



UNIVERSIDAD  
NACIONAL  
DE LA PLATA

**SEDICI**

REPOSITORIO INSTITUCIONAL DE LA UNLP



# Introducción a DSpace

Aspectos Tecnológicos

**PREBI**  
prebi.unlp.edu.ar



Universidad Nacional  
de **Entre Ríos**



Esta obra está bajo una [Licencia Creative Commons Atribución-CompartirIgual 4.0 Internacional](#).

## Hoja de ruta

### Primera Parte

- Introducción a DSpace
- Arquitectura
- Modelo de contenidos
- Autorización y Autenticación
- Servicios para usuarios
- Interoperabilidad

### Segunda Parte

- Herramientas de automatización y Estructura de proyectos
- Configuración
- Vías de ingesta de ítems
- XMLUI
- Máquina virtual vagrant-dspace



# Introducción a DSpace



## Introducción a DSpace - Características generales

- Software libre y open-source
- Proyecto escrito mayoritariamente en lenguaje **JAVA**.

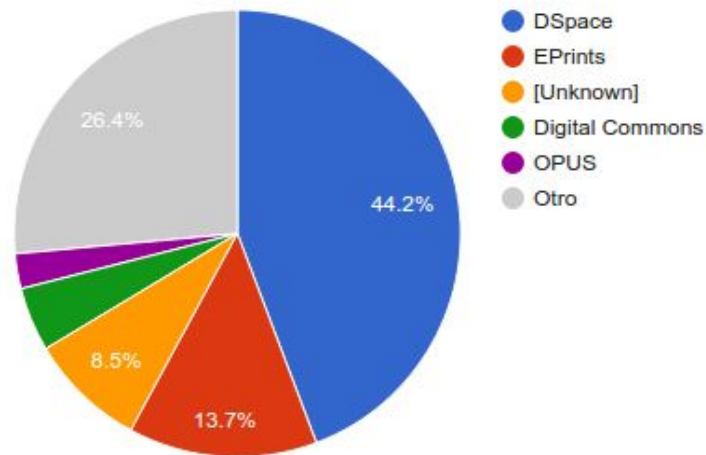


- Desarrollado y mantenido por una amplia **comunidad** de usuarios en todo el mundo
- Altamente **configurable**
- Puede ser fácilmente **extendido**
- Dispone de una gran cantidad de funcionalidades
- Posee un **modelo** de datos simple, con metadatos no jerárquicos e independencia de los formatos de archivos.



## Introducción a DSpace - Uso de Software de repositorios

DSpace	1414
EPrints	437
DigitalCommons	148
OPUS	81
CONTENTdm, dLibra, Fedora, Greenstone, Weko	50~60
Drupal, OJS, Mediawiki, OCS, OMP, Wordpress	<50



Total = 3196 repositories

OpenDOAR - 05-Sep-2016



## Introducción a DSpace - Código Abierto

- Licencia abierta tipo BSD
  - uso gratuito
  - actualizaciones regulares con modificaciones, bugfixes y mejoras
  - personalización
  - comunidad de desarrolladores
  - disponibilidad de mano de obra capacitada



## Introducción a DSpace - Stack

DSpace es una familia de aplicaciones java, algunas de las cuales funcionan en modo CLI (línea de comandos) mientras otras lo hacen como Aplicación Web JAVA.

- Para la ejecución **cli** se requiere una jre7 o jre8
- Para la ejecución de las **aplicaciones web** requiere un web container como: Jetty, Tomcat, entre otros.

En ambos casos se requiere una base de datos relacional para persistir el modelo. Se puede usar:

- postgres, oracle u otras soportadas por Hibernate.



## Introducción a DSpace - Stack

También se utiliza un indexador de texto para

- registros de acceso (estadísticas)
- índice de búsqueda (metadatos y fulltext)
- otros usos.

Solr:

- está optimizado para búsquedas: permite definir relevancia, filtrado, **pesos** en los campos, y más.
- permite gestionar grandes volúmenes de datos



Más información: Solr in DSpace <https://wiki.duraspace.org/display/DSPACE/Solr>



## Introducción a DSpace - Dependencias

**Spring es un** framework que implementa el modelo de **Dependency Injection** (también llamado IoC) sobre objetos POJO. Se compone de

- un bean container que administra el ciclo de vida de beans o POJO: `spring-core`.
- módulos específicos para
  - aplicaciones web mvc: [spring mvc](#)
  - autenticación y autorización: [spring-security](#)
  - entre otros.

DSpace utiliza Spring para administrar:

- Los **servicios** de la Capa de la Lógica de Negocios ([core-services.xml](#))
  - *MetadataValueService*, *ItemService*, *CollectionService*, etc.
- Las **extensiones** al core de DSpace ( configuraciones en [spring/api](#), mayormente)
  - *discovery.xml*, *workflow-actions.xml*, *bitstore.xml*, etc.



# Introducción a DSpace - Dependencias

## Ejemplo de archivo de beans de spring

Archivo `discovery.xml`. donde se declaran los beans que se utilizarán en el módulo.

## Clase `DiscoveryConfigurationService`

```
public class DiscoveryConfigurationService {  
  
    private Map<String, DiscoveryConfiguration> map;  
    private Map<Integer, List<String>> toIgnoreMetadataFields = new HashMap<>();  
  
    public Map<String, DiscoveryConfiguration> getMap() {  
        return map;  
    }  
  
    public void setMap(Map<String, DiscoveryConfiguration> map) {  
        this.map = map;  
    }  
  
    ///...|
```

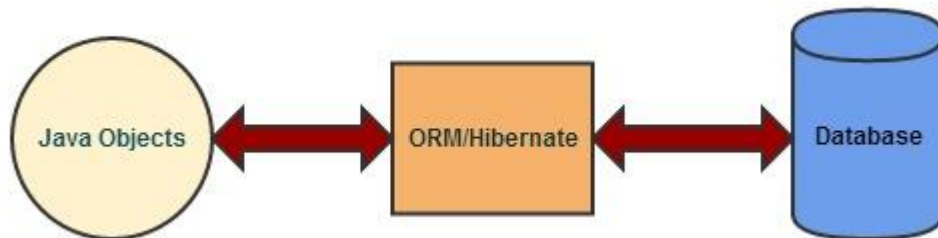
## Clase `DiscoveryConfiguration`

```
public class DiscoveryConfiguration implements InitializingBean{  
  
    /** The configuration for the sidebar facets */  
    private List<DiscoverySearchFilterFacet> sidebarFacets = new  
        ArrayList<DiscoverySearchFilterFacet>();  
  
    public List<DiscoverySearchFilterFacet> getSidebarFacets() {  
        return sidebarFacets;  
    }  
  
    @Required  
    public void setSidebarFacets(List<DiscoverySearchFilterFacet> sidebarFacets) {  
        this.sidebarFacets = sidebarFacets;  
    }  
  
    /// ...
```

```
<!--Bean that is used for mapping communities/collections to certain discovery  
configurations.-->  
<bean id="org.dspace.discovery.configuration.DiscoveryConfigurationService"  
    class="org.dspace.discovery.configuration.DiscoveryConfigurationService">  
    <property name="map">  
        <map>  
            <!--The default entry, DO NOT REMOVE the system requires this-->  
            <entry key="default" value-ref="defaultConfiguration" />  
        </map>  
    </property>  
    <!-- ... -->  
</bean>  
  
<!-- ... -->  
  
<!--The default configuration settings for discovery-->  
<bean id="defaultConfiguration"  
    class="org.dspace.discovery.configuration.DiscoveryConfiguration"  
    scope="prototype">  
    <!--Which sidebar facets are to be displayed-->  
    <property name="sidebarFacets">  
        <list>  
            <ref bean="searchFilterType" />  
            <!-- ... -->  
        </list>  
    </property>  
    <!-- ... -->  
</bean>
```

## Introducción a DSpace - Dependencias

**Hibernate** ORM es un framework para persistencia a partir del mapeo Objeto a Relaciones.



- permite **abstraer** a la aplicación de la BD utilizada
- **mapea automáticamente** los datos de objetos al momento de leer de/persistir en la BD
- permite usar consultas abstractas independientes de la BD usando lenguaje **HQL**
- puede utilizarse en combinación con otras librerías como *HibernateValidation*
- además provee:
  - caching de objetos,
  - transacciones de alto nivel
  - entre otros.

A partir de DSpace 6.x, se comenzó a utilizar **Hibernate** como parte de la capa de almacenamiento.



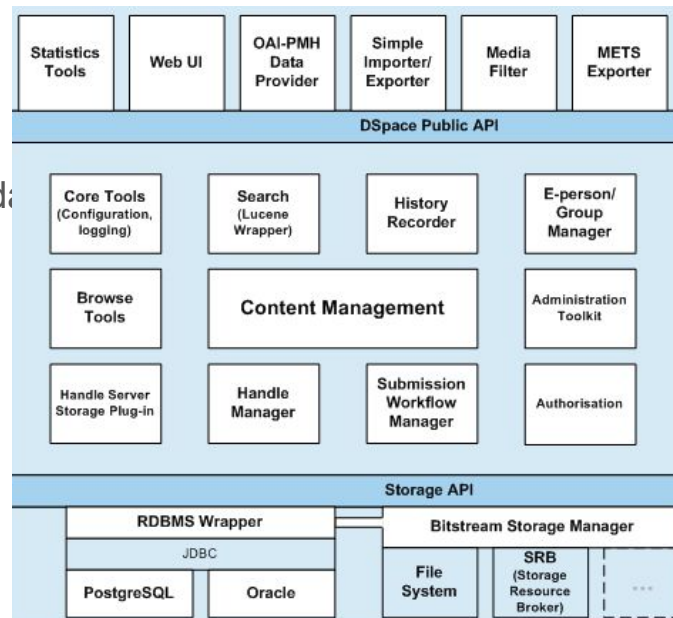
# ARQUITECTURA



# Arquitectura de DSpace - División Modular

DSpace se divide en 3 grandes grupos:

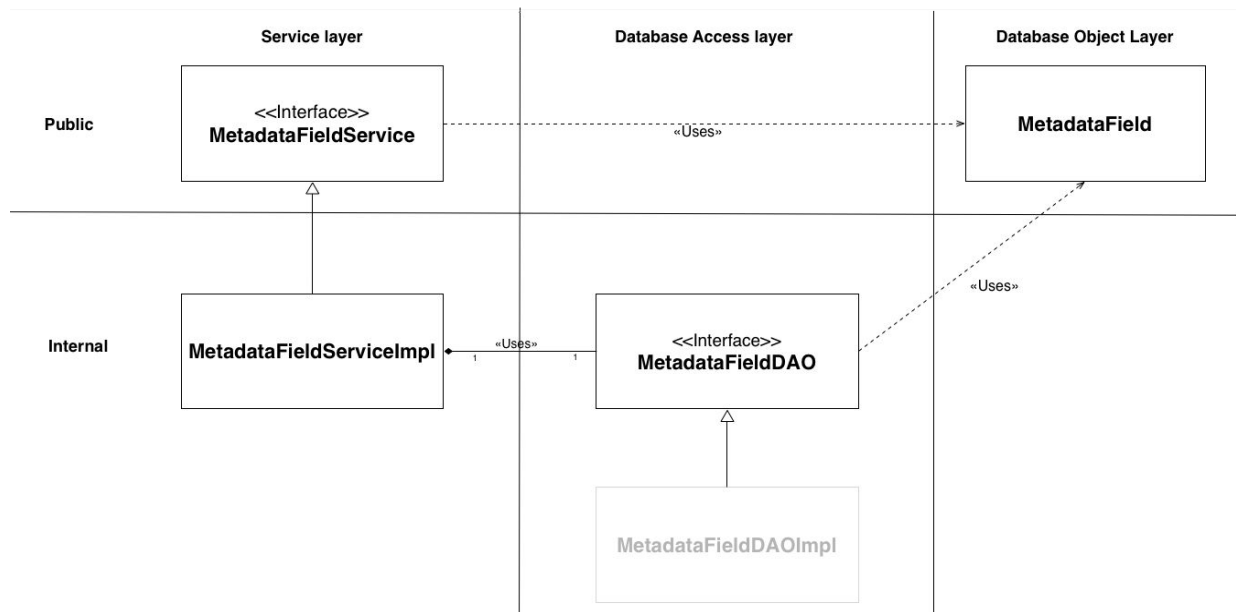
- **Módulos de Aplicación** (XMLUI, JSPUI, SWORD, OAI-PMH, REST-API, etc).
- **Módulos de Lógica de negocio** (DSpace API basada en *servicios*):
  - MetadataValueService
  - DSpaceObjectService
  - etc.
- **Módulos de Almacenamiento**
  - Interfaces DAO, Objetos de base de datos
  - Base de datos específica
  - Almacén de Bitstreams



