

Interoperabilidad entre el Repositorio Institucional y servicios en línea en la Universidad Nacional de La Plata

De Giusti, Marisa R.¹, Villarreal, Gonzalo L.², Terruzzi, Franco A.³, Oviedo, Nestor F.⁴, Lira, Ariel J.⁵

Resumen

La Universidad Nacional de La Plata tiene como objetivo prioritario la difusión de todo el conocimiento generado en la institución, a fin de devolver a la sociedad parte del esfuerzo invertido en la Universidad pública. Para alcanzar este objetivo se han generado diversas iniciativas desde la gestión, la creación de nuevos servicios y la implementación de nuevas líneas de investigación y desarrollo para potenciar tales servicios. El Servicio de Difusión de la Creación Intelectual es el Repositorio Institucional central de la UNLP, y posee en la actualidad todo tipo de materiales científicos y académicos producidos desde la Universidad, incluyendo artículos de revistas científicas, publicaciones en congresos, tesis de posgrado, tesinas, normativas y ordenanzas, libros y libros electrónicos, documentos de audio, materiales educativos, y fotografías de piezas de museos, entre otros. Este repositorio institucional, el mayor de la Argentina y uno de los principales de América Latina, posee un rol central en la difusión de la producción intelectual de la UNLP y en la coordinación con otros servicios y desarrollos, entre los que se destacan el Portal de Revistas de la UNLP, soportado por el software Open Journal Systems, y el Portal de Congresos de la UNLP, soportado por el software Open Conference Systems. En este trabajo se detallan los distintos mecanismos que se han implementado desde la Universidad Nacional de La Plata para facilitar la interacción entre su repositorio institucional SEDICI -construido sobre el software Dspace- y estos portales, evitar la duplicación de esfuerzos y maximizar la difusión abierta del conocimiento. Dichos mecanismos incluyen desarrollos tecnológicos como la utilización de diversos protocolos (RSS/Atom, Sword, OAI-PMH) para establecer un camino unificado de comunicación entre los sistemas, o el desarrollo de plugins que permiten exportar la información desde las herramientas de gestión de producción académica, en un formato comprensible por otros sistemas. Todo esto, además, involucra el establecimiento de flujos de trabajo entre los equipos de personas que conforman los distintos servicios, con la finalidad de obtener un mecanismo unificado que contemple los métodos y las herramientas para exportar los datos e incorporarlos al repositorio. Finalmente, en este trabajo se incluyen otros esfuerzos generados tanto desde el SEDICI como desde la presidencia de la Universidad para asegurar la preservación de toda la producción intelectual y brindar nuevos mecanismos para maximizar la visibilidad de estos documentos científicos y académicos. Esto abarca a otros actores y direcciones de esta institución tales como la Editorial de la Universidad (EDULP), la Dirección de Educación a Distancia (EAD) y la Radio Universidad.

¹ Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC GBA), Proyecto de Enlace de Bibliotecas (PrEBi) y Servicio de Difusión de la Creación Intelectual (SEDICI) de la Universidad Nacional de La Plata; marisa.degiusti@sedici.unlp.edu.ar

² Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Proyecto de Enlace de Bibliotecas (PrEBi) y Servicio de Difusión de la Creación Intelectual (SEDICI) de la Universidad Nacional de La Plata; gonetil@sedici.unlp.edu.ar

³ Portal de Revistas, Portal de Congresos y Servicio de Difusión de la Creación Intelectual (SEDICI) de la Universidad Nacional de La Plata; agustinterruzzi@sedici.unlp.edu.ar

⁴ Proyecto de Enlace de Bibliotecas (PrEBi) y Servicio de Difusión de la Creación Intelectual (SEDICI) de la Universidad Nacional de La Plata; nestor@sedici.unlp.edu.ar

⁵ Proyecto de Enlace de Bibliotecas (PrEBi) y Servicio de Difusión de la Creación Intelectual (SEDICI) de la Universidad Nacional de La Plata; alira@sedici.unlp.edu.ar

Interoperability between online services and the Institutional Repository in La Plata National University

De Giusti, Marisa R.¹, Villarreal, Gonzalo L.², Terruzzi, Franco A.³, Oviedo, Nestor F.⁴, Lira, Ariel J.⁵

Abstract

La Plata National University (UNLP, Universidad Nacional de La Plata) has as priority objective the dissemination of all knowledge generated in the institution, in order to return to the society part of the efforts invested in the public University. To that end, different initiatives have been launched, along with new services as well as research and development projects to foster these services. The Intellectual Creation Dissemination Service (SEDICI, Servicio de Difusión de la Creación Intelectual) is the UNLP Institutional Repository, and holds different classes of scientific and academic materials produced inside the University, including journal articles, conferences' publications, theses, ordinances, books, e-books and audio-documents. This institutional repository, the largest one in Argentina and one of the largest ones in Latin America, possesses a central role in the dissemination of UNLP's intellectual production and in the coordination with other projects and services such as UNLP's Journals Portal, supported by OJS, and UNLP's Conferences Portal, supported by OCS. In this article we explain the different mechanisms implemented from UNLP to ease the interaction between its institutional repository, SEDICI - supported by Dspace - and those portals, to avoid efforts duplication, and to maximize an open and free knowledge dissemination. Such mechanisms include technological developments based on different protocols (such as RSS/Atom, Sword, OAI-PMH) to establish an unified way of communication between systems, and the development of plugins that enable administrators to export information from the different tools in a format that could be understood by other systems. These involve the establishing of workflows among workgroups of the different services too, with the objective of building a unified mechanism that contemplates the methods and tools used to export/import data from and to the repository. In addition, other efforts driven from SEDICI as well as the University presidency are also mentioned in this document. These efforts seek to ensure the preservation of all intellectual production while they provide new mechanisms to maximize the visibility of the scientific and academic documents. Other actors and directions from the UNLP are also involved in such efforts, like the Publishing House EDULP, the Distance Education Direction (EAD), and the University Radio.

