

III Conferencia de Bibliotecas y Repositorios Digitales de América Latina y VIII Simposio Internacional de Bibliotecas Digitales. Universidad de Costa Rica, San José, 2013.

# Infraestructura interoperable alrededor del repositorio institucional SEDICI.

De Giusti, Marisa Raquel, Lira, Ariel Jorge, Villarreal, Gonzalo Luján y Terruzzi, Franco Agustín.

Cita:

De Giusti, Marisa Raquel, Lira, Ariel Jorge, Villarreal, Gonzalo Luján y Terruzzi, Franco Agustín (Octubre, 2013). *Infraestructura interoperable alrededor del repositorio institucional SEDICI. III Conferencia de Bibliotecas y Repositorios Digitales de América Latina y VIII Simposio Internacional de Bibliotecas Digitales. Universidad de Costa Rica, San José.*

Dirección estable: <https://www.aacademica.org/marisa.de.giusti/88>

ARK: <https://n2t.net/ark:/13683/ptyc/WTv>



Esta obra está bajo una licencia de Creative Commons.  
Para ver una copia de esta licencia, visite  
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>.

*Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.*

# Infraestructura interoperable alrededor del repositorio institucional SEDICI

El repositorio **SEDICI** fue creado en 2003 con el objetivo de preservar y socializar la producción intelectual generada desde la **Universidad Nacional de La Plata**. Su naturaleza de repositorio institucional central, en una universidad de más de 100 años, lo ha llevado a incorporar y gestionar un gran volumen de material de tipología y temática muy variada, y a su vez proveniente de distintos orígenes. Para lograr sus objetivos, se ha planteado desde su comienzo la articulación institucional, hacia adentro de la UNLP y hacia otras instituciones, basada en dos ejes de trabajo principales: preservar la producción intelectual de la Universidad y diseminarla tanto en otros espacios de producción de contenidos hacia el interior de la UNLP como hacia el exterior de la misma (bases de datos centralizadas, proveedores de servicios en general, instituciones académicas y de investigación).

Ing. **Marisa R. De Giusti** marisa.degiusti@sedici.unlp.edu.ar  
 Lic. **Ariel J. Lira** alira@sedici.unlp.edu.ar  
 Dr. **Gonzalo L. Villarreal** gonzalo@prebi.unlp.edu.ar  
 Franco A. **Terruzzi** agustinterruzzi@sedici.unlp.edu.ar

1 Proyecto de Enlace de Bibliotecas - Servicio de Difusión de la Creación Intelectual (PREBI-SEDICI), Universidad Nacional de La Plata, La Plata (1900), Argentina  
 2 Investigador independiente, Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC), Proyecto de Enlace de Bibliotecas



# SEDICI

REPOSITORIO INSTITUCIONAL DE LA UNLP

sedici.unlp.edu.ar



## Interacciones

Para Los portales de revistas de la UNLP gestionados con el sistema **OJS**, se desarrolló una extensión o plugin que permite importar estos paquetes **AIP** y extraer de ellos los datos correspondientes a los artículos y números de revistas existentes en el repositorio institucional **SEDICI**.

**DSPACE** provee un generador de contenidos en formato **RSS/ATOM**, accesible para cada comunidad o colección en el repositorio. Este formato no solo es utilizado por usuarios sino por otras aplicaciones, que recuperan las actualizaciones de determinadas partes del repositorio, como por ejemplo "El portal de Revistas Científicas" que recupera y muestra los recursos de algunas revistas seleccionadas.

**SWORD** se ha aprovechado en **SEDICI** para permitir la publicación y envío de los nuevos contenidos de las revistas gestionadas por el **Portal de Revistas de la UNLP** y otros portales de distintas unidades académicas que utilizan el sistema **OJS**. Para esto se ha utilizado una extensión (plugin) desarrollada por el equipo del **SEDICI** y distribuida entre las distintas unidades académicas de la UNLP.

**Dspace** provee un servicio de búsqueda **Open Search** que permite buscar y recuperar los items del repositorio, en un contexto global o acotado a una comunidad o colección específica. Este servicio habilita al portal de la UNLP, entre otros, a recuperar obras de una colección específica de documentos institucionales en formato **ATOM**, igual a una sindicación tradicional.

Los metadatos del repositorio se transforman para cumplir con la especificación del **SNRD**, y se exponen a través de un servidor **OAI-PMH** ad-hoc. El formato del **SNRD** es **DC Simple** más algunas extensiones similares a las establecidas en las directrices **Driver** y **Open-Aire**.

Se adecúan los metadatos al formato **Dublin Core** simplificado, entendido por todos los agregadores **OAI-PMH**.