## Arquitectura - Api basada en Servicios

La API se ha dividido en 3 capas:

1. de Servicios
2. de Acceso a Datos
3. de Objetos Persistentes



Ejemplo con Objeto MetadataField



# Arquitectura - Api basada en Servicios - Capa de Servicios

## Capa de Servicios:

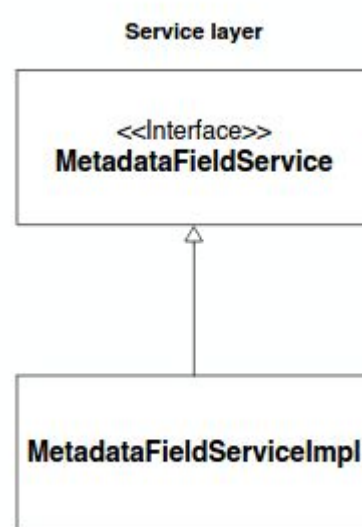
La capa de servicios define bloques con la lógica de negocios, es totalmente pública y se divide en 2:

### 1. Servicios basados en la base de datos:

- Servicios que darán acceso a operaciones CRUD sobre los objetos en base de datos y a operaciones propias de la lógica de cada objeto.

### 2. Servicios de la capa de negocios:

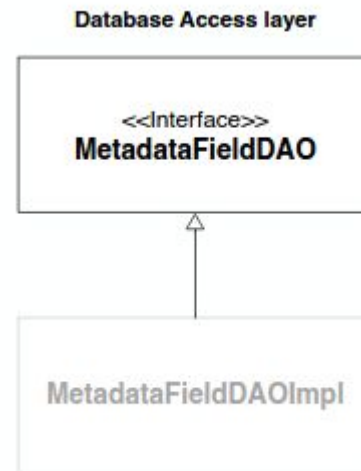
- Servicios que implementan lógica de negocios de alto nivel  
Ejemplo: AuthorizationManager



## Arquitectura - Api basada en Servicios - Capa Acceso a Datos

### Capa de Acceso a Datos (DAOs):

- Accesible sólo a través de la capa de servicios.
- Interactúa con la base de datos
- Contiene lógica de negocios para manipular objetos persistentes.
- Hibernate es elegido como implementación de la JPA pero puede ser fácilmente reemplazado.



## Arquitectura - Api basada en Servicios - Capa de Objetos Persistentes

### Capa de Objetos Persistentes:

- Cada objeto en esta capa se corresponde con una tabla en la base de datos.
- No hay consultas a la base de datos
- Solo getters, setters y anotaciones JPA para asociar variables con tablas y columnas
- No hay lógica de negocios

Database Object Layer

MetadataField



## Arquitectura - Plugins

DSpace dispone de varios plugins (extensiones opcionales) para extender su funcionalidad, entre ellos:

- Plugins para la autenticación
- Plugins para el manejo de autoridades
- Plugins para procesamiento de los bitstreams (*media filters*)



# Arquitectura - Plugins - Autenticación

Dspace cuenta con varios *mecanismos para la autenticación* de usuarios en el sistema:

- LDAP
- IP
- Shibboleth
- x509
- Tradicional

```
#-----#
#-----PASSWORD AUTHENTICATION CONFIGURATIONS-----#
#-----#
# Configuration properties used by the Password #
# Authentication plugin, when it is enabled. #
#-----#
#
# Only emails ending in the following domains are allowed to self-register
# Example - example.com domain : @example.com
# Example - MIT domain and all .ac.uk domains: @mit.edu, .ac.uk
# authentication-password.domain.valid = example.com

##### Password users group #####

# If required, a group name can be given here, and all users who log in
# using the DSpace password system will automatically become members of
# this group. This is useful if you want a group made up of all internal
# authenticated users.
# authentication-password.login.specialgroup = group-name

##### Password hashing algorithm #####

# You may select any digest algorithm available through
# java.security.MessageDigest on your system. At least MD2, MD5, SHA-1,
# SHA-256, SHA-384, and SHA-512 should be available, but you may have
# installed others. If not set, SHA-512 will be used.
# authentication-password.digestAlgorithm = SHA-512
```

http://digital.cic.gba.gob.ar/password-login

## Acceder a DSpace

**Correo electrónico:**

**Contraseña:**  
 [¿Olvidó su contraseña?](#)

## Registrar un nuevo usuario

Registre una cuenta para suscribirse a las colecciones, para recibir n  
Pulse aquí para registrarse.



# Arquitectura - Plugins - Autoridades

DSpace cuenta con un mecanismo para *manejo de autoridades*, que permite la selección de valores autorizados durante la carga de metadatos al repositorio.

- Implementan la interfaz [ChoiceAuthority](#).
- Autoridades predefinidas en DSpace (*DCInputAuthority*, *SHERPARoMEOPublisher*, etc.)
- Capacidad de agregar nuestras **propias autoridades**.

```
public abstract class SPARQLAuthorityProvider implements ChoiceAuthority {  
  
    public Choices getMatches(String field, String text, Collection collection,  
        int start, int limit, String locale) {  
        if (text == null)  
            text = "";  
        else  
            text = text.replace("\n", "");  
  
        ParameterizedSparqlString query = this.getSparqlSearchByTextQuery  
            field, text, locale);  
        Choice[] choices = this.evalSparql(query, start, limit);  
        log.trace(choices.length + "matches found for text " + text);  
        return new Choices(choices, start, limit, Choices.CF_ACCEPTED, fa  
            // TODO hasMore??  
    }  
  
    public final Choices getBestMatch(String field, String text,  
        Collection collection, String locale) {  
        return this.getMatches(field, text, collection, 0, 1, locale);  
    }  
  
    public String getLabel(String field, String key, String locale) {  
  
        ParameterizedSparqlString query = this.getSparqlSearchByIdQuery(field,  
            key, locale);  
        Choice[] choices = this.evalSparql(query, 0, 1);  
        if (choices.length == 0)  
            return null;  
        else  
            return choices[0].label;  
    }  
}
```

**Autor Institucional:**  
Institución creadora del contenido intelectual de la obra

uni

Universidad de Salamanca (USAL)  
AGH University of Science and Technology (AGH)  
Benemérita Universidad Autónoma de Puebla (BUAP)  
Cadi Ayyad University (UCA)   
Centro de Investigación en Economía Política y Comunicación (CIEPYC)  
Editorial de la Universidad Nacional de La Plata (EduLP)  
Facultad de Periodismo y Comunicación Social  
Hanoi University of Mining and Geology (HUMGS)  
Ibn zohr University (UIZ)  
Kyoto University (KYOTOU)



# Arquitectura - Plugins - Media Filters

DSpace cuenta con varias extensiones que permiten procesar los objetos digitales y generar nuevos contenidos para mejorar los servicios ofrecidos (*búsquedas, vistas previas, etc*).

Algunos media-filters:

## EXTRACTORES DE TEXTO

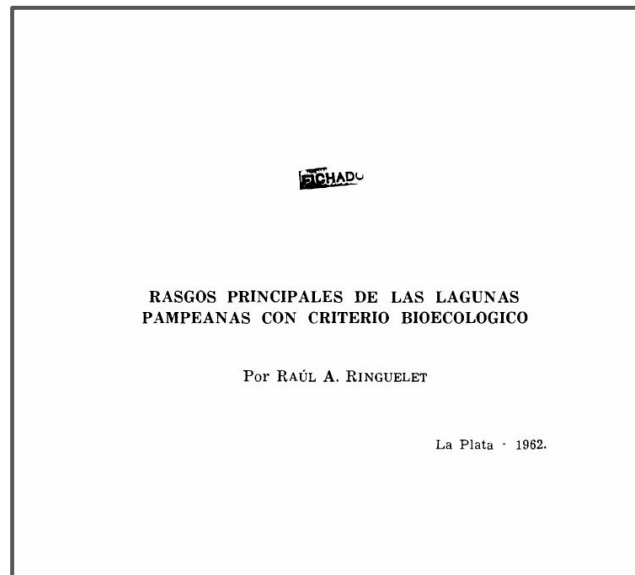
- PDF Text Extractor
- HTML Text Extractor
- Word Text Extractor
- Excel Text Extractor
- PowerPoint Text Extractor

## EXTRACTORES DE IMÁGENES

- JPEG Thumbnail
- ImageMagick PDF Thumbnail

## TEXTO EXTRAÍDO

RASGOS PRINCIPALES DE LAS LAGUNAS  
PAMPEANAS CON CRITERIO BIOECOLOGICO  
Por RAÚL A. RINGUELET  
La Plata - 1962.  
R E S U M E N  
El cuerpo de agua léntico, más común en la llanura chaco-pampeana de la Argentina, se denomina regionalmente "laguna", el cual se estudia en la presente contribución desde un punto de vista integral y complejo, fundamentalmente ecológico. El autor ofrece un panorama sintético de los conocimientos científicos sobre este tipo de ambiente continental acuático, definiéndolo como categoría; detalla los tipos lagunares según una clasificación original con criterio genético, pero con fundamento limnológico, y pasa revista a los rasgos esenciales: geomorfológicos, sedimentológicos, térmicos, químicos, factores bióticos especialmente vida planctónica, productividad, fenómenos rítmicos de varias amplitudes, trofismo y sus tipos, y sucesión.  
S U M M A R Y  
The most common type of inland water in the chaco-pampean plain of Argentina is a lenitic body called regionally "laguna". Correspond to a third order or polymyctic lake by the thermic characteristics: the continuous circulation and the absence of thermal stratification. The bassin is commonly



## THUMBNAIL



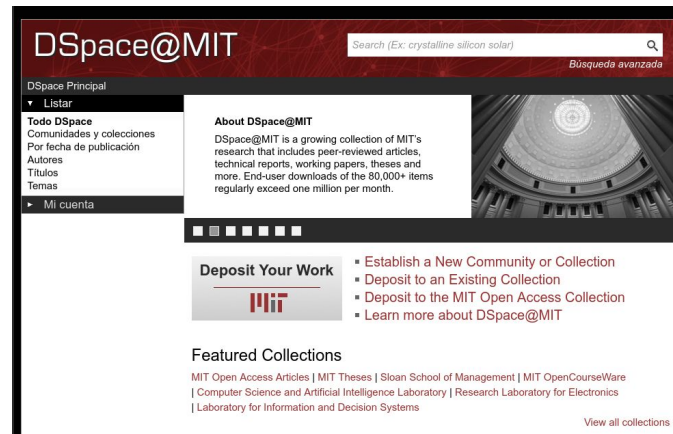
## Arquitectura - Aplicaciones

- **XMLUI**
  - Frontend web basada en framework Manakin
- **JSPUI**
  - Frontend web basada en Java Servlets y Java Server Pages (JSP)
- **REST-API**
  - Interfaz REST
- **OAI2**
  - Data Provider OAI-PMH
- **DSpace-RDF**
  - SPARQL endpoint
- **Sword**
  - servidor sword (v1 y v2)



## Arquitectura - XMLUI

- Es el frontend de DSpace más utilizado.
- Implementa una arquitectura
  - modular en torno a un pipeline
  - extensible
  - multi-tier
    - Capa de estilos
    - Capa de temas
    - Capa de aspectos



# Arquitectura - JSPUI

- Aplicación tradicional MVC orientada a acciones
  - Java Servlets para gestionar la lógica de negocios, y
  - Java Server Pages (JSP) para producir el código HTML.