¹ Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC GBA), Proyecto de Enlace de Bibliotecas (PrEBi) y Servicio de Difusión de la Creación Intelectual (SEDICI) de la Universidad Nacional de La Plata; marisa.degiusti@sedici.unlp.edu.ar

² Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Proyecto de Enlace de Bibliotecas (PrEBi) y Servicio de Difusión de la Creación Intelectual (SEDICI) de la Universidad Nacional de La Plata; gonetil@sedici.unlp.edu.ar

³ Portal de Revistas, Portal de Congresos y Servicio de Difusión de la Creación Intelectual (SEDICI) de la Universidad Nacional de La Plata; agustinterruzzi@sedici.unlp.edu.ar

⁴ Proyecto de Enlace de Bibliotecas (PrEBi) y Servicio de Difusión de la Creación Intelectual (SEDICI) de la Universidad Nacional de La Plata; nestor@sedici.unlp.edu.ar

⁵ Proyecto de Enlace de Bibliotecas (PrEBi) y Servicio de Difusión de la Creación Intelectual (SEDICI) de la Universidad Nacional de La Plata; alira@sedici.unlp.edu.ar

1. Introducción

La Universidad Nacional de La Plata tiene entre sus objetivos prioritarios la difusión en abierto de su producción científica y académica, a fin de aumentar la visibilidad de la misma y devolver a la sociedad parte del esfuerzo invertido en la Universidad pública. La producción creciente de materiales científicos y académicos desde la propia Universidad, que incluye revistas periódicas y actas de congresos organizados por miembros de la institución, y la promoción para la generación constante estos materiales ha sido encarada desde múltiples frentes, entre los que se destacan subsidios para la creación y edición de revistas científicas para todos sus miembros, subsidios para la organización y participación en reuniones científicas para sus investigadores, y la creación de servicios y herramientas que sirvan de soporte para la gestión y publicación de su producción intelectual: revistas científicas, de extensión y de difusión, blogs de cátedras, herramientas de educación a distancia, software para la organización de reuniones científico-académicas, entre otros. Este trabajo se enfocará en particular sobre dos de estos servicios y sus herramientas, el Portal de Revistas [1] y el Portal de Congresos [2] de la UNLP.

La necesidad de albergar, preservar, difundir y dar visibilidad a la propia producción científica e intelectual es cada vez mayor en cualquier institución que realiza tareas de investigación, desarrollo y transferencia. Para alcanzar estas metas, muchas instituciones implementan repositorios institucionales centralizados y de acceso abierto. El directorio de Repositorios de Acceso Abierto (OpenDOAR [3], Directory of Open Access Repositories) cuenta hasta la fecha más de 2200 repositorios institucionales, entre los que se encuentran Saber.ula [4], de la Universidad de los Andes, Venezuela, RiUMA [5] de la Universidad de Málaga, España, la Biblioteca Digital [6] de la Universidad ICESI, Colombia, y el SEDICI [7] de Universidad Nacional de La Plata, Argentina.

El Servicio de Difusión de la Creación Intelectual (SEDICI) de la UNLP fue creado en el año 2003 con el objetivo de preservar y difundir la producción intelectual de la UNLP. Esta iniciativa representa uno de los principales esfuerzos de esta universidad para incrementar la visibilidad y difusión del trabajo de sus docentes, investigadores y alumnos. Para ello, el SEDICI constituye el Repositorio Institucional central de la UNLP, concentra la producción tanto de las Unidades Académicas como de los centros de investigación, y articula sus actividades con otras dependencias e iniciativas de esta institución como por ejemplo la Editorial de la UNLP (EDULP), el programa Radio Universidad y la Dirección de Educación a Distancia, Innovación en el Aula y TIC.

El Portal de Revistas de la UNLP es una iniciativa que tiene como objetivo principal promover la creación de revistas en los distintos ámbitos de la Universidad, incluyendo laboratorios y centros de investigación, facultades, grupos de trabajo y docentes e investigadores independientes. Esta iniciativa sigue en la misma línea de muchas instituciones a nivel mundial que han implementado servicios similares con el objetivo de maximizar el impacto y la visibilidad de sus revistas y sus docentes e investigadores. Por ejemplo, entre las instituciones que cuentan con este tipo de servicios –además de la UNLP– se cuentan la Universidad de Murcia, España [8], la Universidad de Chile [9], la Universidad Pontificia Bolivariana (Colombia)

[10], la Universidad Católica de Brasilia (Brasil) [11], y el Centro Argentino de Información Científica y Tecnológica (Argentina) [12].

Muchas de las instituciones que poseen portales de revistas utilizan el software de código abierto Open Journal Systems (OJS) como soporte. En la actualidad, de hecho, esta plataforma es utilizada por más de 14000 revistas científicas según se destaca en el sitio oficial de PKP OJS, actualizado en julio del 2012 [13]. El uso extendido de esta plataforma se debe al amplio soporte brinda para cada una de las etapas del proceso de publicación de una revista, desde el envío de trabajos, hasta su publicación e indexación, su facilidad de uso y su disponibilidad en múltiples idiomas. La sumatoria entre la cantidad de usuarios de la plataforma, su condición de libre y abierta, su estabilidad y la flexibilidad que ofrece a los gestores y editores fueron los principales factores que motivaron a la UNLP a implementar su Portal de Revistas con esta misma herramienta de software.

De manera similar a lo sucedido con el Portal de Revistas, la UNLP decidió implementar un servicio que brinde apoyo a todos sus investigadores y docentes a la hora de organizar cualquier tipo de reunión científica y/o académica. Este servicio se apoya en el software Open Conferences System (OCS). El equipo que mantiene al Portal de Congresos de la UNLP se encarga de dar soporte técnico a los organizadores, de organizar encuentros periódicos con los usuarios, de crear y brindar documentos y ayudas, y de mantener el servicio actualizado y en funcionamiento.

Uno de los desafíos surgidos desde el comienzo de la implementación de los portales de Revistas y de Congresos de la UNLP fue el establecimiento de un flujo de trabajo y el desarrollo de herramientas tecnológicas que permitan a estos servicios una colaboración de ida y vuelta con el SEDICI, a fin de asegurar la preservación de la producción intelectual de los miembros de la UNLP alojada en dichos portales, y minimizar el re-trabajo por parte de los integrantes de ambos servicios. Para ello, se implementaron diversas políticas de colaboración entre los equipos de trabajo de SEDICI, Portal de Revistas y Portal de Congresos, y se configuraron y ampliaron las herramientas que sustentan estos servicios para favorecer el intercambio de documentos entre ellos. Esto fue posible gracias al uso de herramientas libres y de código abierto: DSPACE como soporte del Repositorio Institucional SEDICI, Open Journal Systems para el Portal de Revistas, y Open Conference Systems para el Portal de Congresos. En este contexto, se buscó también sumar otros actores de diferentes espacios de la UNLP a estos esfuerzos de difusión y preservación.

