

La interoperabilidad de los sistemas de información académico-científica de la UNLP desde la perspectiva del acceso abierto a la información: un tema de tesina.

María del Rosario Molfino^{1,2}

¹Biblioteca de la Facultad de Informática. ²Departamento de Bibliotecología, Facultad de Humanidades y Ciencias de la Educación, Universidad Nacional de La Plata. 48 e/6 y 7, 1900 La Plata, Argentina.

E-mail: rmolfino@info.unlp.edu.ar rmolfino@fahce.unlp.edu.ar

Resumen. Los elementos claves para el desarrollo de la sociedad de la información son la recuperación, el intercambio y cooperación oportuna, efectiva y automática de datos, información, documentos y objetos digitales entre los sistemas de información. En este trabajo, comentaremos las motivaciones que dan origen al tema de la tesina para obtener el título de Licenciada en Bibliotecología y Ciencias de la Información y la razón por la cual se circunscribió a la Universidad Nacional de La Plata. Además, se comentan los primeros resultados obtenidos en la etapa de relevamiento bibliográfico, los cambios efectuados al título, las palabras claves, el índice y los objetivos de la investigación. A modo de cierre se comentarán brevemente las herramientas de apoyo utilizadas hasta el momento.

Introducción

La gestión de la información en pos del acceso constituye el núcleo central de nuestra actividad profesional. En los últimos tiempos, la tarea de organizar la información de manera que esté disponible para apoyar los procesos de investigación y aprendizaje de nuestros usuarios, ha sufrido profundas transformaciones debidas al uso de la tecnología y el trabajo en red. Emerge así un nuevo concepto de unidad de información de naturaleza compleja, que pone en cuestión la capacidad de los sistemas tradicionales de las bibliotecas para tratar esa complejidad y soportar los procesos de trabajo de sus usuarios en el nuevo entorno digital basado en la Web y las posibilidades de colaboración que esta ofrece.

Durante la mayor parte de su historia, las bibliotecas han vivido en el mundo de lo impreso. Todo nuestro sistema de gestión se basaba en un contexto en el que la información era estática, estaba pensada para ser utilizada por las personas en entornos

físicos y aislados. La primera gran revolución digital fue la automatización del catálogo, que generó la primera información digital que entró en nuestras bibliotecas: los metadatos, es decir, el registro bibliográfico legible por computadora. Teníamos una información que no estaba pensada para ser "leída" únicamente por las personas. A partir de ahí, con la penetración de Internet y sus tecnologías en nuestros procesos, los cambios se sucedieron cada vez más rápido. A los registros bibliográficos legibles por computadora siguieron lo que entonces denominamos "recursos electrónicos", las revistas y libros electrónicos primero, y luego las colecciones digitales que fueron incorporándose como resultados de nuestros procesos de digitalización. Nuestros catálogos empezaron a incluir hiperenlaces a los textos descritos por esos y con ello comenzamos a gestionar no sólo información referencial sino también información de contenido (el texto completo). Empezamos a enriquecer nuestros catálogos con nuevas prestaciones y funcionalidades y nos incorporamos a la Web. El último hito de esta segunda fase ha sido la creación de los repositorios institucionales para asumir la responsabilidad de la gestión de la información digital que se genera en nuestras instituciones. Sin embargo, el desafío continúa, y en el ámbito académico/científico comienza ya a percibirse un interés creciente por la gestión de los datos primarios de investigación. Un nuevo horizonte, donde la heterogeneidad de la información y de los sistemas que la gestionan conforma un escenario cada vez más complejo.

Se parte de considerar que nuestra actividad, además de realizar una gestión eficiente intra-muros, se preocupa por garantizar el más amplio acceso a los recursos externos de utilidad para su comunidad de usuarios. Si bien es cierto que la cantidad de sistemas de información crece -lo que en términos de potencialidad siempre se verá como positivo- es una realidad también que existe una gran disparidad en las formas de tratamiento de la información y de los sistemas informáticos que la gestionan. Existe un problema de heterogeneidad de productos de hardware y software, de modelización de sistemas, de tratamiento semántico de la información, de forma de ingreso de datos. Pese a lo anterior, en la biblioteca universitaria existe una tendencia general hacia la integración de sistemas, de la misma manera en que se tiende a la convergencia de sistemas digitales de toda la Universidad (podemos poner de ejemplo la integración del sistema integrado de gestión bibliotecaria KOHA-UNLP con el sistema de gestión de alumnos SIU-Guaraní). La inversión en sistemas de gestión de información, de gestión

bibliotecaria, colecciones documentales físicas y digitales y en sistemas de gestión de activos digitales diseñados para apoyar y facilitar las funciones de enseñanza e investigación de las universidades suponen costos muy elevados para éstas. Para mejorar el retorno de la inversión realizada, es deseable que todos estos sistemas se integren en un entorno de información digital homogénea y coherente.

