Nacional de La Plata. Responsable como Directora de PrEBi de la participación en los Talleres realizados por la UNLP durante 2007 (3 talleres comenzando el 23/8/2007) y Marzo de 2008 dedicados a la Planificación Estratégica de la Universidad para el período 2008-2010.

18.2 Dentro del Consorcio Iberoamericano para Educación en Ciencia y Tecnología (ISTEC): En este contexto he cumplido desde 1990 con las siguientes funciones:

18.2.1. Miembro de la Junta Ejecutiva desde 1991 hasta Julio de 2007.

18.2.2. Representante de la Universidad Nacional de La Plata en todas las Asambleas Generales (anuales) de ISTECS desde 1990 a la fecha.

18.2.3. Representante de la Universidad Nacional de Noroeste de la Provincia de Buenos Aires desde 2005 y hasta la fecha.

Nota: El ISTECS es una organización dedicada a apoyar proyectos de docencia e investigación ligados a Ciencia y tecnología, educación presencial y a distancia en grado y posgrado, actualización curricular y actividades de transferencia de docentes y alumnos entre los distintos países de América y España que lo integran.

19. TAREAS DOCENTES DESARROLLADAS EN EL PERIODO. *Indicar el porcentaje aproximado de su tiempo que le han demandado.*

19.1 Tareas de grado:

19.1.1. Cargo actual: Profesor Adjunto Interino Dedicación Simple, por concurso de la Facultad de Informática, UNLP, para la materia: "Modelos y Simulación". Resolución Nro. 165/01 de Junio de 2001. (Corresponde el mismo cargo previo pero con rotación de materia)

19.2 Tareas de postgrado:

19.2.1. Profesor del Curso de Doctorado en Ciencias Informáticas: "Gestión de la calidad según normas ISO. Aplicaciones a Software". Duración 20 horas. Facultad de Informática, Universidad Nacional de La Plata. Agosto de 2006 y Junio de 2007.

19.2.2. Profesor invitado para el dictado del curso Taller "Celsius 1.6 y Celsius Network", el día 6 de Junio de 2007. Duración 8 horas. Evento organizado por Consorcio Iberoamericano para la Educación en Ciencia y Tecnología en la Universidad Técnica Particular de Loja, Ecuador.

20. OTROS ELEMENTOS DE JUICIO NO CONTEMPLADOS EN LOS TITULOS ANTERIORES. *Bajo este punto se indicará todo lo que se considere de interés para la evaluación de la tarea cumplida en el período.*

- 21. TITULO Y PLAN DE TRABAJO A REALIZAR EN EL PROXIMO PERIODO.** *Desarrollar en no más de 3 páginas. Si corresponde, explicita la importancia de sus trabajos con relación a los intereses de la Provincia.*

TITULO Y PLAN DE TRABAJO 2008-2009

Investigación sobre ontologías y web semántica. Desarrollos de software para la biblioteca digital y modelización de sistemas.

1) Investigación sobre ingeniería ontológica y web semántica.

La Web semántica es una extensión de la web actual en la cual el “significado” de la información está bien definido y permite un mejor trabajo cooperativo entre máquinas y personas. El esqueleto de la Web semántica se compone de metadatos asociados a los documentos y archivos de la Web actual. Las ontologías proveen un dominio de conocimiento compartido que es capaz de permitir la comunicación entre agentes (agentes, agentes inteligentes, agentes semánticos) y humanos. Las ontologías son claves para la web semántica y pueden representarse a través del uso de lenguaje específico tales como OWL (lite, DL, Full). Dado que las bibliotecas digitales ofrecen el acceso a grandes volúmenes de contenidos se pretende estudiar algunos temas claves relacionados con la web semántica para permitir una mejor descripción de los objetos digitales y de los repositorios. El objetivo prioritario es permitir la interoperabilidad expresada fundamentalmente por la facilidad de acceso, la consistencia y la coherencia de los objetos digitales y de los servicios que ofrece el repositorio SeDiCI. El estudio de las denominadas “tecnologías semánticas” permitirá incluir a posteriori nuevas herramientas en el repositorio que ayuden a mostrar de manera confortable la información, permitiendo una navegación más cómoda y atendiendo al perfil de los usuarios. Otro desafío a atender en esta etapa de investigación es dar al usuario la posibilidad de visualizar los diferentes repositorios de una manera consistente, como si se tratara de un repositorio simple.

Se estudiarán las tecnologías más prominentes de la web semántica: editores ontológicos, herramientas de anotación, máquinas de inferencia. El enfoque propuesto para esta investigación será consistente con el adoptado por los grupos de trabajo de W3C (<http://www.w3c.org/>) es decir un enfoque en capas o layers como el propuesto por Berners-Lee.