En este trabajo se mencionan cuáles fueron las políticas implementadas, junto a los mecanismos y las tecnologías desarrolladas para concentrar, preservar y aumentar la visibilidad y la difusión de toda la creación intelectual de la UNLP. Si bien este trabajo se enfoca principalmente en los tres servicios mencionados en el párrafo precedente, con SEDICI como eje central, se mencionan también otros esfuerzos y desarrollos que van en el mismo sentido, tanto desde la presidencia de la UNLP como desde las distintas facultades.

2. OJS y Dspace

2.1 De OJS a Dspace

La articulación entre el Portal de Revistas de la UNLP y el SEDICI tiene como objetivo simplificar el mecanismo de depósito en el repositorio de los artículos alojados en el Portal de Revistas. Esto puede realizarse en dos momentos del ciclo de vida de las revistas dentro del portal: en el momento inicial, donde se requiere una importación masiva de todos los números anteriores de una revista para ser incorporados al repositorio (para aquellas revistas añejas que comienzan a utilizar el portal y migrar sus artículos al formato digital), y en el día a día, al momento de publicar nuevos números de las distintas revistas. Para realizar la integración entre OJS y DSpace se evaluaron diversas alternativas que consideraban dos posibles casos de uso para esta integración: por un lado, OJS aloja documentos (artículos) que deben ser mostrados en Dspace y por otro, muchas revistas que ya poseían un mecanismo de preservación y difusión a través de SEDICI han comenzado a utilizar el Portal de Revistas, con lo cual sus artículos deben ser enviados desde DSPACE hacia OJS. Para realizar esto se plantearon dos alternativas basadas en el intercambio de datos en línea y en herramientas de importación y exportación compatibles entre ambos sistemas:

Alternativa 1: Exportar los artículos desde OJS utilizando un formato de metadatos compatible entre los dos sistemas –en este caso Dublin Core [14]–. La exportación generaría un único archivo XML [15], que contiene para cada número de revista, sus secciones, los artículos organizados en estas secciones y los autores de cada artículo. Además, se añade al XML de cada artículo el documento completo (originalmente en un archivo pdf), codificado en base-64 [16]. La ventaja principal de esta alternativa reside en que el portal ya contiene herramientas de exportación a XML, por lo que sólo se requiere reforzar el desarrollo en las funciones de transformación de múltiples archivos XML aislados hacia un único archivo integrador, que luego deberá ser interpretado desde el DSpace de SEDICI. No obstante, el proceso intermedio de adaptar cada uno de los XML exportados y generar un XML global podría ser demasiado costoso para la importación de revistas con un gran historial de números y artículos.

Alternativa 2: Realizar la exportación desde OJS utilizando el plugin para Sword v1. Con esta alternativa se buscan aprovechar e incrementar las herramientas que ya ofrece el portal para realizar este tipo de operaciones; además Dspace también permite realizar depósitos vía sword v1 o sword v2, por lo que la comunicación entre los dos sistemas puede hacerse de manera simple y uniforme. Sin embargo, esta alternativa requiere adaptar de manera correcta los metadatos exportados en un formato que sea apto para la compartición mediante sword. Una gran ventaja de esta opción es que las herramientas que ofrecen ambos sistemas se encuentran disponibles y sólo requieren adaptación para el caso particular de la UNLP. Más allá de esto, no es posible asegurar una automatización completa con este método.

La alternativa 2 fue seleccionada en especial por su facilidad de desarrollo, ya que permitía obtener una solución aplicable en menor tiempo. La implementación se realizó a partir del plugin de exportación de Sword v1 que posee OJS. Esta herramienta construye lo que se conoce como un “paquete de depósito”, el cual contiene toda la información que puede considerarse pertinente sobre el o los artículos que se quieren depositar, transformados a base-64. Las restricciones propias de la estructura de metadatos interna utilizada por el SEDICI requirió algunas modificaciones menores en el mecanismo de construcción de este paquete, al cual se le agregó una mayor cantidad de datos relativos a cada artículo. Una primera adaptación se realizó para incorporar una fecha de publicación a cada artículo de OJS, ya que

en este sistema los artículos no poseen un campo propio con la fecha de publicación –pues esta información se encuentra en el número al que pertenecen– y esto es requerido para cada registro de SEDICI. La modificación en la herramienta de exportación de OJS permite adicionar la fecha de publicación del número en cuestión al que pertenece el artículo o, en última instancia, considere la fecha de incorporación al repositorio como fecha de publicación del artículo, a fin de evitar que este campo quede con valor nulo.

Nombre de campo en lenguaje natural	Nombre de campo en dspace	Nombre de campo en OJS
DOI	sedici.identifier.doi	doc.doi
Título	dc.title	doc.title
Título Alternativo	dc.title.alternative	doc.title
Autor	sedici.creator.person	doc.author.fullName
Fecha de publicación	dc.date.issued	doc.DatePublished
Idioma	dc.language	doc.language
Resumen	dc.description.abstract	doc.abstract
Tipo de Documento	dc.type	doc.type
Identificación del número	sedici.relation.journalVolumeAndIssue	journal.year, journal.number, journal.volume
Título de la revista	sedici.relation.journalTitle	journal.title

Tabla 1. Relación entre metadatos de OJS y de Dspace. La primer columna muestra el significado de cada campo, la columna central indica cómo se denomina cada campo en Dspace, ya sea definido por Dublin Core (dc) o propio de SEDICI (sedici) y la columna de la izquierda señala qué atributos del artículo (doc) o de la revista (journal) se utilizan para obtener cada metadato.