Los datos de recursos disponibles en Acceso Abierto dentro de los repositorios **ISTEC**, son recuperados desde el Metabusador bajo el protocolo **SRU**, utilizando un adaptador **SRU-a-SOLR** que traduce una solicitud **SRU** a una búsqueda interna **Solr**, el indexador utilizado por el **OPAC** para guardar y recuperar la información.

Antes de exponer los datos en su propio **OAI**, el **SNRD** transforma los registros almacenados para que sean compatibles con la especificación de **LaReferencia**.

## Glosario Aplicaciones

- DSPACE** Plataforma de código abierto para el archivo y publicación de recursos digitales. Permite organizar el repositorio partir de comunidades jerárquicas y colecciones que contienen los recursos, representados con metadatos configurables y archivos digitales.
- PAZPAR2** Es un software de metabúsqueda de alta performance, independiente del modelo de datos y de la interfaz de usuario, y posee características avanzadas como unión de resultados, ranqueo por relevancia, ordenamiento y faceting, generado por la empresa **Index Data** www.indexdata.com.
- OJS** **Open Journals System** es un software de código abierto para la administración de revistas creado por el **Public Knowledge Project**, liberado bajo licencia **GNU General Public License**.
- OCS** **Open Conferences System** es un software de código abierto para la administración de congresos que propone un sistema de gestión estándar que se aplica en varios lugares del mundo. Fue creado por el **Public Knowledge Project** y liberado bajo licencia **GNU General Public License**.
- SOLR + SRU** Este desarrollo permite realizar búsquedas remotas sobre bases de datos **Solr** bajo el estándar **SRU (Z39.50)**, lo que permite a repositorios funcionar como otros **OPAC** de bibliotecas.

## Portal de revistas

El Portal de Revistas es un servicio provisto por la UNLP para la gestión de publicaciones periódicas que se producen en el marco de la universidad. La plataforma que sustenta el portal está construida sobre el software **OJS** de **PKP**.

## Revistas científicas UNLP

Espacio provisto por la **Secretaría de Ciencia y Técnica** con el fin de difundir y promover las publicaciones periódicas científicas desarrolladas en el ámbito de la Universidad. De esta manera se consolida una larga tradición en las políticas de edición científica que se remonta a la fundación de nuestra Universidad.

## Otros portales de revistas UNLP

Existen otras revistas que son gestionadas por distintos sitios de las correspondientes unidades académicas de la UNLP. Estas pueden variar en contenido y formato, pero son en general administradas por sistemas similares al del Portal de Revistas.

## Portal de congresos

Plataforma de gestión de congresos que presta un servicio para todos los miembros de la UNLP que deseen organizar un congreso del ámbito académico de la UNLP. Este portal, construido sobre la plataforma **OCS** de **PKP**, permite a cada congreso gestionarse de manera completamente autónoma.

## Sitios UNLP

La UNLP y sus proyectos internos poseen muchos portales en los cuales se muestran recursos alojados en **SEDICI**. Algunos ejemplos son sitios institucionales de: cátedras, laboratorios, departamentos y facultades.

## Repositorios UNLP

Además de **SEDICI**, el repositorio central de la UNLP, algunas facultades también han implementado sus propios repositorios para preservar y difundir materiales específicos generados desde su unidad académica.

## SNRD

El **Sistema Nacional de Repositorios Digitales (SNRD)** es una iniciativa gubernamental cuyo propósito es conformar una red interoperable de repositorios digitales en ciencia y tecnología en Argentina, a partir del establecimiento de políticas, estándares y protocolos comunes a todos los integrantes del Sistema.

## OPAC ISTEAC

Servicio brindado desde el consorcio **ISTEC**, a través de su iniciativa **LibLink**, y mantenido por el **PREBI-UNLP** con el fin de apoyar la difusión de la producción científica de las instituciones que forman parte de esta comunidad y facilitar el acceso público y sin barreras a los recursos. La información expuesta en este portal, se recupera a partir de cosechas **OAI-PMH** de recursos publicados en Acceso Abierto en repositorios miembros del consorcio y de Latinoamérica, entre los que se encuentra **SEDICI**.

## Metabusador ISTEAC

Servicio de metabúsqueda proporcionado por **ISTEC** para facilitar y agilizar el acceso recursos bibliográficos disponibles en las bibliotecas del consorcio. Cada metabúsqueda dispara en simultáneo decenas de búsquedas distribuidas en los servidores **Z39.50** y **SRU** de las bibliotecas cooperantes, y a medida que recibe los resultados, los normaliza, deduplica y se los entrega al usuario.