2) Desarrollos de software para la Biblioteca Digital

2.1.) Gestión inteligente de la información: El análisis realizado el período 2006-2007, sobre diversos motores de bases de datos no permite aun arribar a una solución definitiva al problema de la gestión de grandes volúmenes de información. Sin embargo, a partir del mismo se ha logrado definir una línea a seguir en las próximas pruebas que deberán ser efectuadas, que consiste principalmente en asignar un papel más protagónico a los indexadores de texto.

La problemática de la gestión y recuperación de grandes volúmenes de información subyace en el hecho que las operaciones a realizar son extremadamente diferentes, y por ello sucede que ciertas bases de datos son performantes para la gestión (inserción, modificación, relaciones) pero no para las búsquedas y recuperación, mientras que con otras sucede la situación opuesta. Las bases de datos analizadas, a excepción de Apache Lucene, pueden ser colocadas en la primer categoría, mientras que los indexadores de texto en general pueden ser colocadas en la segunda.

Una alternativa a evaluar consiste en mantener los datos en dos bases de datos separadas, y mantener relaciones entre ellos de forma que la gestión pueda ser realizada desde una base de datos y la indexación y búsqueda pueda ser realizada desde un indexador de texto, es decir, utilizar un motor de base de datos que se ajuste a la necesidad eficientemente.

Otra cuestión que surge del almacenamiento de grandes volúmenes de datos es la del ordenamientos de los resultados de una búsqueda. Si bien este problema puede parecer menor, no lo es. Cuando el número de resultados es elevado, el ordenamiento

de los mismos determina la usabilidad de la herramienta de búsqueda (en este caso SeDiCI) debido a que mayoritariamente los usuarios no examinan todos los resultados arrojados, sino solo los primeros 10 o 20. La desventaja que poseen la mayoría de los motores de base de datos de la actualidad es que carecen de criterios de ordenamientos complejos en forma nativa; la funcionalidad que proveen en general se limita a ordenar los resultados por un campo específico. Para lograr ordenar los resultados en base a criterios específicos se deben definir configuraciones complejas que indiquen la relevancia o peso de cada campo, las combinaciones de campos aceptados, y muchas más restricciones semánticas que implícitamente existen dentro de cada registro de información.

Los próximos análisis serán realizados inicialmente sobre Apache Lucene junto a Apache Solr, y consistirán en lograr explotar todas las facilidades que la plataforma provee. A continuación, la nueva versión de Berkeley dbxml será reevaluada, ya que el error existente en la versión anterior ha sido solucionado según indica la empresa responsable de su desarrollo.

2.2) Reingeniería de SeDiCI: Con el objetivo de comenzar un proceso de renovación del Servicio de Difusión de la Creación Intelectual de la UNLP, se deben realizar varias extensiones, entre las que se pueden citar: administración distribuida de documentos, políticas de seguridad, capacidad de controlar mayor volumen de materiales externos, modificación de la interfaz, internacionalización del sitio, interoperabilidad por web services, y búsquedas semánticas entre otras.

La administración distribuida permitirá controlar la creación, modificación y catalogación de documentos a través de Internet sin necesidad de una configuración específica desde dentro del PrEBi, permitiendo de esta forma descentralizar la carga de trabajo a realizar dentro del repositorio.

Las extensiones a las políticas de seguridad son necesarias debido principalmente a la distribución de la administración, ya que al dar acceso público vía Internet los riesgos de accesos indebidos se acrecentarán.

Se debe incrementar la capacidad de control de material recolectado bajo el protocolo OAI-PMH desde repositorios externos, con el fin de poder abarcar más repositorios de diferentes temáticas específicas y realizar cosechas actualizadas y periódicas.

La visualización del portal deberá ser actualizada para poder brindar un servicio más usable e intuitivo para todo tipo de personas, independientemente de su capacitación técnica.

La internacionalización del sitio permitirá definir un número arbitrario de idiomas en los que se podrá visualizar el portal. En cualquier momento se podrán agregar nuevos idiomas soportados simplemente cargando las traducciones correspondientes.

SeDiCI posee varios web services que permiten el acceso desde motores de búsqueda externos, sin embargo, debe rediseñarse la arquitectura de web services para poder contemplar nuevos tipos de servicios automáticos, como por ejemplo, servicios de búsqueda de documentos para bibliotecas externas que tienen sus documentos alojados en SeDiCI, o protocolos de Intercambio de material/metadatos del tipo OAI, SRU/SRUW , z3950, etc.