Para mantener la compatibilidad entre las estructuras de metadatos, algunos de ellos fueron suprimidos al realizar la exportación, como por ejemplo la *afiliación* del autor que no es considerada en la estructura de metadatos de SEDICI. Asimismo, el metadato *título* permite ahora múltiples instancias, lo que permitió el envío de múltiples títulos para un mismo artículo; esta es una característica propia de OJS, cuando un mismo artículo se presenta en múltiples idiomas. Por el mismo motivo, se alteró el metadato *resumen* a fin de permitir que también sean enviados varios resúmenes por artículo. El metadato de identificación de cada revista se construye la configuración de las mismas en OJS, ya sea mediante volumen, volumen y número, sólo número, volumen, número y año, etcétera. El metadato *idioma*, requerido por SEDICI en todos los documentos, fue incorporado también como campo adicional en cada documento de OJS enviado vía sword. En la Tabla 1 se muestran las adaptaciones de compatibilidad implementadas entre los metadatos del Portal de Revistas y del SEDICI. Además de realizar las modificaciones pertinentes en OJS al momento de exportar los documentos, se implementaron algunas modificaciones en el software de Dspace de SEDICI, en particular en su *crosswalk*, encargado de transformar los datos recibidos mediante SWORD (ver ejemplo 1).

Una vez que se realizaron los ajustes necesarios tanto en OJS como en Dspace, se implementó el mecanismo de migración en un contexto real, enviando el último número de una de las revistas del Portal de Revistas hacia el repositorio institucional. Los artículos fueron incorporados en una colección de DSpace, utilizada también para el autoarchivo por parte de los docentes e investigadores de la UNLP, y permaneciendo ocultos a la espera de una revisión general por parte de los administradores de SEDICI para ser publicados definitivamente. Los resultados satisfactorios fomentaron un segundo ciclo de mejoras, con el objetivo complementario de aprovechar las ventajas que ofrece Sword 2. Esta versión permite dar soporte durante todo ciclo de vida de un documento depositado, agregando funcionalidades tales como la posibilidad de modificar el documento una vez que ya ha sido depositado, de completar datos incompletos, de ofrecer una mayor tolerancia a los errores que surjan en cualquier parte del proceso y de eliminar remotamente un documento depositado [17].

The image shows a web interface for the SWORD connector. On the left, there is a form titled "Conector de depósito de importación/exportación SWORD" with fields for "Punto de depósito" (set to "sedici"), "Nombre usuario/a", "Contraseña", and "Punto de depósito" (set to "Autoarchivo"). There are checkboxes for "Depositar galeradas" and "Depositar los archivos de editorial más recientes". Below the form is a table of articles to be uploaded:

NÚMERO	TÍTULO	AUTORES/AS
<input checked="" type="checkbox"/> VOL_1 (2009)	Aplicação de processo oxidativo avançado baseado em fotocatalise heterogênea (TiO2 /UVsolar) para o pré-tratamento de efluente lácteo	
<input checked="" type="checkbox"/> VOL_3 (2011)	Avaliação comparativa de iscas atrativas a partir da riqueza de espécies de formigas (Hymenoptera: Formicidae) numa floresta de Eucalyptus grandis, em Santa Maria, Rio Grande do Sul, Brasil	
<input checked="" type="checkbox"/> VOL_3 (2011)	Avaliação da Redução da Poluição do Chorume Tratado por Processo Fotoquímico	

On the right, the SEDICI logo and navigation menu are visible. Below, the "Tareas del flujo de trabajo" section shows a list of tasks:

Estas tareas son ítems que están esperando aprobación antes de ser añadidas al repositorio. Hay dos colas de tareas, una para tareas que ha aceptado y otra para tareas que todavía no han sido asumidas por nadie.

Sus tareas

Tarea	Ítem	Colección	Remitente
No tiene asignada ninguna tarea			

Tareas en cola

Tarea	Ítem	Colección	Remitente
<input type="checkbox"/> Revisión SEDICI	Aplicação de processo oxidativo avançado baseado e ...	Autoarchivo	Portal de Revistas UNLP
<input type="checkbox"/> Revisión SEDICI	Avaliação comparativa de iscas atrativas a partir ...	Autoarchivo	Portal de Revistas UNLP
<input type="checkbox"/> Revisión SEDICI	Avaliação da Redução da Poluição do Chorume Tratad ...	Autoarchivo	Portal de Revistas UNLP

Imagen 1: Envío de artículos vía Sword desde el Portal de Revistas de la UNLP hacia el SEDICI. Los artículos son enviados a la colección de Autoarchivo del repositorio institucional, para luego ser revisados antes de publicarse. Para este tipo de autoarchivo, el remitente es el Portal de Revistas, lo cual ayuda a los administradores de SEDICI a diferenciarlo del autoarchivo tradicional realizado por los usuarios.

```

<?xml version="1.0" encoding="utf-8"?>
<xsl:stylesheet
  xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
  xmlns:dim="http://www.dspace.org/xmlns/dspace/dim"
  xmlns:epdcx="http://purl.org/eprint/epdcx/2006-11-16/" version="1.0">

  <xsl:template match="text()"></xsl:template>

  <xsl:template match="/epdcx:descriptionSet">
    <dim:dim>
      <xsl:call-template name="el_titulo"/>
      <xsl:apply-templates/>
    </dim:dim>
  </xsl:template>

  <xsl:template match="/epdcx:descriptionSet/epdcx:description/epdcx:statement">
<xsl:if test="./@epdcx:propertyURI='http://purl.org/dc/terms/isPartOf'">
  <dim:field mdschema="sedici" element="relation" qualifier="journalVolumeAndIssue">
    <xsl:value-of select="substring-before(epdcx:valueString, '*')"/>
  </dim:field>
  <dim:field mdschema="sedici" element="relation" qualifier="journalTitle">
    <xsl:value-of select="substring-after(epdcx:valueString, '*')"/>
  </dim:field>
</xsl:if>
<xsl:template name="el_titulo">
  <xsl:if
test="epdcx:description/epdcx:statement[@epdcx:propertyURI='http://purl.org/dc/elements/1.1/title
']][1]">
    <dim:field mdschema="dc" element="title">
      <xsl:value-of
select="epdcx:description/epdcx:statement[@epdcx:propertyURI='http://purl.org/dc/elements/1.1/tit
le']][1]/epdcx:valueString"/>
    </dim:field>
  </xsl:if>
</xsl:template>

<xsl:template
match="/epdcx:descriptionSet/epdcx:description/epdcx:statement[@epdcx:propertyURI='http://purl.or
g/dc/elements/1.1/title']">
  <dim:field mdschema="dc" element="title" qualifier="alternative">
    <xsl:value-of select="epdcx:valueString"/>
  </dim:field>
</xsl:template>
<xsl:template
match="/epdcx:descriptionSet/epdcx:description/epdcx:statement[@epdcx:propertyURI='http://purl.or
g/dc/elements/1.1/title']][1]">
  </xsl:template>
</xsl:stylesheet>

```

Ejemplo 1: Parte del código XSLT adaptado del *crosswalk* de DSpace que transforma los metadatos enviados desde OJS via Sword 1 al formato de metadatos de SEDICI.