Motivación

El principal motivo para la elección del tema a desarrollar en mi tesina de licenciatura es la percepción de la situación problemática real, objetiva y observable, antes descripta. Parto de considerar que los bibliotecarios podemos aportar conocimientos de alta significación, metodologías de trabajo, aptitudes, actitudes y visiones de mucha utilidad al momento de pensar la integración de sistemas. La conformación de equipos de trabajo compartidos con los informáticos ya constituye una realidad en nuestro entorno, pero la permanencia y consolidación de estos equipos depende en gran medida de que los bibliotecarios nos adaptemos constantemente a los nuevos entornos tecnológicos. Por todo ello, se cree encontrar en el área de la Interoperabilidad de sistemas de información, un tema de investigación y trabajo que permite desplegar y hacer avanzar nuestro perfil profesional.

Dado que los sistemas de información funcionan en contextos determinados, con comunidades de usuarios específicas que los mantienen, mejoran, alimentan y usan se hace imprescindible circunscribir cualquier estudio que los involucre a un ámbito específico. En tal caso, el desafío siempre será trabajar la particularidad sin perder de vista las normas y estándares internacionales y las tendencias tecnológicas globales. Para este trabajo se ha elegido el caso de la UNLP, fundamentalmente porque no siendo ajeno a la situación problemática general, permite delimitar un universo tratable, al menos en un nivel de profundidad intermedia. La factibilidad de acceso a la estructura de la organización para la realización del estudio también fue motivo de evaluación al momento de la selección del caso.

Solo para mencionar de manera general algunos de los principales problemas que se observan, antes de comenzar en el estudio sistemático, señalamos que:

1. No existe un repositorio institucional digital de la UNLP completo que albergue y difunda de manera sistemática toda la producción académico-científica producida por los miembros de la comunidad universitaria a través de sus diversas dependencias (Facultades, Centros y Laboratorios de investigación, entre otros).
2. No es posible participar de repositorios temáticos nacionales, regionales o internacionales que vuelvan visible las investigaciones y la producción intelectual de la UNLP en su conjunto.
3. Los OPAC y los repositorios de la UNLP no interoperan entre ellos.
4. Existe poca información global sobre la cantidad, tipología, antigüedad, volumen de datos, software y protocolos que soportan los sistemas de información que existen actualmente en la UNLP. Interesan especialmente los sistemas que pueden ofrecer información abierta, y los de datos primarios de investigación.

Exploración bibliográfica preliminar

Las primeras estrategias de búsqueda fueron interoperabilidad, protocolos de intercambio de información, OAI-PMH. Algunas de las fuentes consultadas fueron:

E-LIS

Library Literature & Information Science Full

LISTA Library, Information Science & Technology Abstracts (EBSCO)

Redalyc

Scopus

De las bases de datos referenciales, se seleccionaron los documentos más relevantes que luego se buscaron a texto completo. Estos documentos fueron subidos al Google Sites diseñado para la tesina. Más adelante se comentará el valor de esta herramienta.

Cambios en la propuesta del plan de trabajo

Durante la etapa que actualmente se está transitando -que corresponde a la elaboración del plan de trabajo- las ideas han ido cambiando. En el siguiente tabla se resumen las modificaciones que han sufrido el título y las palabras claves.

Tabla 1: Cambios en la propuesta del plan de trabajo de la tesina

	1ª etapa	2ª etapa	3ª etapa
TITULO	Open Access y Open Archive: el acceso abierto a la información científica.	Estudio de interoperabilidad en los sistemas de información documental. El caso de la UNLP	La interoperabilidad de los sistemas de información académico – científica de la UNLP desde la perspectiva del acceso abierto a la información
PALABRAS CLAVES	ACCESO ABIERTO; INFORMACIÓN CIENTÍFICA; MOVIMIENTO OPEN ACCESS; PROTOCOLO DE INTERCAMBIO DE INFORMACIÓN; OAI-PMH	INTEROPERABILIDAD; SISTEMAS DE INFORMACIÓN DOCUMENTAL; SISTEMA DE INFORMACIÓN CIENTÍFICA; PROTOCOLOS DE INTERCAMBIO DE INFORMACIÓN; METADATOS	INTEROPERABILIDAD; SISTEMAS DE INFORMACION DOCUMENTAL; AMBITO ACADEMICO-CIENTIFICO; PROTOCOLOS DE INTERCAMBIO DE INFORMACIÓN; METADATOS; MODELO DE REFERENCIA OAIS; INTERCAMBIO DE DATOS CIENTIFICOS; ACCESO ABIERTO A LA INFORMACIÓN; UNLP

Avances en la definición de la metodología

Este es un trabajo que se encuentra enmarcado claramente dentro del campo de las Tecnologías de Información (TI), por lo tanto es conveniente realizar algunas precisiones sobre el marco metodológico en el que nos movemos.