The screenshot displays the DSpace web interface with several key components highlighted:

- (Header):** The top section containing the DSpace logo and navigation links.
- (Location Bar):** A box containing the text "DSpace at My University >".
- (Navigation Bar):** A vertical sidebar on the left with links for Home, Communities & Collections, Titles, Authors, Subjects, and By Date.
- Search Area:** A central section with a search box and a "Go" button.
- (Browse Layout):** A table listing items with columns for Issue Date, Title, and Author(s).

Issue Date	Title	Author(s)
1996	The 1995 review and extension of the Nuclear Non-Proliferation Treaty	edited by Thomas Morgan, Jason Zych and Clifford E. Singer
1998	Aadhi raat ke baad : "After midnight"	S. Rashid Alam
2005	Anatomy of interlinking rivers in India : a decision in doubt	A.C. Shukla and Vandana Asthana, Asthana, Vandana
Feb-2004	Asymptotic theory of ignition and failure of self-sustained detonations	-
Jun-2001	Atomistic scale experimental observations and micromechanical/continuum models for the effect of hydrogen on the mechanical behavior of metals	-



## Arquitectura - Aplicaciones - REST

# RESTful API

GET PUT POST DELETE

### → Simple

- ◆ Basado sobre HTTP
- ◆ típicamente usa métodos HTTP: GET (retornar), POST (crear), PUT (actualizar) y DELETE (eliminar).
- ◆ Evita las abstracciones adicionales de otros protocolos de intercambio de mensajes (ej. SOAP)

### → Permite:

- ◆ obtener datos
- ◆ ejecutar operaciones sobre los datos
- ◆ Intercambiar datos en formatos procesables por máquina: XML, JSON, etc

### → Se pueden implementar aplicaciones complejas sobre cualquier lenguaje

- ◆ cliente server: php, java, ruby, javascript, etc.
- ◆ cliente userAgent: javascript



## Arquitectura - Aplicaciones - REST

- Disponible desde DSpace 4
- API funcionalmente completa
  - ◆ acceso a DSO:
    - Comunidades, Colecciones, Ítems y Bitstreams
  - ◆ gestión de DSO
  - ◆ autenticación
- Soporta XML y JSON
  - ◆ curl -s -H "Accept: application/xml"  
<http://localhost:8080/rest/communities>
  - ◆ curl -s -H "Accept: application/json"  
<http://localhost:8080/rest/communities>

Specific Bitstream	/bitstreams/:bitstreamID
Download a Bitstream	/bitstreams/:bitstreamID/retrieve
Bitstream Expands	parent, all

List Communities	/communities/
Specific Community	/communities/:communityID
Community Expands	parentCommunity, collections, subCommunities, logo, all

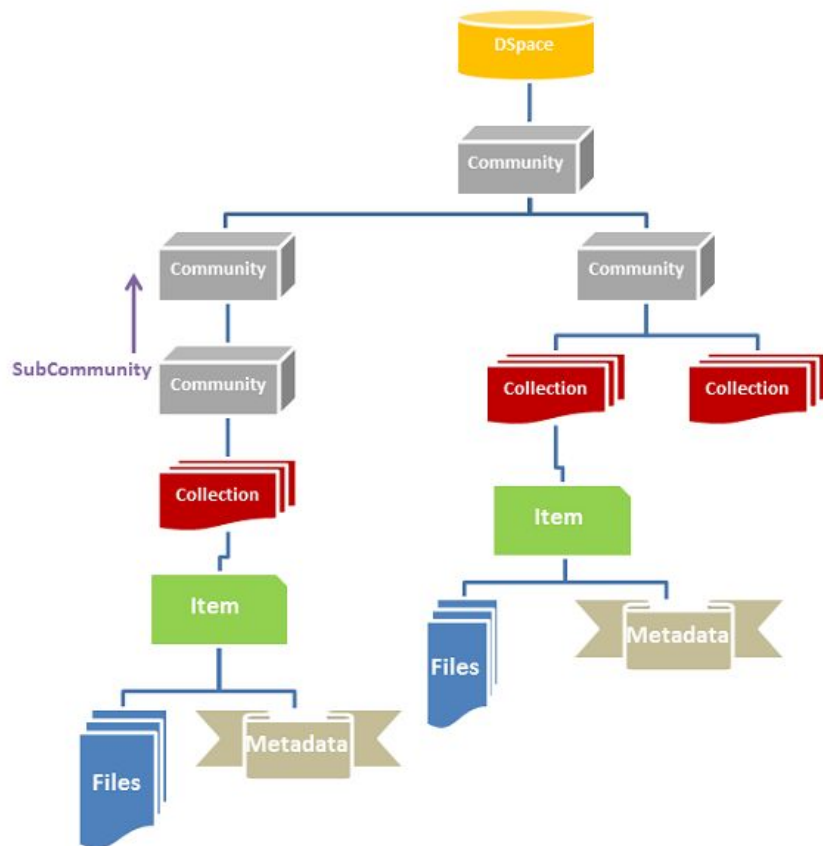
Más doc DSpace REST API <https://wiki.duraspace.org/display/DSDOC5x/REST+API>



# Modelo de contenidos



## Modelo de contenidos - Estructura típica



1. El repositorio se organiza en una o más **comunidades** de nivel base que se organizan jerárquicamente en subcomunidades.
  - Son como espacios de trabajo
2. Las **colecciones** son los “estantes” dentro de las comunidades, que agrupan contenido relacionado.
3. Los ítems son las obras que van en los estantes y que se pretende que el público encuentre.
4. Los metadatos describen al recurso
5. Los bitstreams son la representación digital del recurso.




# Modelo de contenidos - Comunidad y Colecciones



Repositorio Institucional  
Comisión de Investigaciones Científicas

[Inicio](#) [Explorar](#) [Aportar Material](#) [Mas información](#) [Contacto](#)

 [Mi cuenta](#) [ES](#)

## Comunidades en DSpace

Elija una comunidad para listar sus colecciones

### ▼ Centros [2053]

Centros de la Comisión de Investigaciones Científicas

#### ▶ CIDCA [1]

Centro de Investigación y Desarrollo en Criotecnología de Alimentos

#### ▶ CESGI [1]

Centro de Servicios en Gestión de Información

### ▼ LEMIT [806]

Laboratorio de Entrenamiento Multidisciplinario para la Investigación Tecnológica

Tesis [10]

Artículos y presentaciones en Congresos [20]

Libros [2]

#### ▼ Revista Ciencia y Tecnología de los Materiales [32]

Número 01 [7]

Número 02 [7]

Número 03 [6]

Número 04 [6]

Número 05 [6]

#### ▶ Informes de Investigación [34]

▶ Jornadas de Técnicas de Reparación y Conservación del Patrimonio [419]

▶ Revista Ciencia y Tecnología del Hormigón [46]

▶ Anales LEMIT. Serie II [237]

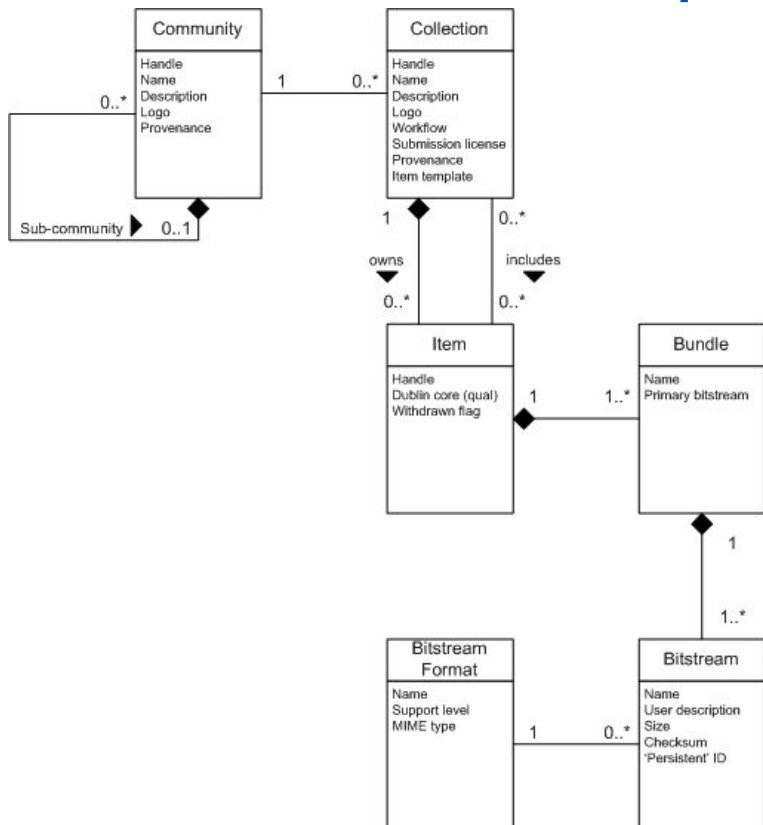
▶ Anales LEMIT. Serie III [6]

#### ▶ CIOp [198]

Centro de Investigaciones Ópticas



## Modelo de contenidos - DSpace Data Model



- Las comunidades poseen subcomunidades y colecciones
- Las colecciones poseen ítems
- Los ítems se componen de bundles
- Cada bundle posee bitstreams (archivos)



## Modelo de contenidos - Bundles

Los ítems están vinculados a sus bitstreams a través de entidades llamadas Bundles. Los **bundles** agrupan bitstreams según ciertas categorías:

- **ORIGINAL**: contiene los bitstreams a publicar.
- **THUMBNAILS**: son archivos con thumbnails extraídos a partir de los bitstreams originales (p.e. thumbnails de PDFs, imágenes, etc.).
- **TEXT**: contiene el texto completo (full-text) de otros bitstreams. Se genera a partir de la extracción automática de texto sobre otros bitstreams y se usa durante la indexación para mejorar los resultados de búsqueda.
- **LICENSE**: contiene la licencia que el usuario aceptó al depositar el contenido.
- **CC\_LICENSE**: contiene la [Licencia CC](#) (si es que existe) seleccionada por el usuario durante la carga.



# Modelo de contenidos - Bitstreams

Los ítems del repositorio se componen de **objetos digitales** que representan la obra en sí que se quiere publicar. Los objetos digitales pueden ser:

- Audios
- PDFs
- Documentos de texto (.doc, .odt., etc)
- Diapositivas de presentaciones
- Planillas de cálculo
- Imágenes
- Videos
- etc.

En el lenguaje de DSpace, estos objetos digitales reciben el nombre de **BITSTREAMS**.



The screenshot shows the CIEDIGITAL website interface. At the top, there is a navigation menu with options: Inicio, Explorar, Aportar Material, Mas información, and Contacto. On the right, there are links for Mi cuenta, Estadísticas, and ES. The main content area displays a document titled "El Falcon Verde como lugar simbólico y la posibilidad de su activación para un uso turístico responsable" by "Sahores Avalis, Virginia". The document is categorized under "Año 2013". A "Resumen:" section is visible, starting with "La última dictadura que sufrió la Argentina, desde 1976 hasta 1983, dejó en la memoria una serie de lugares simbólicos. Entre estos símbolos de la represión que siguen vigentes se encuentran diversos Centros Clandestinos de Detención (CCD) y otros, más relacionados con los espacios cotidianos, como el Falcon Verde. Dicho vehículo era utilizado por las Fuerzas Armadas para el cumplimiento del plan de Reorganización Nacional. A partir de la necesidad de conservar este lugar simbólico, es que se busca demostrar que el Falcon, como objeto de uso turístico, puede adquirir un reconocimiento y valoración como símbolo identitario y de esa manera sortear el obstáculo que presenta la discontinuidad de su fabricación. Establecer que el uso turístico puede colaborar con la conservación y sostenibilidad en el tiempo del patrimonio cultural. El Turismo, como actividad cultural y social, puede colaborar con la conservación de los lugares simbólicos y evitar la mercantilización de estos, otorgando la posibilidad de pensar en construcciones de sentidos y resignificaciones. En este sentido, se pone en evidencia que el turismo puede no ser sólo una actividad económica sino también una expresión cultural y una práctica social que ayude a la conservación del patrimonio. Se intenta mostrar la conexión directa existente entre el Patrimonio y el turismo, más específicamente el Patrimonio Cultural y el Turismo Cultural. Visto este último como una actividad social que involucra tanto a la comunidad receptora como a los visitantes. Se busca ahondar en la significación del Falcon Verde, en tanto símbolo de la historia reciente de la República Argentina, a través de fuentes y testimonios. Así como la posibilidad de su activación para el uso turístico responsable."