2.2 De Dspace a OJS

El repositorio institucional de la UNLP fue creado casi 6 años antes de la implementación del Portal de Revistas. Esto generó la existencia de publicaciones en el SEDICI desde hace muchos años, y algunas de ellas han decidido recientemente migrar su sistema de gestión y comunicación al Portal de Revistas. En este caso, surgió la necesidad de invertir el proceso de integración entre OJS y Dspace, ya que las revistas deseaban también que el nuevo sitio web muestre todos los volúmenes y números anteriores, pero resultaba muy costoso cargar manualmente todos estos números con sus metadatos y artículos completos.

Dado que Dspace cuenta con herramientas de exportación de colecciones, y que OJS posee mecanismos de importación masiva, se buscó de aprovechar estos desarrollos, realizando las adaptaciones pertinentes a fin de asegurar la compatibilidad entre ambos sistemas. Durante la etapa de análisis de alternativas, se propusieron diversas opciones para integrar los documentos de SEDICI en el Portal de Revistas:

Opción 1: Establecer la comunicación desde Dspace hacia OJS mediante la utilización de canales o *feeds* RSS/Atom [18]. Para volcar información desde Dspace se podría utilizar un canal RSS que, ya sea periódicamente o a demanda, actualice la información requerida por OJS. Esto requiere implementar también un plugin lector de RSS para OJS, lo cual es relativamente simple puesto que esta herramienta ya utiliza RSS para exponer cierta información. La principal ventaja de esta alternativa es su facilidad de implementación y la rapidez con la cual se podía llegar a una solución que funcione. No obstante, existen algunas desventajas a considerar: los canales RSS son asincrónicos, por lo que para que el funcionamiento sea adecuado, el lector de *feeds* incorporado a OJS debía estar activo en todo momento a la espera de recibir información desde Dspace. Esto se puede evitar si se cuenta con una persona que realice esta tarea periódicamente, que sincronice el envío desde Dspace y la captura desde OJS. De todos modos, existe una dificultad mayor: para el caso de los artículos contenidos en Dspace, no es posible enviar adjunto el archivo con el artículo, ya que el estándar RSS no soporta *a priori* el envío de documentos adjuntos.

Opción 2 : Realizar la exposición desde Dspace y recolección en OJS por medio del protocolo OAI PMH [19]. Dspace ya implementa un data-provider de OAI, lo que simplifica la implementación de esta opción. Desde el lado de OJS, también se dispone de un plugin que proporciona un data-provider para OAI; no obstante, en este caso resultaba necesario desarrollar un service-provider de OAI capaz de recibir el contenido expuesto desde Dspace. La ventaja principal de esta opción es el uso de un mecanismo ampliamente utilizado para compartir contenidos, por lo que implementar un plugin con un service-provider para OJS podría ser reutilizado en el futuro. La desventaja principal de esta solución es básicamente la misma que presenta la solución anterior, puesto que OAI es un mecanismo pensado para compartir contenido pero no para enviar archivos adjuntos, lo cual aún requería que el editor de la revista importe los documentos completos manualmente.

Opción 3: Conectar Dspace con OJS mediante Sword. En este caso se plantea el camino inverso de la solución escogida para la integración de OJS hacia Dspace. Para esto, es posible utilizar el sword-client incorporado en Dspace, y es necesario construir un plugin o sistema similar que permitiera ejecutar un sword-server en el contexto de OJS. La ventaja de esta resolución radica en la unificación del mecanismo de comunicación entre los dos sistemas. Además, como ya se había visto antes, sword permite enviar archivos adjuntos, por lo que se podían enviar todos los artículos y sus datos desde Dspace, previa transformación entre los datos de uno y de otro sistema. La desventaja más importante aquí es la necesidad de implementar un servidor sword por completo, lo cual requiere un mayor esfuerzo de desarrollo. Si bien de momento esta opción fue rechazada, no se descarta de plano su implementación a mediano plazo.

Opción 4: Considerar las formas de exportación ya implementadas en Dspace, como por ejemplo la exportación a un paquete AIP [20], y realizar un plugin para OJS que sea capaz de

interpretar la estructura de esta exportación, tomar los datos y crear los artículos a partir de los mismos. Esta opción permite una implementación relativamente rápida para establecer un camino de comunicación directo entre Dspace y OJS. La idea aquí es establecer un mecanismo disparado manualmente, en el cual un administrador del repositorio genere un único archivo con la exportación completa de una revista en el repositorio y lo envíe a un administrador del Portal de Revistas, quien deberá tomar este archivo e importarlo en OJS por medio del plugin desarrollado para tal fin. La implementación de esta solución es relativamente simple, brinda la posibilidad de procesar un gran número de registros juntos, y permite incluir también los archivos adjuntos como parte del mecanismo de exportación-importación. No obstante, una desventaja de esta solución es su grado de acoplamiento entre las estructuras que gestionan ambos sistemas, lo que restringe su reutilización en otros contextos.

La opción 4 fue la escogida aquí, mediante la implementación de un nuevo plugin de importación para OJS, el cual toma el paquete AIP generado por Dspace, analiza la información que éste contiene y, a partir de la misma, crea los números y artículos de la revista en OJS. Esto debe aplicarse a cada revista que se desea incorporar a OJS. El plugin desarrollado debe ser utilizado por un usuario con rol de gestor de revista, así como también por el administrador del sistema. Este plugin dispone de varias formas de importación para escoger y distintos medios desde donde importar, como por ejemplo importaciones desde sistemas como OCS y OJS, e importación desde Dspace por medio del formato AIP. Además para cada una de estas mejoras, se permitió el procesamiento del paquete de exportación completo, descomprimiendo el archivo zip y creando los artículos necesarios, junto a sus metadatos (autores, fechas, títulos, números) correspondientes.