Las búsquedas semánticas permitirán devolver a los usuarios un conjunto de resultados recuperados criteriosamente y ordenados por relevancia. Esta recuperación se realiza en base a criterios variables que frecuentemente dependen del perfil de usuario, de metainformación de los documentos almacenada en el repositorio y/o de normalizaciones y correcciones a los términos de búsqueda cargados por los usuarios. En cambio, el ordenamiento de los resultados conseguidos, depende principalmente de la relevancia de cada ítem particular dentro del conjunto de todos los resultados, y asignándose la relevancia en base a los pesos de cada tipo de coincidencia de la expresión de búsqueda dentro de los ítems.

Implementar todas las extensiones arriba mencionadas sobre la plataforma actual de SeDiCI es una tarea prácticamente imposible debido a diversos motivos, dentro de los

cuales se puede nombrar la ausencia de una arquitectura informática extensible que permita agregar estas nuevas funciones, las tecnologías usadas originalmente casi han caducado, y finalmente el hecho de que se debería reimplementar alrededor del 60 % de la administración de carga a causa de la descentralización.

En resumen, el elevado costo asociado a modificar la plataforma SeDiCI nos lleva a tener que realizar un proceso de reingeniería sobre la plataforma, ya que el costo de rediseñar e implementar la misma es muy similar al costo modificarla, mientras que el beneficio es mucho menor. Sumado al aumento del costo/beneficio, realizar una reingeniería nos permite utilizar tecnologías más novedosas.

3)Desarrollos en relación a los compromisos internacionales del lugar de trabajo bajo mi dirección: habiendo cumplido la planificación de informe 2006-2007 en lo que respecta al desarrollo de Celsius Network (CelsiusNT) que permite la interconexión de todas las instancias de Celsius en una red virtual entre las partes y con ello la automatización en las tareas de solicitud y entrega de documentos, realizado el primer modelo probabilístico de funcionamiento, para el período 2008-2009 se propone el agregado al software de las llamadas "Estadísticas Globales", entre las cuales podemos mencionar: número total de solicitudes entre sitios, cantidad de páginas, retardos, etc.. La importancia de estas estadísticas estriba en que con ellas será posible obtener datos del funcionamiento del sistema de intercambios en su totalidad.

4)Desarrollos de software en PrEBi: En PrEBi se realiza la búsqueda de material bibliográfico sobre catálogos expuestos en Internet de las Bibliotecas cooperantes (no se trata aquí de bases de datos abiertas). Existen actualmente muchos protocolos que permiten obtener la información expuesta por catálogos: Z39.50, OAI PMH, webservices, etc... Se destaca el protocolo Z39.50, debido a su gran expansión a nivel global que lo ha convertido en un estándar mundial. Este protocolo es muy flexible, independiente de la plataforma y posee soporte en la mayoría de los lenguajes de programación de aplicaciones web actuales (java, php, perl, etc...). Además, existe una enorme cantidad de herramientas, librerías y documentación disponible abiertamente para bibliotecarios y desarrolladores de aplicaciones. Se propone la realización de una plataforma denominada Celsius Bloodhound dedicada a realizar "off line" la búsqueda en los catálogos de las bibliotecas cooperantes de manera automática entregando al administrador la localización del recurso con el fin de agilizar la etapa de búsqueda que demanda un 30% del tiempo total desde que se inicia la solicitud hasta la entrega. En este trabajo, se requiere que la aplicación realizada interactúe al menos utilizando el protocolo Z39.50, pero debe estar pensado para poder incluir eventualmente otras tecnologías (por ejemplo, incorporar webservices). El diseño preliminar de esta herramienta se adjunta en el trabajo presentado en el punto 13.7.

5) Modelización de sistemas:

5.1. Modelización de CelsiusNT: Transcurrido un tiempo de utilización concreta por parte de las instancias de CelsiusNT en funcionamiento, se propone, mediante la utilización de un sistema de simulación de eventos discretos, el desarrollo del nuevo modelo probabilístico para determinar, ya con bases en datos obtenibles de la utilización real, la performance y capacidad de respuesta frente a situaciones anómalas diferentes a las previas y derivadas de las nuevas capacidades del software.

5.2.Modelización de Celsius Bloodhound: se realizará el modelo global del sistema, el cual deberá servir tanto para desarrollar la aplicación como para la simulación. Se considerarán diferentes niveles de carga, caídas remotas, etc... Como no todas las tecnologías son igualmente eficientes, requieren el mismo procesamiento y consumen el mismo ancho de banda estas diferencias se pondrán en evidencia al momento de realizar las simulaciones y serán útiles en las diferentes etapas de avance del desarrollo.