On the right side of the document page, there is a "Descargas" section with a red-bordered box containing a thumbnail of the document and a download link: "Documento completo Archivo PDF (5.525Mb)".



## Modelo de contenidos - BitStore

Los bitstreams son alojados en un **Bitstore**, que puede estar alojado de forma local o en la nube (según la configuración en [dspace/config/spring/api/bitstore.xml](#)):

- **ASSETSTORE:** si tenemos esta forma de almacenamiento, los bitstreams se alojan en un directorio local en el servidor:

```
{dspace-install-dir}/assetstore/
```

- **S3Store:** si tenemos esta forma de almacenamiento, los bitstreams se alojarán en la nube, más específicamente en el Cloud Storage de [Amazon S3](#). Para que esto funcione debemos configurar diferentes parámetros como, p.e, una *accessKey* al sistema de Amazon.



## Modelo de contenidos - Metadatos

dc.format.extent	116 p.	es
dc.language	es	es
dc.title	Un metalenguaje de programación orientado al diseño de interfaces gráficas	es
dc.type	Tesis	es
sedici.creator.person	Santana, Carlos Ariel	es
sedici.creator.person	Coni, Cintia Vanesa	es
sedici.subject.materias	Ciencias Informáticas	es
sedici.subject.keyword	arquitectura dirigida por modelos	es
sedici.subject.keyword	transformaciones de modelos	es
sedici.subject.keyword	interfaz de usuario	es
sedici.description.fulltext	true	es
mods.originInfo.place	Facultad de Informática	es
sedici.subtype	Tesis de grado	es
sedici.rights.license	Creative Commons Attribution 4.0 International (CC BY 4.0)	
sedici.rights.uri	<a href="http://creativecommons.org/licenses/by/4.0/">http://creativecommons.org/licenses/by/4.0/</a>	
sedici.contributor.director	Pons, Claudia	es
thesis.degree.name	Licenciado en Sistemas	es
thesis.degree.grantor	Facultad de Informática	es

### Los metadatos

- son datos estructurados que describen otros datos
- son datos sobre datos

← Ej: representación del recurso

<http://sedici.unlp.edu.ar/handle/10915/47008>



## Representación - Metadatos



### Esquemas:

- **Dublin Core Element Set (DCES)** es uno de los esquema de metadatos estándar más utilizados <http://dublincore.org/documents/dces/>
  - **Título** → dc.title
- **DCMI Terms (dc extrendido o calificado)** es una extensión al esquema básico de DC <http://dublincore.org/documents/dcmi-terms/>
  - **Título alternativo** → dcterms.alternative
- Un repositorio puede utilizar un *perfil de metadatos* basado en esquema propio, estándar y/o mixto
  - Ej: subconjunto de DC, uso de *application profiles* (esquema de metadatos + restricciones de uso) específicos.



# Representación - Metadatos

Ejemplo: SEDICI utiliza un perfil de metadatos propio basado en:

→ **DC simple**

- ◆ dc.title

→ **DC calificado**

- ◆ dc.date.accessioned

→ **ETD**

- ◆ thesis.degree.name

→ **MODS**

- ◆ mods.location

→ **Esquema de metadatos SEDICI**

- ◆ sedici.subtype

dc.date.issued	2009	
dc.identifier.uri	<a href="http://hdl.handle.net/10915/1102">http://hdl.handle.net/10915/1102</a>	
dc.description.abstract	Tomo I: Memoria descriptiva; Informe medioambiental; Trazado; Proyecto de arquitectura la Estación Plaza Moreno. Tomo II: Proyecto estructural de la Estación Plaza Moreno. Tomo III: Plan de trabajos; Evaluación económica	es
dc.format.extent	3 vol. + planos	es
dc.language	es	es
dc.title	Línea "A" de subte - La Plata	es
dc.type	Tesis	es
sedici.creator.person	Liaudat, Joaquin	es
mods.location	<a href="http://biblio.ing.unlp.edu.ar/cgi-bin/koha/opac-detail.pl?bib=INGC-MON-017361">http://biblio.ing.unlp.edu.ar/cgi-bin/koha/opac-detail.pl?bib=INGC-MON-017361</a>	es
sedici.subject.materias	Ingeniería	es
sedici.subject.materias	Ingeniería Civil	es
sedici.subject.eurovoc	transporte subterráneo	es
sedici.subject.eurovoc	medio de transporte	es
sedici.description.fulltext	false	es
mods.originInfo.place	Departamento de Construcciones	es
sedici.subtype	Tesis de grado	es
sedici.contributor.director	Venier, Leonardo P.	es
thesis.degree.name	Ingeniero Civil	es
thesis.degree.grantor	Facultad de Ingeniería	es

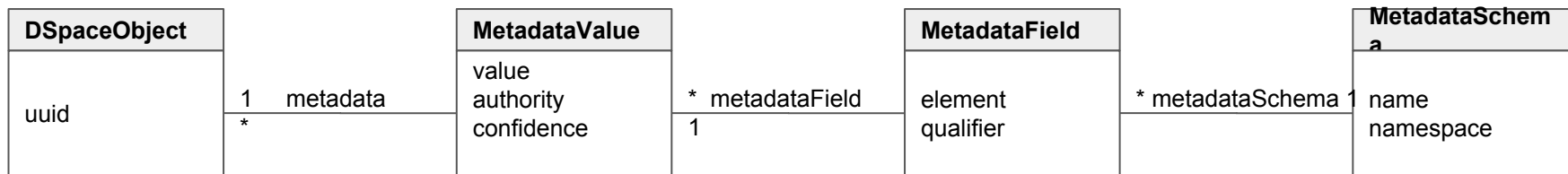


## Modelo de contenidos - Metadatos

DSpace permite definir **perfiles de metadatos** a partir de

1. combinación de elementos de diferentes schemas
2. usa de calificadores:
  - schema.element.qualifier

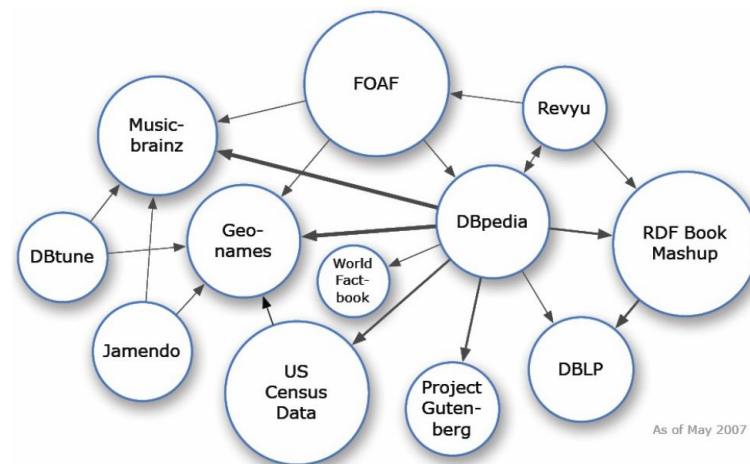
DC Simple (estándar DCES)	DC simple (perfil en DSpace)
dc.date	dc.date.accessioned dc.date.issued dc.date.created dc.date.available dc.date.submitted



## Modelo de contenidos - Autoridades

Uso de Vocabularios controlados gestionados en otros sistemas

- Tesoros
- Sistemas de clasificación / Taxonomías
- Lista de encabezamientos de materias
- Bases de datos de investigadores
- Jerarquía de instituciones
- Grados alcanzados



El software del repositorio debe integrarse con estos sistemas.

En DSpace estos vocabularios controlados son llamados **autoridades**.



## Modelo de contenidos - Autoridades

- **Autoridad:** conjunto de entidades normalizadas
  - **Registro de autoridad:** información asociada con una entidad
  - **Clave de autoridad:** código que identifica unívocamente una entidad.



## Modelo de contenidos - Autoridades

### Ventajas:

1. Permite probar que dos valores son idénticos comparando por la clave de las autoridades.
2. Ayuda a completar metadatos con valores correctos.
3. Permite mejorar la calidad de los metadatos.
4. Mejora la interoperabilidad compartiendo un nombre de autoridades con otra aplicación.
5. Reduce el tiempo de carga



## Modelo de contenidos - Autoridades

DSpace incluye una API para el uso de autoridades en los metadatos.

- la fuente puede ser externa o interna al repositorio
- para integrar nuevas autoridades deben usarse las clases e interfaces:
  - **Choice:**
    - Clase que contiene los atributos authority, label, confidence y value.
  - **Choices:**
    - Clase que contiene un conjunto de Choice.
  - **ChoiceAuthority:**
    - Interfaz para suplir el mecanismo del control de autoridades.



## Modelo de contenidos - Autoridades

La elección del valor de un metadato pueden ser abiertas o cerradas.

- Abiertas :
  - Se pueden usar valores no incluidos en las elecciones presentadas.
- Cerradas:
  - Elección restringida solo para un conjunto de valores ofrecidos.

La elección del valor de un metadato pueden ser obligatorio u opcional.



## Modelo de contenidos - Autoridades

### Choice management:

Mecanismo para selección o elección de los posibles valores de un metadato, a partir de valores propuestos o consultas.

Ejemplo: Submission

#### Autor Institucional:

Institución creadora del contenido intelectual de la obra

- Universidad de Salamanca (USAL)
- AGH University of Science and Technology (AGH )
- Benemérita Universidad Autónoma de Puebla (BUAP)
- Cadi Ayyad University (UCA)
- Centro de Investigación en Economía Política y Comunicación (CIEPYC)
- Editorial de la Universidad Nacional de La Plata (Edulp)
- Facultad de Periodismo y Comunicación Social
- Hanoi University of Mining and Geology (HUMG)
- Ibn zohr University (UIZ)
- Kyoto University (KYOTOU)

Add

Add



## Modelo de contenidos - Autoridades

### Indexación de autoridades

DSpace permite indexar las autoridades para crear facets de búsqueda que permiten al usuario filtrar resultados

Materia
Ciencias Físicas (84)
Óptica, Acústica (40)
Física (23)
Ingeniería de los Materiales (12)
Ciencias Químicas (9)
Diseño Arquitectónico (8)
Ingeniería de los Materiales (7)
Física Atómica, Molecular y Química (3)
Ingeniería Eléctrica y Electrónica (3)
Físico-Química, Ciencia de los Polímeros, Electroquímica (2)
... ver más



## Modelo de contenidos - Autoridades

### Control de Autoridades en DSpace - Base de datos

1. No es un reemplazo del valor de los metadatos.
2. Es configurado mediante un campo en la base de datos.

text_value	authority	Confidence
Universidad Nacional de La Plata (UNLP)	<a href="http://digital.cic.gba.gob.ar/auth/node/86555">http://digital.cic.gba.gob.ar/auth/node/86555</a>	600
Cañueto, Matías F.	<a href="http://digital.cic.gba.gob.ar/auth/node/204702">http://digital.cic.gba.gob.ar/auth/node/204702</a>	600
Attribution 4.0 International (BY 4.0)	<a href="http://creativecommons.org/licenses/by/4.0/">http://creativecommons.org/licenses/by/4.0/</a>	600
UNLP	<a href="http://digital.cic.gba.gob.ar/auth/node/86555">http://digital.cic.gba.gob.ar/auth/node/86555</a>	500
Cañueto, Matías G.		-1



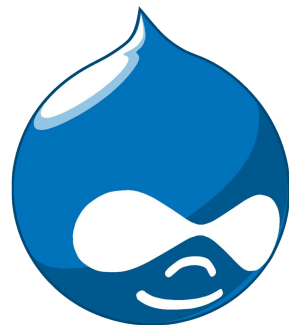
## Modelo de contenidos - Gestión de Autoridades

Las autoridades controladas utilizadas en **CIC-DIGITAL** se encuentra en una instalación de Drupal.