Para incorporar a OJS e indexar cada artículo, se reutiliza parte del plugin de importación denominado *QuickSubmit*. Este plugin permite a distintos usuarios publicar artículos fácilmente, mediante el envío de todos los datos necesarios a través de un formulario web. Se realizaron modificaciones al código de dicho plugin para que tome los datos previamente descomprimidos, como si hubiesen sido enviados a través del formulario web, y cree el artículo correspondiente del mismo modo que lo haría si un usuario lo hubiera cargado manualmente. Esta solución, denominada *QuickImport plugin*, fue diseñada de forma modular para simplificar la incorporación de otros medios de exportación desde Dspace o desde otros sistemas, y su utilización ha resultado productiva tanto desde el Portal de Revistas como desde el SEDICI. Como consecuencia de todo esto, ya existen ciertas iniciativas para aprovechar la solución diseñada y utilizarla en otros ámbitos de exportación o de importación, como por ejemplo, la integración del Portal de Congresos con el Repositorio Institucional.

3. Portal de Congresos y SEDICI

El Portal de Congresos de la UNLP es un servicio que se encuentra disponible desde principios del 2009, con el objetivo de ofrecer una herramienta de gestión virtual a todos los miembros de la institución que deseen organizar cualquier tipo de reunión científica y/o académica. Para realizar esto, el equipo del portal realiza reuniones y capacitaciones en el uso del software Open Conference Systems, que se utiliza para la gestión de los congresos. Durante estas reuniones, se informa a los organizadores del congreso sobre las ventajas y la importancia de publicar en abierto todos los artículos y presentaciones. Para maximizar su difusión y visibilidad, se les ofrece también la inclusión de la producción científico-académica del

congreso en el repositorio de la UNLP. Desde luego, este trabajo se realiza de manera transparente para los organizadores, y sólo deben autorizar explícitamente la difusión de estos materiales.

Una vez que las actas del congreso han sido publicadas, el equipo de SEDICI realiza la recolección de todas las publicaciones. Para simplificar esta labor, el equipo del portal de congresos desarrolló un plugin para OCS [21] capaz de exportar las publicaciones y autores en múltiples formatos (texto plano, .doc, .pdf, .html, .csv, .zip, entre otros). La información generada desde el plugin sirve de guía al equipo de SEDICI para asegurarse la correcta carga y catalogación de todos los documentos generados en el congreso. Dado que esta herramienta ha sido muy utilizada recientemente, el equipo del Portal de Congresos ha comenzado a diseñar una solución semi-automática, capaz de aprovechar la flexibilidad del plugin ya generado, y también de generar un archivo de exportación de OCS que pueda ser importada en el Dspace que sustenta a SEDICI. Asimismo, se está evaluando la integración de OCS con DSpace por medio de Sword2, de manera similar al trabajo de análisis que se está realizando para la integración OJS-Dspace.

4. Otros canales de difusión

Las actividades dedicadas a maximizar la visibilidad de la producción intelectual de la UNLP abarcan también a otros actores, servicios y políticas implementadas desde la Universidad. Los mecanismos de difusión impulsados por los distintos grupos de la UNLP son posibles gracias al constante estímulo generado desde la propia presidencia de la Universidad. En este sentido, en el año 2011 se firmó la resolución 78/11 que establece que las tesis de Maestría y Doctorado desarrolladas en la UNLP serán difundidas y preservadas en formato digital a través del SEDICI. Esta resolución contempla un período máximo de 24 meses en la cual la publicación de la tesis puede ser demorada, para aquellos casos donde sea necesario, por ejemplo para tramitar patentes. Al momento del depósito, los tesistas deben firmar una licencia de distribución no exclusiva [22], la cual indica que “la obra se pondrá a disposición del público para que haga de ella un uso justo y respetuoso de los derechos de autor siendo requisito cumplir con las condiciones de la licencia de uso”. En el mismo cuerpo de la licencia de distribución, el autor debe además elegir una licencia de uso (qué usos permite para su obra) entre las distintas posibilidades que ofrecen las licencias Creative Commons [23]. Para asegurar el depósito de la tesis por parte de los graduados de las diferentes carreras de posgrado, la Universidad ha establecido que el SEDICI debe emitir un certificado de depósito, el cual es requerido a la hora de tramitar el título de posgrado. De este modo, la Universidad se asegura, a través de su repositorio institucional central, la preservación y difusión de estos valiosos materiales.

Además de SEDICI, cada vez más facultades de la UNLP cuentan también con repositorios propios, en lo cuales almacenan documentos similares a los que se almacenan en el repositorio central SEDICI (tesis, artículos, trabajos en congresos) pero también otros documentos como planes de estudio de las materias y resoluciones internas. Muchos de estos repositorios implementan el protocolo OAI PMH, mediante el cual exponen todos sus documentos. En estos casos, el equipo de SEDICI realiza un harvesting inicial por medio de dicho protocolo, con el cual se obtiene un subconjunto de la gran cantidad de registros alojados en estos repositorios.

Este subconjunto incluye todos los documentos que SEDICI gestiona, y excluye otros que no deben formar parte del repositorio central de la UNLP. Los registros obtenidos son luego procesados y normalizados, para finalmente incorporarlos a la base de datos de SEDICI con el objetivo de exponerlos en su portal web; no obstante, en el portal de SEDICI se mantiene un enlace al registro original en el repositorio desde el cual se obtuvo. Una vez finalizada la exportación inicial, el equipo de SEDICI realiza una tarea de monitoreo permanente, en busca de nuevos registros incorporados a estos repositorios pasibles