## LA Referencia

Esta iniciativa busca compartir y dar visibilidad a la producción científica generada en las instituciones de educación superior y de investigación científica Latinoamericanas. Su principal propósito es la creación de una estrategia consensuada y un marco de acuerdos para la construcción y mantenimiento de una Red Federada de Repositorios Institucionales de Publicaciones Científicas destinada a almacenar, compartir, dar visibilidad y acceso abierto a su producción científica de América Latina.

## Scientific Commons

Es un proyecto de la Universidad de St. Gallen cuyo objetivo es desarrollar el mayor archivo de conocimiento científico con textos completos y disponibles para el público. Esta plataforma incluye una herramienta de búsqueda de publicaciones y perfiles de autor, entre otros servicios.

## Base-Search

**BASE** (del acrónimo inglés **Bielefeld Academic Search Engine**) es un motor de búsqueda multidisciplinario creado por la Biblioteca de la Universidad de Bielefeld, Alemania. Es usado generalmente para encontrar textos abiertos en recursos académicos a través de Internet. Los metadatos almacenados son el resultado de cosechas **OAI-PMH** en repositorios externos y transformaciones para mejorar su calidad.

## ARGENTINA

Software Open Journal System (OJS)

Software WordPress

Software Open Conference System (OCS)

OAI-PMH

OAI-PMH

## LATINOAMÉRICA

## MUNDO

## Interacciones

**SEDICI** utiliza **SWORD** para simplificar el depósito de nuevos contenidos de las revistas gestionadas por el Portal de Revistas de la **UNLP** y otros portales de distintas unidades académicas de la UNLP que utilizan el sistema **OJS**.

El servidor es provisto por **Dspace**, mientras que el cliente es una adaptación local del plugin **SWORD** para **OJS** disponible públicamente.

Los portales **OJS/OCS** pueden exponer sus metadatos en formato **DC Simple**. **SEDICI** toma estos datos, los mapea al formato interno y los utiliza como precarga.

## Mapeo / Crosswalk de metadatos

Es un proceso de transformación de datos utilizado para permitir el traspaso de información entre sistemas desacoplados. El objetivo es llevar la información de un sistema origen hacia un formato común, comprensible por otros sistemas externos.

En el ámbito de repositorios digitales, se utiliza como mecanismo principal para el intercambio de metadatos sobre protocolos estándares como **OAI-PMH** para exposición y cosecha, o **SWORD** para depósito delegado. El caso típico de transformación consiste en llevar/traer el formato interno de metadatos del repositorio a **Dublin Core** Simple o extendido



## Glosario

- DSPACE AIP** Formato de empaquetado de recursos provisto por **Dspace** para exportación/importación de recursos que incluyen toda la información descriptiva, administrativa y de preservación sobre un recurso y sus archivos. El término **AIP** proviene de la Norma **OAIS - ISO 14721:2012**.
- RSS ATOM** Son dos formatos similares para redifusión de contenidos en la web que permiten, entre otras cosas, que los usuarios se suscriban a un sitio y recuperen automáticamente las actualizaciones.
- SWORD** **Sword** es un protocolo de depósito semi-automático de recursos en repositorios construido sobre el protocolo de publicación **AtomPub**. Permite construir "paquetes de depósito" que posibilitan la interoperabilidad de dos sistemas a través de depósitos semi-automáticos y asíncronos en formatos de datos y metadatos comunes a ambos sistemas.
- OPEN SEARCH** Es un conjunto de tecnologías utilizado para acceder de manera simple a un servicio de búsqueda y recuperar los resultados devueltos. En su especificación, establece una serie de prácticas que permiten indicar cómo una aplicación cliente puede solicitar una búsqueda a un servidor y recuperar luego los resultados en uno de los formatos disponibles, como **RSS** y **ATOM**.
- OAI-PMH** Es un protocolo de recolección de metadatos definido por la **OAI** para intercambio de metadatos de forma independiente de las aplicaciones de borde. Hoy en día, es el estándar de facto utilizado para implementar agregación de recursos e interoperabilidad de repositorios digitales.
- Z39.50 SRU/SRW** **Z39.50** es un protocolo cliente-servidor dirigido a facilitar la búsqueda y recuperación de información en distintos sistemas a través de una misma interfaz. Su aplicación en el mundo de las bibliotecas y de los centros de documentación permite la consulta de recursos distribuidos en distintas bases de datos, desde un mismo punto de acceso. **SRU** y **SRW** son protocolos derivados de **Z39.50**, que funcionan sobre **HTTP**, utilizando **URLs (SRU)** o **webservices (SRW)**.