Drupal:

- Cuenta con un conjunto de **módulos** que facilitan el modelado, la importación de los contenidos con sus relaciones y su indexación, gracias a su soporte de nodos y taxonomías.
- La comunicación entre esta plataforma y CIC-DIGITAL es mediante consultas **SPARQL**, a través de un endpoint configurado en la instalación de Drupal el cual permite realizar las búsquedas.

**SEDICI** utiliza un software propio de gestión de autoridades



**Drupal**™



sedici.unlp.edu.ar

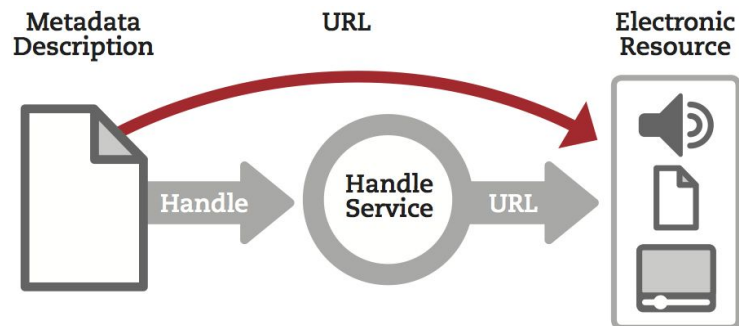
## Modelo de contenidos - Identificadores persistentes

Permiten identificar un recurso de manera independiente a la dirección del repositorio o sitio donde dicho recurso se encuentra.

[sedici.unlp.edu.ar/handle/10915/54681](http://sedici.unlp.edu.ar/handle/10915/54681) ⇒ <http://hdl.handle.net/10915/54681>

[revistas.ucr.ac.cr/index.php/eciencias/article/view/23690](http://revistas.ucr.ac.cr/index.php/eciencias/article/view/23690) ⇒ <http://dx.doi.org/10.15517/eci.v6i2.23690>

<http://revistas.unlp.edu.ar/raab/article/view/2177> ⇒ <http://dx.doi.org/10.17139/raab.2016.0018.02.09>



Fuente de la imagen <https://www.clarin.eu/content/persistent-identifiers>



## Modelo de contenidos - Identificadores persistentes

Por qué son necesarios?

- Las URLs pueden cambiar
  - en el dominio
  - en la ruta
- por
  - Cambios en el software
  - Cambios en políticas institucionales

**10045/13546**  
↑ prefijo ↑ sufijo ↑

**Handle**

**10.4100/jhse.2010.52.15**  
↑ prefijo ↑ sufijo ↑  
**DOI**

http://purl.org/net/example/myFirstPURL  
Scheme host domain PURL name  
PURL id

**PURL**

Los cambios de URLs deben notificarse al **manejador de redirecciones**, para que éste actualice sus reglas de redireccionamiento



## Modelo de contenidos - Identificadores persistentes

Los recursos pueden tener más de un identificador persistente:

- al momento de la publicación (ej. artículo de revista)
- al momento del depósito en el repositorio institucional

### Información general

Fecha de publicación: diciembre 2016

Idioma del documento: Español

Revista: Revista Argentina de Antropología Biológica; vol. 18, no. 2

Dossier: Los estudios de histología ósea en Antropología Biológica

Institución de origen: Asociación de Antropología Biológica de la República Argentina (AABRA)

DOI: 10.17139/raab.2016.0018.02.01

ISSN: 1853-6387

Referencias geográficas: Argentina

Extensión: 3 p.

Materias: Antropología

Descriptor: Técnicas Histológicas; paleohistología; tejido óseo

### Descargar archivos



Documento completo  
Descargar archivo (428.7Kb) - PDF



Enlace externo  
[www.revistas.unlp.edu.ar/...](http://www.revistas.unlp.edu.ar/)

Creado el: 3 de junio de 2016

Disponible en SEDICI desde: 3 de junio de 2016

Por favor, utilice este identificador (URI) para citar o enlazar este ítem: <http://hdl.handle.net/10915/53190>

## Modelo de contenidos - Identificadores persistentes - Usos

Los identificadores persistentes pueden aplicarse:

- documentos de texto: artículos, tesis, libros
- datos de investigación: bases de datos, imágenes, audios
- personas

Dónde se usan?

- Catalogación y publicación de recursos
- Citas bibliográficas
- Menciones en línea (blogs, redes sociales, etc.)



The screenshot shows a Scopus profile for Gonzalo L. Villarreal. At the top, the name "Gonzalo L. Villarreal" is displayed in green. Below it is a green bar with the text "ORCID ID" and a link to "orcid.org/0000-0002-3602-8211". A button labeled "Ver versión pública" is visible. The word "Scopus" is written in orange below the ORCID information. Below this is a dark teal navigation bar with "Search" and "Alerts" buttons. Underneath, there is a "Back to results" link and a "1 of 1" indicator. The main profile information includes the name "Villarreal, Gonzalo Luján", affiliation "Universidad Nacional de La Plata, Servicio de Difusión de la Creación Intelectual, La Plata, Argentina", and author ID "25628901000". At the bottom, there is a link to the ORCID profile: "http://orcid.org/0000-0002-3602-8211".



## Modelo de contenidos - Identificadores persistentes en DSpace

Dspace permite utilizar identificadores

- internos: uuid auto-generados
- handle: servidor local
- doi: registro externo
- custom: según se implemente

Se utilizará un servidor handle.net local al servidor de DSpace con un prefijo único propio de UNER.



# Autorización y Autenticación



## Autenticación

A partir de diversos métodos configurables:

- por password
- por ip o rango de ips
- por LDAP
- por Shibboleth
- X509
- custom

Más info <https://wiki.duraspace.org/display/DSDOC6x/Authentication+Plugins>



## Autorización - Modelo

El esquema de autorización de dspace funciona a partir de 3 elementos:

### 1. E-Person

- representa un usuario dentro del sistema.
- puede ser una persona u otra aplicación

### 2. E-Group:

- representa un grupo de trabajo.
- Se compone de :
  - múltiples E-Person, y/o
  - otros E-Group
- Existen dos grupos del sistema: Anonymous, Administrator
- Todos los usuarios pertenecen al grupo Anonymous

Un E-Person puede pertenecer a muchos grupos (directa o transitivamente)



## Autorización - Modelo

### 3. Resource Policies: derechos explícitos de la forma

(Objeto, Usuario, Derecho, Fecha inicio, Fecha fin)

<p><b>Objeto</b> es cualquier DSO</p> <ul style="list-style-type: none"><li>• Comunidad</li><li>• Colección</li><li>• Ítem</li><li>• Bundle</li><li>• Bitstream</li></ul>	<p><b>Usuario</b> es:</p> <ul style="list-style-type: none"><li>• E-Person</li><li>• E-Group</li></ul>	<p><b>Derecho</b> puede ser:</p> <ul style="list-style-type: none"><li>• READ: ver o descargar</li><li>• WRITE: modificar datos</li><li>• ADD: agregar a un contenedor</li><li>• REMOVE: eliminar de un contenedor</li><li>• ADMIN: administración plena de elementos</li></ul>
---	--	---

*Por defecto, todas las comunidades, colecciones e ítems tienen permiso de READ para el grupo Anonymous*



# Servicios para usuarios



## Acceso a la información - Exploración

El repositorio debe brindar mecanismos para acceder a los contenidos que aloja:

### → 1. Exploración a partir de metadatos

- ◆ Autores
- ◆ Temas
- ◆ Años
- ◆ etc

### Listar por autor

0-9 A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

O introducir las primeras letras:  Ir

Orden:  Resultados:

Mostrando ítems 1-60 de 2093 [Página siguiente](#)

Abásolo, María José [5]

Abbattista, L. (Facultad de Odontología) [1]

Abbattista, M (Facultad de Odontología) [1]

Abdallah, Chaouki T. (Ibero American Science and Technology Education) [1]

Abedini, Walter Ismael [5]

Aboujo, Ana Lía (Facultad de Ciencias Naturales y Museo) [1]

Abrahamovich, Eliana (Centro de Investigaciones en Fitopatología ) [1]

Abreu de Utermohlen, Hilda [1]

Acciaresi, Horacio Abel [12]

Achinelli, Fabio Germán (Facultad de Ciencias Agrarias y Forestales) [7]

Aguirre, Verónica (Instituto de Investigación en Informática) [4]

Agulló, Luis (Universitat Politècnica de Catalunya) [1]

Aguyaro, Mariano Javier (Centro de Investigación de Tecnología del Cuero) [2]

Agüero, Matías Emanuel [1]

Ahmad, N. (Pakistan Atomic Energy Commission) [1]

Aiello, Hugo [1]

Aimetta, Corina (Centro de Estudios en Rehabilitación Nutricional y Desarrollo Infantil) [10]

Ainchil, Jerónimo Enrique (Universidad Nacional de La Plata) [4]

Almaraz, Laura Beatriz [5]

Almassio, Marcela Fabiana [3]

Almeida, Karina Figueiredo de (Universidade Federal do Rio de Janeiro) [1]

Almirón, Adriana E. (Facultad de Ciencias Naturales y Museo) [4]

Alonso Guzmán, Elia Mercedes [4]

Altolaquirre, María Fátima (Instituto de Hidrología de Llanuras) [3]

Altuzarra, César (Universidad Nacional de Rosario) [1]

Alvarez Davila, Manuel [1]

Alvarez Martini, Carlos Alberto (Centro de Investigación en Metrología y Calidad) [2]



# Acceso a la información - Exploración

## → 2. A partir de su organización lógica de comunidades, subcomunidades y colecciones

### Unidades académicas

- o Facultad de Arquitectura y Urbanismo
- o Facultad de Bellas Artes
- o Facultad de Ciencias Agrarias y Forestales
- o Facultad de Ciencias Astronómicas y Geofísicas
- o Facultad de Ciencias Económicas
- o Facultad de Ciencias Exactas
- o Facultad de Ciencias Jurídicas y Sociales
- o Facultad de Ciencias Médicas
- o Facultad de Ciencias Naturales y Museo
- o Facultad de Ciencias Veterinarias



### Red de Universidades con Carreras en Informática (RedUNCI)

La Red de Universidades con Carreras en Informática (RedUNCI) coordina actividades académicas relacionadas con el perfeccionamiento docente, la actualización curricular, la utilización de recursos compartidos, y establece un marco de colaboración para el desarrollo de las actividades de postgrado en las carreras Ciencias de la Computación y/o Informática.

 [Acceder al sitio web](#)

#### Explorar:

[Listar todo](#) [Autores](#) [Temas](#)

#### Subcomunidades en esta comunidad

[Eventos](#)

#### Colecciones en esta comunidad

[Libros, actas de congresos y otros documentos](#)



### CACIC 2015

**Nombre del evento:** XXI Congreso Argentino de Ciencias de la Computación

**Fecha:** 5 al 9 de octubre de 2015

**Lugar:** Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA); Sede Junín

**Comité organizador:** Claudia Russo; Tamara Ahmad; Natalia Anolles; Paula Lencina; Leonardo Mangold; Daniela Pérez; María Linda Trinidad Picco; Sabina Rodríguez; Eliana Serrano; Oscar Spada; Carlos Di Cicco; Mónica Sarobe

**ISBN:** 978-987-3724-37-4

**Materias:** Ciencias Informáticas

 [Acceder al sitio web](#)


**Explorar:** [Listar todo](#) [Autores](#) [Temas](#)



# Acceso a la información - Exploración

→ 3. A partir de su presentación en la web

Navegue por nuestras colecciones

-  **Tesis**  
Tesis de grado, post-grado y otros documentos
-  **Revistas**  
Publicaciones en revistas científicas
-  **Eventos**  
Ponencias realizadas en congresos y conferencias
-  **Libros**  
Libros digitalizados y e-books
-  **Red UNCI**  
Artículos y ponencias de la Red UNCI

## Archivo Institucional



Documentos de carácter institucional, informativos, normativos o administrativos de la Universidad de Salamanca.