Otra mecanismo de difusión que está comenzando a aplicarse consiste en la publicación en otros portales de un subconjunto de registros, con publicaciones y trabajos que posee el repositorio central de la UNLP y que cumplen con algún filtro en particular: un autor, un tema, un laboratorio, un tipo de documento, etcétera. Los registros son recuperados mediante una consulta por medio del protocolo OpenSearch [24], cuyos resultados son estructurados bajo el formato Atom y enviados al *cliente* mediante HTTP. Esto permite, por ejemplo, que un centro de investigación publique todos los documentos de sus investigadores en su portal web, o incluso que un investigador publique en su sitio web personal sus propias publicaciones. Si bien hoy en día este servicio ya funciona, es necesario contar con un mecanismo que permita exponer estos registros en otros portales de manera simple. Para ello, se ha integrado un plugin de Wordpress [25], el popular CMS de código abierto. Este plugin está siendo utilizado durante el desarrollo del nuevo portal web del Proyecto de Enlace de Bibliotecas de la UNLP (PREBI) [26], en el que se exponen por un lado todas las publicaciones del grupo de investigadores, y por el otro se muestran las publicaciones de cada investigador en las páginas personalizadas de cada uno. Este servicio ofrece grandes ventajas, como ser la facilidad para mantener actualizada la lista de publicaciones en las distintas páginas del portal de PREBI, el incremento de la visibilidad de las publicaciones de la UNLP por medio de su difusión en otros portales, y la promoción del uso del repositorio de la Universidad por parte de los investigadores y centros de investigación, como principal medio para resguardar y difundir sus trabajos y desarrollos.

5. Conclusiones

En este trabajo se han mencionado muchos de los esfuerzos que realiza la Universidad Nacional de la Plata para asegurar la preservación y maximizar la difusión y la visibilidad de su producción académica y científica. Muchos de estos esfuerzos se han centrado en mejorar los mecanismos de interacción entre los distintos servicios que brinda la Universidad a sus miembros. En muchos casos esto ha requerido la implementación de mejoras y ampliaciones en las distintas herramientas de software utilizadas para brindar dichos servicios, como ser Open Journals Systems, Open Conferences Systems y DSpace. En otros casos, los esfuerzos son posibles gracias al trabajo conjunto de los distintos grupos y actores de la Universidad, mediante la implementación de políticas administrativas y de servicios que promuevan el uso del repositorio como principal medio de difusión del conocimiento generado.

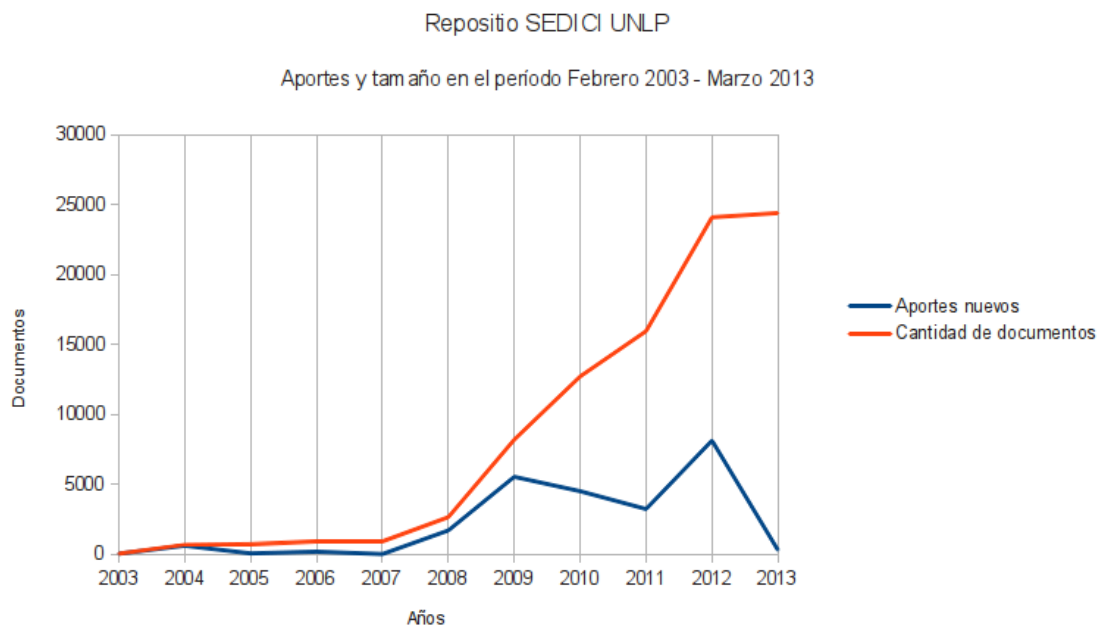


Imagen 2. Crecimiento del SEDICI en los últimos 10 años (más es mejor), desde Febrero de 2003 hasta Marzo de 2013. El tamaño del repositorio de la UNLP ha presentado un crecimiento exponencial desde el período 2008-2009, en el cual la cantidad de documentos incorporados interanualmente triplica los aportes en periodos anteriores.

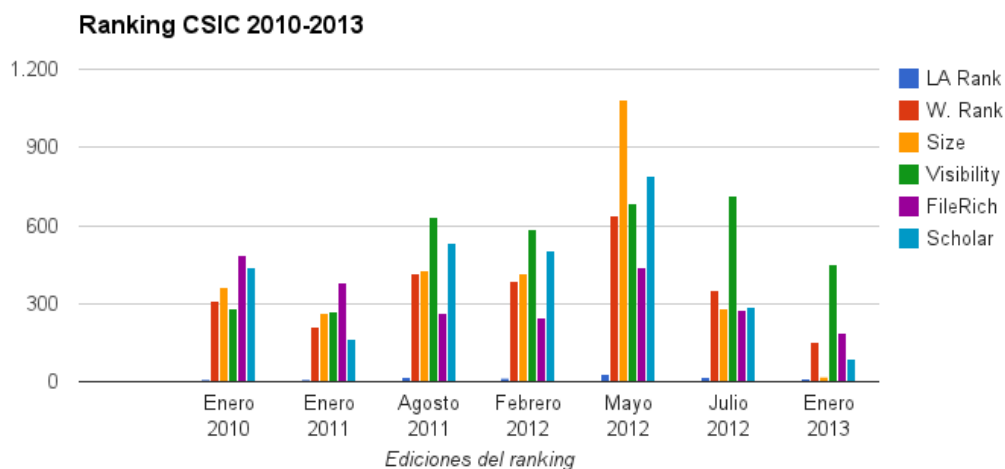


Imagen 3. Posición de SEDICI en el ranking CSIC desde el año 2010 (menos es mejor). Si bien se observa un retroceso en mayo de 2012, período durante el cual se realizó la migración hacia Dspace, en las últimas ediciones todos los indicadores (Tamaño, visibilidad, archivos ricos y Google Scholar) han mejorado notablemente, lo que ha repercutido tanto en el ranking de Latinoamérica (LA Rank) como en el ranking global (W. Rank)