Para las TI una definición posible de información es “datos que han sido procesados de manera significativa para el receptor y tienen un valor real o percibido para la acción y la toma de decisiones actuales o futuras”. Por su parte, la tecnología ha sido definida como “aplicación práctica de la inteligencia”, es decir que se concibe como algo práctico y útil, que se encarna en instrumentos o artefactos, en lugar de ser exclusivamente conceptual. Por lo tanto, la Tecnología de la Información es la tecnología utilizada para adquirir y procesar la información en apoyo de propósitos humanos.

Para los autores March y Smith [1995], la atención científica sobre esta área de estudio se debió en un primer momento a su potencial para impactar –positiva o negativamente-

en la eficacia de la organización. El interés científico se basa en la hipótesis de que estos fenómenos pueden ser explicados por las teorías científicas y que la investigación científica puede mejorar la práctica.

Sin embargo para ellos hay dos tipos de interés científico en esto, el descriptivo y el prescriptivo. En sintonía con las ideas de Herbert Simon [1996], si la investigación en TI estudia lo artificial en lugar de los fenómenos naturales, es decir las creaciones humanas tales como las organizaciones y los sistemas de información, el hecho de que estos fenómenos artificiales pueden ser a la vez creados y estudiados, hace que los científicos del campo puedan contribuir con cada una de estas actividades. Esta es básicamente la doble naturaleza de la investigación en TI. Los autores consideran que los productos de las ciencias del diseño son de cuatro tipos: construcciones, modelos, métodos y puestas en práctica. Las construcciones son las conceptualizaciones usadas para describir los problemas dentro del dominio y especificar sus soluciones. Se pueden combinar en construcciones de orden superior, a menudo denominados modelos, que se utiliza para describir las tareas, situaciones o artefactos. Los científicos del diseño también desarrollan métodos, formas de llevar a cabo actividades dirigidas a un fin. Finalmente, estos se pueden concretar en productos específicos, implementaciones físicas destinadas a realizar ciertas tareas. En lugar de plantear teorías, los científicos del diseño se esfuerzan por crear modelos, métodos y puestas en práctica que sean innovadores y valiosos. Las ciencias del diseño involucran dos actividades básicas, construir y evaluar. Esto sería el paralelismo del par descubrimiento/justificación en las ciencias de la naturaleza.

Al tomar estas ideas como punto de partida, la tesina se apoyará en la construcción conceptual existente a nivel de Sistemas de Información en general y sobre la Interoperabilidad de esos sistemas en particular; así como también sobre todas las conceptualizaciones existentes a nivel de Bibliotecas Digitales y Sistemas Automáticos de Gestión Bibliotecaria. Haciendo uso de ese vocabulario, se aspira a construir un modelo que represente la situación problema y su solución. La información sobre la situación actual deberá ser relevada. Para ello, se propone la realización de un inventario de los sistemas de información existentes en la UNLP que de cuenta de sus correspondientes capacidades de interoperabilidad. Dado que para el contexto de este trabajo interesa el tipo de información que es factible de ofrecer de manera abierta,

quedan excluidos todos los sistemas que dentro de la universidad gestionan información personal. Para la realización del inventario se propone comenzar con un cuestionario específico que será acompañado de una entrevista abierta a informantes clave dentro de la organización. El primer grupo seleccionado corresponde a los directores de bibliotecas, quienes permitirán, siguiendo una estrategia de “bola de nieve”, detectar otros informantes dentro de la estructura de su propia Unidad Académica. La Directora de la Biblioteca Pública de la UNLP será la principal referente para aportar los informantes a nivel de la estructura general de la universidad.

El objetivo principal de la tesina lo constituye la elaboración del modelo, por lo cual, el trabajo sobre el método, entendido como el conjunto de pasos -algoritmo o guía- usado para realizar la tarea de interoperabilidad propuesta, se abordará en relación a un subcaso concreto a seleccionar, el cuál podrá concluir en una implementación, dependiendo de la posibilidad concreta que se presente de integrar un grupo de trabajo de desarrolladores.