## Biblioteca Digital



Colecciones patrimoniales de documentos históricos y fondos específicos digitalizados de la USAL.

## Repositorio Científico



Investigación científica producida o editada por los departamentos y centros de la Universidad de Salamanca

## Repositorio Docente



Documentos de carácter didáctico producidos por la Universidad de Salamanca y entidades colaboradoras



## Acceso a la información - Búsqueda

### → 4. Búsqueda

Se realiza a partir de

#### A) metadatos de los objetos digitales

- por autor
- por título
- por área temática
- etc.

B) el texto completo de la obra (si está disponible)

#### Search

Search: All of DSpace

health

**Filters**  
Use filters to refine the search results.

Subject + Equals + public

**New Filters:**

Title + Contains +

Showing 10 out of a total of 227 results. (0.227 seconds)

1 2 3 4 ... 23 [Next Page](#)

[How to improve public health systems : lessons from Tamil Nadu](#)

Das Gupta, Monica; Desikachari, B.R.; Somanathan, T.V.; Padmanaban, P. (2009-10-01)

[How might India's public health systems be strengthened ?](#)

Das Gupta, Monica; Shukla, Rajendra; Somanathan, T.V.; Datta, K.K. (2009-11-01)



## Acceso a la información - Búsqueda facetada

Un facet es un **filtro** generado dinámicamente a partir de una búsqueda o exploración previa

Permite refinar los resultados a partir de múltiples criterios

Muy útiles para localizar contenidos en grandes conjuntos de datos

Refine su búsqueda

Tipo de documento

- Artículo (9)
- Clase (2)
- Documento de trabajo (6)
- Objeto de conferencia (67)
- Preprint (6)
- Reporte (5)
- Tesis de doctorado (3)
- Tesis de grado (2)

Fecha de publicación

- 2010 - 2016 (100)**
- 2016 (5)
- 2015 (26)
- 2014 (22)
- 2013 (22)
- 2012 (14)
- 2011 (7)
- 2010 (4)

Materia

- Ciencias Informaticas (89)
- Bibliotecologia (77)
- Educacion (10)
- Informatica (4)
- Ciencias de la Educacion (1)

Unidades académicas → Presidencia →

### PREBI-SEDICI

Resultados de su búsqueda...

**Objeto de conferencia** Curso de Posgrado dictado en la Facultad de Informática  
Curso de posgrado: Bibliotecas y repositorios digitales. Tecnología  
De Giusti, Marisa Raquel; Servicio de Difusión de la Creación Intelectual (SE

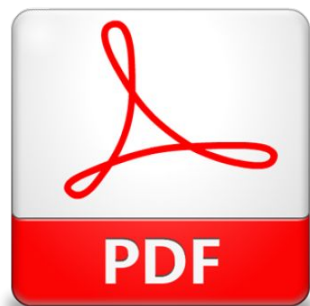
**Objeto de conferencia** Semana del Acceso Abierto (Montevideo, Uruguay, 2013)  
Jornada de sensibilización: Hacia la construcción de un repositori  
De Giusti, Marisa Raquel

**Objeto de conferencia** III Conferencia de Bibliotecas y Repositorios Digitales de A  
Infraestructura interoperable alrededor del repositorio institucional  
De Giusti, Marisa Raquel; Lira, Ariel Jorge; Villarreal, Gonzalo Luján; Terruzz

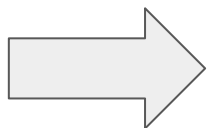
**Objeto de conferencia** III Conferencia de Bibliotecas y Repositorios Digitales de A  
Discurso de inauguración del BIREDIAL 2013

## Acceso a la información - Búsqueda por texto completo

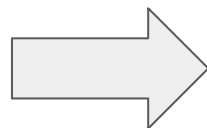
- Permite localizar objetos digitales más allá de sus metadatos
- Las búsquedas no se realizan directamente sobre los documentos sino sobre indexaciones previas de su contenido.
- Requiere extraer el texto de los OD e integrarlo en un sistema de búsquedas



Documento



Texto extraído



Índice de texto



## Depósito - Autoarchivo

El **autoarchivo** es un proceso mediante el cual los usuarios envían materiales a través de un formulario de carga sumamente sencillo.

Este proceso es siempre **mediado** por algún administrador del sistema, quien decide si el contenido es depositado o es rechazado.



## Estadísticas

- Registra datos de:
  - accesos, descargas, navegación
  - en un índice Solr
- A diferencia de herramientas como Google Analytics y PIWIK, los registros se crean por cada acceso al servidor y no desde un agente de usuario.
- provee estadísticas en vivo
  - acceso a ítems por origen y fecha,
  - descargas de bitstreams por origen y fecha,
  - ítems más accedidos del repositorio



# INTEROPERABILIDAD



## Interoperabilidad - Introducción

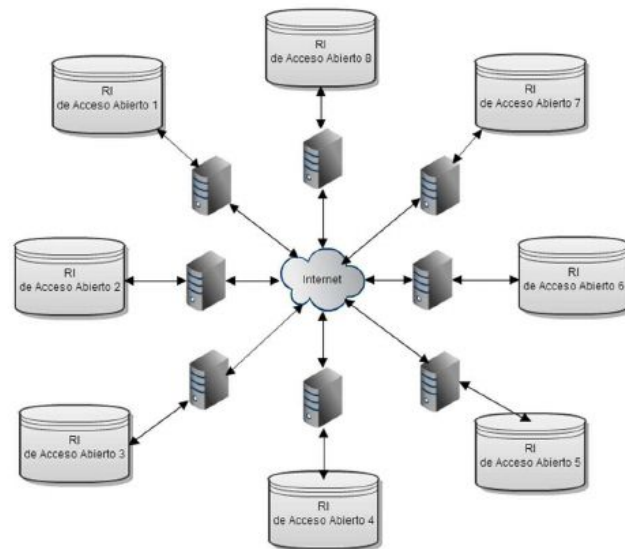
### ¿Cómo se define la interoperabilidad?

Capacidad de intercambio transparente de información entre dos o más sistemas.

### ¿Para que interoperar?

- Para formar parte de una red de repositorios (y no ser una isla de información).
- Para aumentar la visibilidad del contenido del repositorio a nivel internacional.
- Para normalizar la información expuesta (mediante el cumplimiento de estándares).

Los repositorios digitales deben pensarse como sistemas interoperable desde el principio.



## Interoperabilidad - Usos

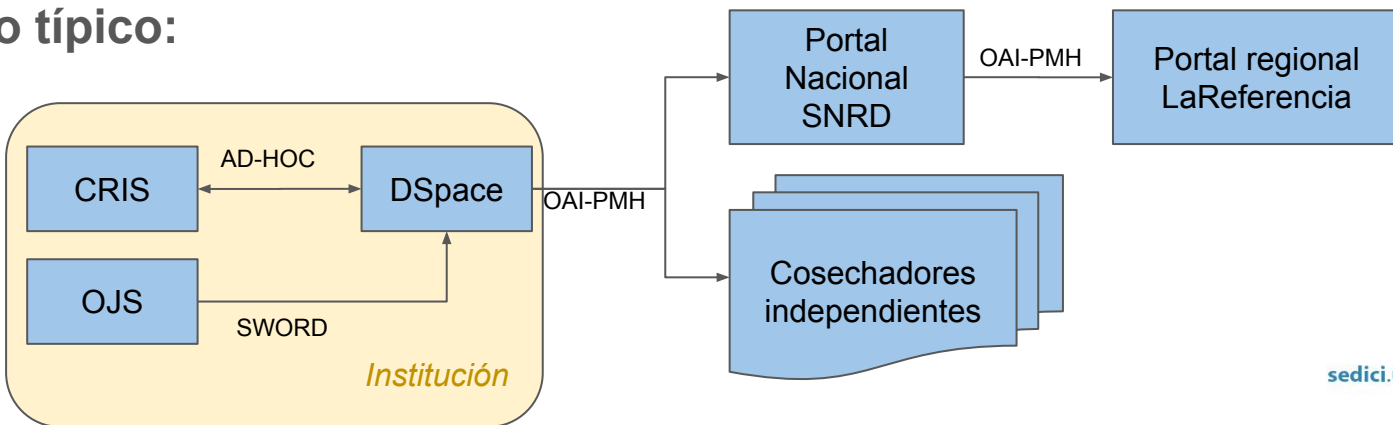
### Desde el repositorio

- integrarse con sistemas de la institución
- ampliar el alcance y difusión de los contenidos
- incorporarse a sistemas o redes regionales e internacionales

### Hacia el repositorio

- facilitar y/o agilizar la ingesta de contenidos
- modificar remotamente los contenidos

### Ejemplo típico:

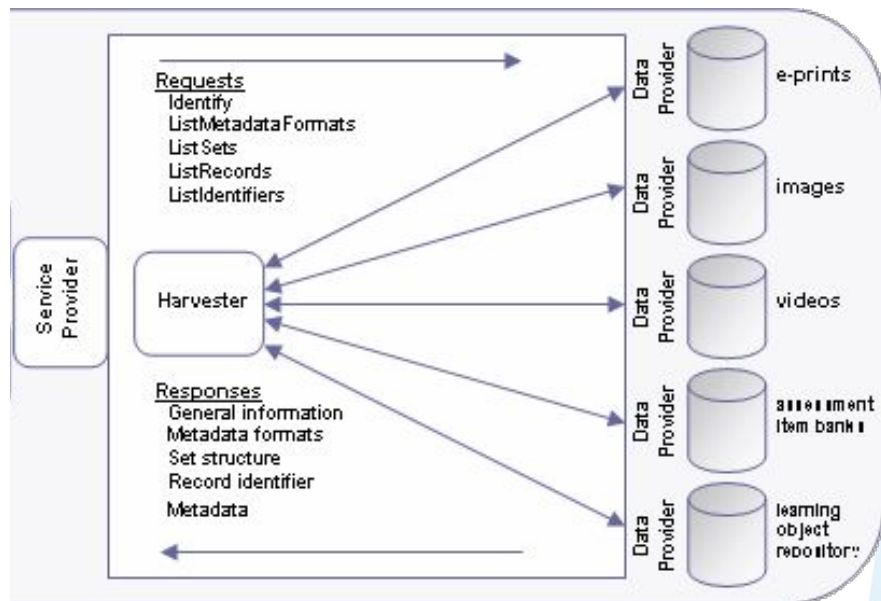


## Interoperabilidad - Protocolo OAI-PMH

*Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH) es un mecanismo para permitir la interoperabilidad entre repositorios

Prevé 2 roles:

- Data Provider
- Service Provider



## Interoperabilidad - Verbs del protocolo OAI-PMH

<i>Identify</i>	Se utiliza para obtener información del repositorio <a href="http://sedici.unlp.edu.ar/xoai/snrd?verb=Identify">http://sedici.unlp.edu.ar/xoai/snrd?verb=Identify</a>
<i>Get Record</i>	Se utiliza para obtener un registro específico <a href="http://sedici.unlp.edu.ar/xoai/snrd?verb=GetRecord&amp;metadataPrefix=oai_dc&amp;identifier=oai:sedici.unlp.edu.ar:10915/1063">http://sedici.unlp.edu.ar/xoai/snrd?verb=GetRecord&amp;metadataPrefix=oai_dc&amp;identifier=oai:sedici.unlp.edu.ar:10915/1063</a>
<i>List Identifiers</i>	Se utiliza para obtener los identificadores de los registros <a href="http://sedici.unlp.edu.ar/xoai/snrd?verb=ListIdentifiers&amp;metadataPrefix=oai_dc">http://sedici.unlp.edu.ar/xoai/snrd?verb=ListIdentifiers&amp;metadataPrefix=oai_dc</a>
<i>List Metadata Formats</i>	Se utiliza para listar los <i>metadata format</i> disponibles en el repositorio para diseminar contenido <a href="http://sedici.unlp.edu.ar/xoai/snrd?verb=ListMetadataFormats">http://sedici.unlp.edu.ar/xoai/snrd?verb=ListMetadataFormats</a>
<i>List Records</i>	Se utiliza para obtener todos los registros <a href="http://sedici.unlp.edu.ar/xoai/snrd?verb=ListRecords&amp;metadataPrefix=oai_dc">http://sedici.unlp.edu.ar/xoai/snrd?verb=ListRecords&amp;metadataPrefix=oai_dc</a>
<i>List Sets</i>	Se utiliza para obtener los sets <a href="http://sedici.unlp.edu.ar/xoai/snrd?verb=ListSets">http://sedici.unlp.edu.ar/xoai/snrd?verb=ListSets</a>



## Interoperabilidad - Protocolo OAI-PMH

¿Cómo se **entienden entre sí** los *service* y *data provider*?