Como resultado, el repositorio SEDICI de la UNLP ha incrementado su tamaño de aproximadamente 2000 documentos en el año 2008 [27] a más de 24000 a fines del año 2012,

lo que representa un incremento superior al 1200% en 4 años (Imagen 2). Gracias a este crecimiento exponencial, el SEDICI se ha colocado como el principal repositorio de la Argentina en cuanto a visibilidad, el segundo más grande en tamaño [28], y como uno de los 10 repositorios más grandes de Latinoamérica [29]. Esto ha servido para incrementar la visibilidad de la producción intelectual y de la Universidad en su conjunto, ubicandola en el puesto 13 en América Latina y en el puesto 323 a nivel mundial (imagen 3) según el último ranking de Universidades [30] publicado por el Consejo Superior de Investigaciones Científicas de España (CSIC) [31]. Estos resultados confirman que los esfuerzos dedicados a la preservación y difusión de la creación intelectual de la UNLP han resultado muy positivos, e incentivan a todos los actores a continuar trabajando en el mismo sentido.

Referencias

- [1]. Portal de Revistas de la UNLP. Web. 5 de Marzo de 2013. www.revistas.unlp.edu.ar
- [2]. Portal de Congresos de la UNLP. Web. 5 de Marzo de 2013. www.congresos.unlp.edu.ar
- [3]. Directory of Open Access Repositories. Web. 5 de Marzo de 2013. www.openoar.org
- [4]. SABER-ULA. Repositorio Institucional de la Universidad de Los Andes. Web. 5 de Marzo de 2013. www.saber.ula.ve/
- [5]. RIUMA - Repositorio Institucional de la Universidad de Málaga. Web. 5 de Marzo de 2013. www.riuma.uma.es/xmlui.
- [6]. Biblioteca Digital - Universidad Icesi. Web. 5 de Marzo de 2013. http://bibliotecadigital.icesi.edu.co/biblioteca_digital/
- [7]. Servicio de Difusión de la Creación Intelectual de la UNLP. Web. 5 de Marzo de 2013. <http://sedici.unlp.edu.ar/>
- [8]. Universidad de Murcia. Web. 5 de Marzo de 2013. www.um.es/
- [9]. Universidad de Chile. Web. 5 de Marzo de 2013. www.uchile.cl/
- [10]. Universidad Pontificia Bolivariana. Web. 5 de Marzo de 2013. www.upb.edu.co
- [11]. Universidade Católica de Brasília. Web. 5 de Marzo de 2013. www.ucb.br/
- [12]. CAICYT - CONICET. Centro Argentino de Información Científica y tecnológica. Web. 5 de Marzo de 2013. www.caicyt.gov.ar
- [13] "OJS Keeps on Growing". Web. 5 de Marzo de 2013. <http://pkp.sfu.ca/node/5323>
- [14]. Harper, Corey A. *Dublin Core metadata Initiative: beyond the element set*. National information standards Organization (NISO), winter 2010. Impreso
- [15]. Norman, Walsh. *Extensible Markup Language (XML) 1.0 (Fifth Edition)*. W3C Recommendation, 26 November 2008. En línea
- [16]. S, Josefsson. *The Base16, Base32, and Base64 Data Encodings*. RFC 4648, Octubre 2006. En línea
- [17]. R, Jones, Stuart Lewis. *Sword 2.0 Profile*. swordapp.org, 20 de Marzo de 2012. En línea
- [18]. M, Nottingham, R. Sayre. *The Atom Syndication Format*. RFC 4287, Diciembre 2005. En línea
- [19]. Rusch-Feja, D. "The Open Archives Initiative and the OAI Protocol for Metadata Harvesting: rapidly forming a new tier in the scholarly communication infrastructure." *learned Publishing*, 15(3), 179-186.
- [20]. R, Wolfe, W. Reilly. *DSpace METS document profile for submission information packages (SIP)*. 2006. En línea.

- [21]. "Multiple report plugin". Web. 5 de Marzo de 2013.
<http://pkp.sfu.ca/support/forum/viewtopic.php?f=28&t=8810>
- [22]. "Licencia de distribución no exclusiva SeDiCI". Web. 5 de Marzo de 2013.
http://sedici.unlp.edu.ar/static/resources/licencia_de_deposito.pdf
- [23]. *Creative commons licences*. Commons, C. (2005). En línea.
- [24]. D., Clinton. *OpenSearch Specifications, 1.1, Draft 3. A9. com*. (2006). En línea
<http://www.opensearch.org/Specifications/OpenSearch/1.1>.
- [25]. T., Polychniatis, S. van der Rijst, R. van Vliet, G. Wirken. *WordPress*. 2010. En línea
- [26]. PrEBi - Proyecto de Enlace de Bibliotecas. Web. 5 de Marzo de 2013
www.prebi.unlp.edu.ar
- [27]. Villarreal, G. L., De Giusti, M. R., Sobrado, A., Lira, A. J., Vila, M. M., & PrEBi, U. N. L. P. (2008). SeDiCI (Servicio de Difusión de la Creación Intelectual). *D-Lib Magazine*, 14(9/10), 1082-9873.
- [28]. Country matches any of "Argentina" - Registry of Open Access Repositories. Web. 5 de Marzo de 2013. http://roar.eprints.org/cgi/roar_search/advanced?location_country=ar&software=&type=&order=-recordcount%2F-date
- [29]. Latin America - Ranking Web of Repositories. Web. 5 de Marzo de 2013.
http://repositories.webometrics.info/en/Latin_America
- [30]. LatinoAmérica - Ranking Web de Universidades. Web. 5 de Marzo de 2013
http://www.webometrics.info/es/Latin_America_es
- [31]. Consejo Superior de Investigaciones Científicas - CSIC - csic.es. Web. 5 de Marzo de 2013. <http://www.csic.es/>