Herramientas de apoyo

Durante la etapa de la elaboración del plan de la tesina se han utilizado aplicaciones que facilitan la organización de las actividades típicas involucradas en la realización de una tesina. Por ejemplo:

- **Google Sites** que es una aplicación online que permite crear un sitio web de una forma tan sencilla como editar un documento. Cualquier usuario con unos pocos clicks puede construir una web a su medida reuniendo en un único lugar y de una forma rápida información variada, vídeos, calendarios, presentaciones, fotos, archivos adjuntos y texto. Se pueden crear sitios colaborativos y hacerlos públicos o privados. Aquí se lo ha usado fundamentalmente como una herramienta de comunicación con el director y co-director. La posibilidad de concentrar todos los materiales y comentarios en un solo espacio, siguiendo una lógica de organización compartida por los tres, no solo ahorra tiempo, sino que también es un gran ordenador de lo que se va haciendo. Por ahora es un sitio privado ya que hay colgada allí mucha bibliografía que no es de acceso abierto, pero la propuesta es que una vez terminado el trabajo se reestructure y quede a disposición de todos.



Figura 1: Página principal del sitio creado en Google Sites

- **ZOTERO** que es un complemento de Firefox 2.0 pensado para estudiantes e investigadores y funciona como un gestor bibliográfico. Ha sido escogido por PC Software como una de las mejores aplicaciones del software libre. Es gratuito, de código abierto, se implementa continuamente, tiene varios idiomas y permite elaborar bibliografías en los principales estilos.

Zotero recopila información de una amplia variedad de documentos, la almacena y nos permite etiquetar, modificar, añadir y quitar datos. Además nos permite relacionar unos documentos con otros, a través de sus datos, hacer colecciones y subcolecciones. Importa datos directamente desde:

- las páginas Web visualizadas en el momento
- catálogos Web: Google Scholar, Library Of Congress, Scirus, y multitud de bibliotecas y redes científicas
- catálogos comerciales como Amazon

Exporta datos a diversos formatos (RIS, TeX, MODS y otros) y genera citas bibliográficas con estilos APA, MLA y Chicago.

- **L^AT_EX**: Es un sistema de composición de textos, orientado especialmente a la creación de libros, documentos científicos y técnicos que contengan fórmulas matemáticas.

La idea principal de LaTeX es que el autor se centra en el contenido y no en la forma del documento. Para lograr esto, LaTeX está provisto de una serie de macros y estilos predefinidos. El propósito de LaTeX es facilitar el uso del lenguaje de composición tipográfica.

Con LaTeX, la elaboración del documento requiere normalmente de dos etapas: en la primera, hay que crear mediante cualquier editor de texto llano un fichero fuente que, con las órdenes y comandos adecuados, contenga el texto que queramos imprimir. La segunda consiste en procesar este fichero; el procesador de textos interpreta las órdenes escritas en él y compila el documento, dejándolo preparado para que pueda ser enviado a la salida correspondiente, ya sea la pantalla o la impresora. Ahora bien, si se quiere añadir o cambiar algo en el documento, se deberá hacer los cambios en el fichero fuente y procesarlo de nuevo. Esta idea, que puede parecer poco práctica a priori, es conocida por los que están familiarizados con el proceso de compilación que se realiza con los lenguajes de programación de alto nivel (C, C++, etc.), ya que es completamente análogo.

El modo en que LaTeX interpreta la «forma» que debe tener el documento es mediante *etiquetas*. Por ejemplo, `\documentclass{article}` le dice a LaTeX que el documento que va a procesar es un artículo.

Las ventajas de trabajar con LaTeX son:

- No hay que recordar que "estilo" se usó para las secciones previas, las subsecciones, los capítulos y demás.
- Si la fuente del título de sección resulta ser demasiado grande no es necesario cambiarla en todo el documento, se puede hacer con una modificación sencilla.
- Al estar el documento marcado con secciones, subsecciones y demás, resulta más sencillo crear tablas de contenido (de hecho, se crean automáticamente).

Bibliografía

- Google Sites [en línea] www.google.com/sites/overview.html [Consultado el 29-11-10]
- Latex: A document preparation system [en línea] www.latex-project.org [Consultado el 29-11-10]
- March, S. y Smith, G. 1995. Design and natural science research on information technology. *Decision Support Systems*; 15, 251-266.
- Simon, H. 1996. *The Sciences of the Artificial*, 3rd ed. MIT Press.
- Zotero [en línea] www.zotero.org [Consultado el 29-11-10]