1. Esquema Dublin Core simple (oai\_dc)
  - USO OBLIGATORIO
2. Uso de múltiples esquemas estándares
  - mods,
  - mets,
  - etc
3. Acuerdo de uso de un esquema específico
  - Ej. arXiv
4. Directrices de interoperabilidad
  - SNRD: [http://repositorios.mincyt.gob.ar/pdfs/Directrices\\_SNRD\\_2015.pdf](http://repositorios.mincyt.gob.ar/pdfs/Directrices_SNRD_2015.pdf)
  - Open Aire: <https://guidelines.openaire.eu/en/latest/>
  - Driver2



## Interoperabilidad - Directrices de interoperabilidad sobre OAI-PMH

### ¿Cuál es su objetivo?

Establecer un perfil de metadatos a partir de reglas de uso de los metadatos que definen:

- Sintaxis: uso libre, estructura definida o vocabulario controlado.
- Obligatoriedad: uso obligatorio, recomendado u opcional.
- Otras restricciones: orden de instancias de metadatos.



**SNRD**

Sistema Nacional de  
Repositorios Digitales

### ¿Qué utilidad tiene cumplir con las directrices ?

- SNRD <http://repositoriosdigitales.mincyt.gob.ar/>
- LA Referencia <http://lareferencia.redclara.net/rfr/>
- EUROPEANA <http://www.europeana.eu/portal/en>
- Tesis Latinoamericanas

LA REFERENCIA



europeana

[sedici.unlp.edu.ar](http://sedici.unlp.edu.ar)



## Arquitectura - Interoperabilidad - OAI PMH - Usos

- Ejemplos

	Driver 2.0	SNRD 2015	Open Aire 3.0
<b>Uso de dc:rights</b>	Uso: Recomendado Esquema: no especificado	Uso: Obligatorio Esquema: vocabulario OpenAIRE 3.0. Se debe exponer como primera instancia	Uso: Obligatorio Esquema: vocabulario propio de 4 niveles de acceso: closed, embargoed, restricted y open Access



**SNRD**

Sistema Nacional de  
Repositorios Digitales



## Interoperabilidad - Módulo OAI2 en DSpace

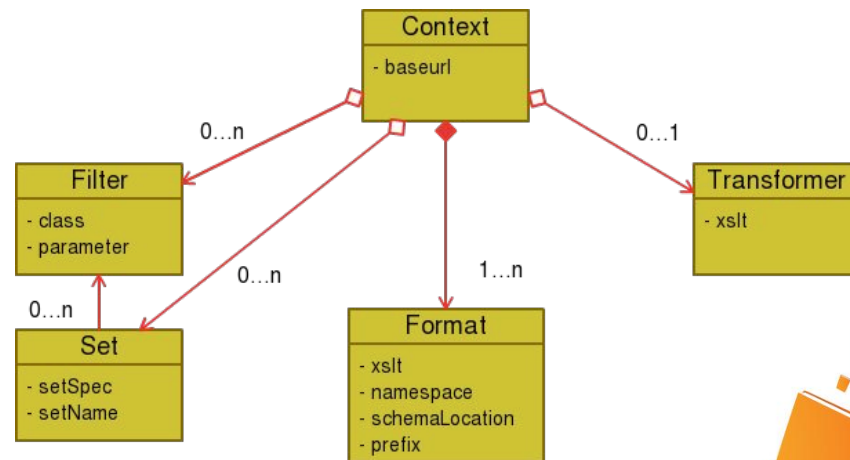
### DSpace como data provider

DSpace provee un módulo OAI 2.1 el cual permite al repositorio funcionar como *data provider*.

### Características

Configuración flexible basada en:

- Contexto (*Context*)
- Filtro (*Filter*)
- Transformador (*Transformer*)
- Mapeador (*Format*)



## Interoperabilidad - RSS y OpenSearch



RSS se utiliza para la sindicación de noticias y contenidos en línea

OpenSearch es un conjunto de tecnologías que permiten publicar los **resultados de una búsqueda** en un formato adecuado para la sindicación y agregación (fuente: Wikipedia)

- Permite que otras aplicaciones y sitios web expongan contenidos del repositorio
- Se integra fácilmente mediante RSS/Atom

### Últimos documentos agregados

Pamidronato endovenoso vs. Alendronato oral en el tratamiento de la osteoporosis establecida  
*Ferrari, Eliseo;*

Efecto de la fertilización con nitrógeno y fósforo sobre la acumulación de biomasa de pasturas de *Panicum coloratum* implantadas sobre suelos hidro-halomórficos de la Pampa Deprimida  
*Insausti, Mariano;*

Aislamiento, caracterización genética y análisis filogenético de *Rickettsia massiliae*, Ciudad Autónoma de Buenos Aires  
*Cicuttin, Gabriel L.;*

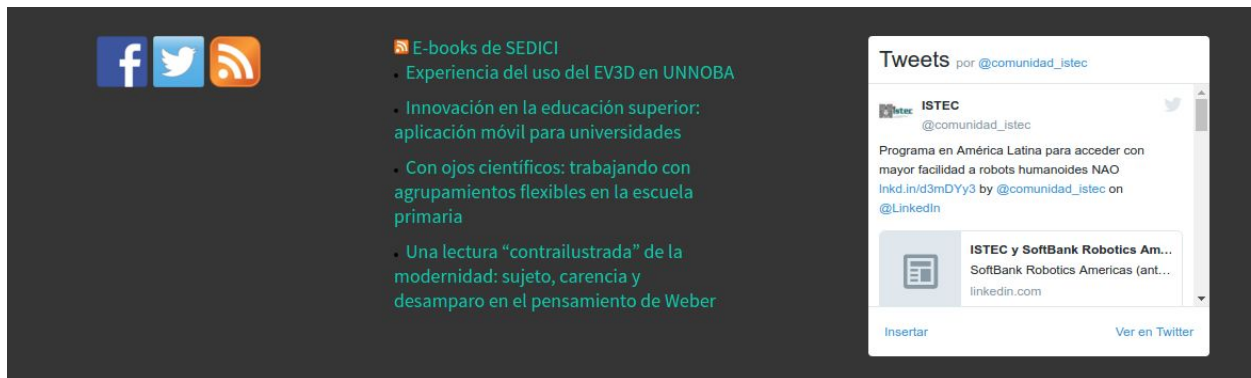
Propuesta pedagógica para la Cátedra de Climatología y Fenología Agrícola  
*Pinciroli, María;*

Agujeros negros astrofísicos  
*Pérez, Daniela;*



## Arquitectura - Interoperabilidad - RSS y OpenSearch

- Caso de uso: permitir que una web incluya un listado de recursos cargados en el repositorio. Ej últimos libros cargados, obras de un autor, etc.



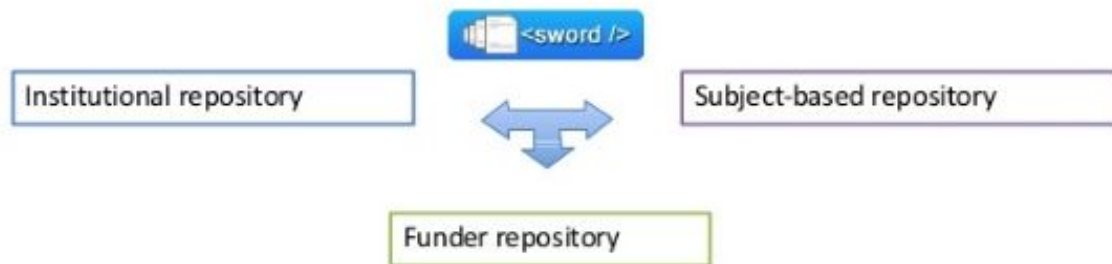
The screenshot shows a dark-themed interface. On the left, there are three social media icons: Facebook, Twitter, and RSS. To the right of these icons is a list of items under the heading "E-books de SEDICI". The list contains three items, each with a small icon and a title. The first item is "Experiencia del uso del EV3D en UNNOBA". The second item is "Innovación en la educación superior: aplicación móvil para universidades". The third item is "Con ojos científicos: trabajando con agrupamientos flexibles en la escuela primaria". Below these items is a fourth item, "Una lectura 'contrailustrada' de la modernidad: sujeto, carencia y desamparo en el pensamiento de Weber". To the right of the list is a white box containing a tweet. The tweet is from the account "ISTEC" (@comunidad\_istec) and contains the text: "Programa en América Latina para acceder con mayor facilidad a robots humanoides NAO". Below the text is a link "lnkd.in/d3mDYy3 by @comunidad\_istec on @LinkedIn". There is also a small image of a document icon and the text "ISTEC y SoftBank Robotics Am... SoftBank Robotics Americas (ant... linkedin.com". At the bottom of the tweet box are the buttons "Insertar" and "Ver en Twitter".

- En dspace se configura desde el archivo dspace.cfg ([RSS](#) - [OpenSearch](#))



## Interoperabilidad - Protocolo SWORD

- SWORD (Simple Web-service Offering Repository Deposit) <http://swordapp.org/>
  - a. permite el **depósito remoto** de contenidos en un repositorio
- ¿Para qué sirve?
  - a. transferencia desde/hacia otros sistemas de publicación (ej OJS)
  - b. depósito de múltiples documentos (metadatos y binarios) a la vez
  - c. depósito simultáneo en múltiples repositorios



## Interoperabilidad - Protocolo SWORD

### → ¿Cómo funciona?

- a. El cliente se autentica con el servidor
- b. El cliente solicita al servidor en reglamento de uso denominado ServiceDocument
- c. El cliente envía el paquete a depositar indicando el punto de depósito
- d. El servidor recibe el paquete, lo abre, extrae la información de acuerdo a la representación interna del repositorio y completa el depósito.
- e. El cliente recibe el resultado del depósito

### → Se integra con otras aplicaciones

- a. Existen librerías para desarrolladores en PHP, Java, Ruby, Python
- b. Existen plugins para la dspace, ojs, ocs, entre otros.



# Interoperabilidad - SWORD

## Caso de uso

### Portal de Revistas de la UNLP

- Una revista de la UNLP publica un nuevo número en su sitio web (OJS)
- Un administrador del sitio selecciona los artículos del nuevo número y los envía a una colección del repositorio SEDICI (< 1 min)
- Un administrador de SEDICI verifica y completa los registros, y publica el nuevo número en el repositorio (~30 min)

DSpace 5 SWORD v2 server

<https://wiki.duraspace.org/display/DSDOC5x/SWORDv2+Server>

Conector de depósito de importación/exportación SWORD

Punto de depósito: **sedici** ADMINISTRADOR

Nombre usuario/a:

Contraseña:

Punto de depósito: **Autoarchivo** **Actualizar**

Opciones:  depositar galerías  
 depositar los archivos de editorial más recientes

NÚMERO	TÍTULO	AUTORES/AS
<input checked="" type="checkbox"/> VOL_1_2020	Aplicação de processo oxidativo avançado baseado em fotocatalise heterogênea (TiO <sub>2</sub> /UV-vis) para o pré-tratamento de efluente têxtil	
<input checked="" type="checkbox"/> VOL_1_2021	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_1_2021	Avaliação da Redução da Poluição do Cloruro Tratado por Processo Fotoquímico	



## SEDICI

REPOSITÓRIO INSTITUCIONAL DE LA UNLP

Início Buscar material Subir material

Administración

### Tareas del flujo de trabajo

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	Item	Colección	Ramitente
No tiene asignada ninguna tarea			

#### Tareas en cola

Tarea	Item	Colección	Ramitente
<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



# SEGUNDA PARTE



# Herramientas de automatización y Estructura de proyectos



## Herramientas de automatización - Maven

Es una software para gestión de proyectos Java que permite controlar actividades recurrentes como:

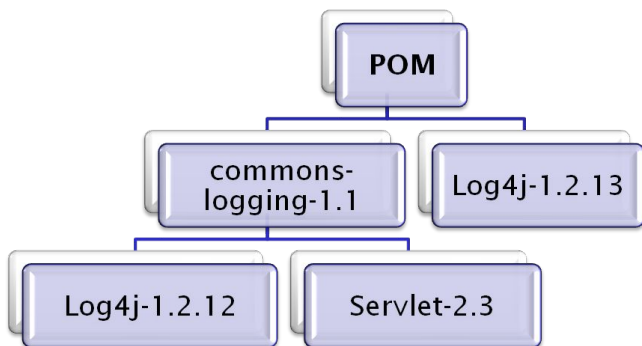
- recuperación de dependencias y resolución de conflictos entre ellas
- filtrado de archivos de configuración
- compilación,
- empaquetado (jar, war, pom),
- deploy,
- generación de documentación,
- e infinidad de acciones más.

**Maven**<sup>TM</sup>



## Herramientas de automatización - Maven

Su funcionamiento se rige por uno o varios archivos **POM.xml** que describen el proyecto, sus dependencias a otros módulos y librerías, y configuran las acciones de cada etapa.



```
<project xmlns="http://maven.apache.org/POM/4.0.0" xmlns:xsi="http://www.w3.org/2001/XMLSchema-ins
<modelVersion>4.0.0</modelVersion>
<groupId>com.example.pom</groupId>
<artifactId>sampleProject</artifactId>
<version>0.0.1-SNAPSHOT</version>
<packaging>war</packaging>
<name>sampleProject</name>
<description>This is a sample maven project for adding dependencies.</description>

<dependencies>
  <dependency>
    <groupId>org.springframework.security</groupId>
    <artifactId>spring-security-parent</artifactId>
    <version>3.0.8.RELEASE</version>
  </dependency>
</dependencies>

</project>
```

### Comandos típicos:

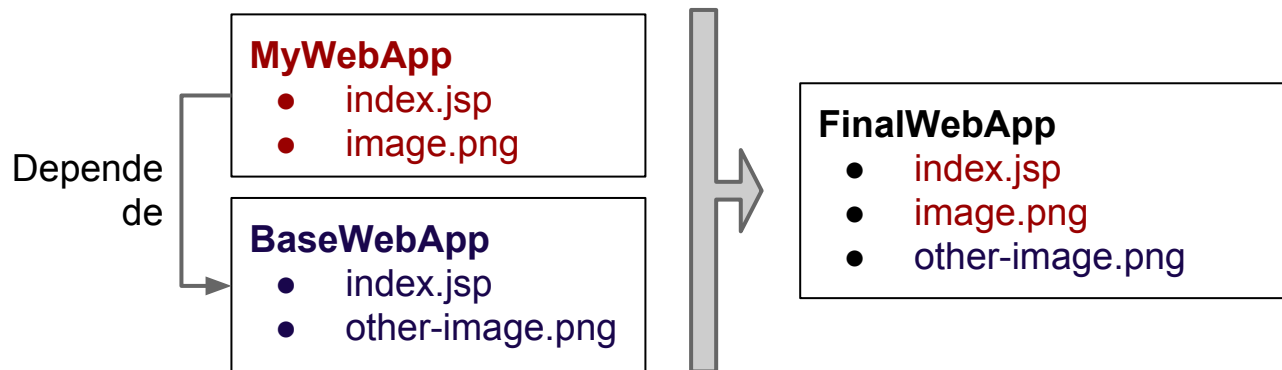
Compilar proyecto: `mvn package`

Limpiar proyecto: `mvn clean`



## Herramientas de automatización - Maven Overlays

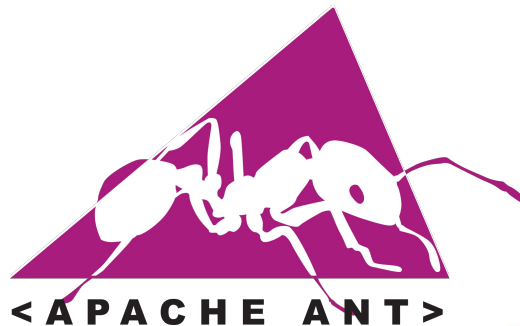
- Permiten "extender" una aplicación web (war), agregando nuevos elementos o redefiniendo los existentes.
- En maven, se hace overlay siempre que una aplicación web depende de otra aplicación web.
- Maven permite configurar explícitamente el orden en el que se realizará el overlay y qué recursos deben considerarse en el proceso



## Herramientas de automatización - Ant

Apache Ant es una librería de Java y una herramienta que permite automatizar tareas de cualquier tipo.

- Su funcionamiento se rige por el archivo `build.xml`
- Ant provee tareas predefinidas que permiten compilar, testear y correr aplicaciones Java.



## Herramientas de automatización - Empaquetado e instalación

**Maven** es necesario en las etapas de compilación y empaquetado de DSpace.

- Compila y empaqueta todas las aplicaciones y el core de DSpace
- Genera el paquete de instalación en  
[dspace-source]/dspace/target/dspace-installer

```
cd [dspace-source] && mvn package
```

**ANT** es necesario en la etapa final de instalación/actualización

- se ejecuta desde [dspace-source]/dspace/target/dspace-installer
- Instalación
  - `ant fresh_install`
- Actualización
  - `ant update`



## Estructura de proyectos

Un proyecto principal *dspace-parent* con múltiples subproyectos, en varios niveles:

- Un subproyecto por módulo
  - dspace-api
  - dspace-jspui
  - dspace-oai
  - dspace-rest
  - dspace-solr
  - dspace-swordv[1|2]
  - dspace-xmlui
- Un subproyecto especial para la generación del instalador:
  - dspace



## Estructura de proyectos

La relación entre proyectos se mantiene con Maven

- verticalmente entre proyectos y subproyectos, configurando **módulos**
- horizontalmente entre proyectos no emparentados, mediante **dependencias**

Tenemos la posibilidad de **personalizar** DSpace:

1. *modificando* los módulos Maven existentes pensados para este fin:
  - se encuentran en el directorio `dspace/modules`
  - dependen de los módulos originales y hacen Overlay
  - el más importante es *additions*
2. *agregando* por nuestra cuenta un proyecto Maven aparte y marcándolo como subproyecto en el POM de DSpace.



# Configuración de DSpace



## Configuración - Archivos local.cfg y dspace.cfg

En local.cfg puede ser usado para almacenar cualquier cambio en la configuración locales a nuestra instalación.

- Ubicado en `[dspace-source]/local.cfg`

En dspace.cfg es el archivo principal de configuración. En este se mantiene las configuraciones más importantes de DSpace de aspecto general.

- Ubicado en `[dspace-source]/dspace/config/dspace.cfg`



# Configuración - local.cfg y dspace.cfg - Configuraciones típicas

## dspace.cfg

- datos de conexión con el motor de base de datos
- configuraciones del módulo de exploración
- configuraciones de idioma
- configuraciones de plugins (autoridades, media-filters, authentication)
- configuraciones de visualización de metadatos (del formulario de carga)
- ***se indica qué otros módulos de configuración se cargarán***
  - discovery.cfg
  - oai.cfg
  - etc.

## local.cfg

- directorio de instalación de DSpace
- configuración de email saliente
- **código** de Google Analytics
- usuario y **password** de acceso a base de datos
- se pueden sobrescribir propiedades de otros módulos de configuración



# Configuración - local.cfg y dspace.cfg - Ejemplo

## local.cfg

```
dspace.dir = /var/repo/dspace/cic-digital/install
dspace.baseUrl = http://digital.cic.gba.gob.ar
#Database host, user && password
db.url =
jdbc:postgresql://pgdatabase.digital.cic.gba.gob.ar:5432/dspace
db.username = cic_digital
db.password = cic-f1bdf7bce8b
#Handle config
handle.prefix = 11746
```

## dspace.cfg

```
dspace.dir = /dspace
dspace.baseUrl = http://localhost:8080
#Database host, user && password
db.url = jdbc:postgresql://localhost:5432/dspace
db.username = dspace
db.password = dspace
#Handle config
handle.canonical.prefix = http://hdl.handle.net/
handle.prefix = 123456789
### i18n - Locales / Language ####
default.locale = es
webui.supported.locales = es, en, pt
```

## Configuración resultante

```
dspace.dir = /var/repo/dspace/cic-digital/install
dspace.baseUrl = http://digital.cic.gba.gob.ar
#Database host, user && password
db.url =
jdbc:postgresql://pgdatabase.digital.cic.gba.gob.ar:5432/dspace
db.username = cic_digital
db.password = passssssword
#Handle config
handle.canonical.prefix = http://hdl.handle.net/
handle.prefix = 11746
### i18n - Locales / Language ####
default.locale = es
webui.supported.locales = es, en, pt
```



## Configuración - input-forms.xml.

En el archivo **input-forms.xml** podemos configurar la visibilidad de los metadatos a completar durante la carga de ítems. Ejemplo:

### Agregar campo:

Para agregar un campo para el metadato 'thesis.degree' de forma sencilla, es necesario configurar el siguiente `<field>` en **input-forms.xml**:

```
<field>
  <dc-schema>thesis</dc-schema>
  <dc-element>degree</dc-element>
  <dc-qualifier>level</dc-qualifier>
  <label>Nivel de grado de la Tesis</label>
  <input-type>onebox</input-type>
  <hint>Indique el nivel de grado de su trabajo de Tesis según los valores especificados en el estándar de
NDLTD.</hint>
</field>
```



## Configuración - discovery.cfg

En discovery.cfg se encuentra la configuración general de Discovery.

### Ejemplo:

Para determinar la ubicación del índice de Solr se edita la propiedad **discovery.search.server**.

```
discovery.search.server=[http://localhost:8080/solr/search]
```

**NOTA:** La modificación de esta propiedad sobrescribe las propiedades definidas en **local.cfg**



## Configuración - discovery.xml.

En **discovery.xml** se encuentra la configuración de la interfaz de Discovery.

Ejemplo: No indexar ciertos metadatos.

Excluir metadato dc.rights para comunidades/colecciones:

```
<property name="toIgnoreMetadataFields">  
  <map>  
    <entry>  
      <key><util:constant static-field="org.dspace.core.Constants.COMMUNITY"/></key>  
      <list>  
        <value>dc.rights</value>  
      </list>  
    </entry>  
    <entry>  
    </map>  
</property>
```



## Configuración - discovery.xml.

Agregar un nuevo filtro de búsqueda que no está incluido como **facet**:

```
<bean id="searchFilterTitle"
class="org.dspace.discovery.configuration.DiscoverySearchFilter">
  <property name="indexFieldName" value="title"/>
  <property name="metadataFields">
    <list>
      <value>dc.title</value>
    </list>
  </property>
</bean>
```

### Search

Search: All of DSpace

health

Filters  
Use filters to refine the search results.

Subject Equals public

New Filters:  
Title Contains

Apply

Showing 10 out of a total of 227 results. (0.227 seconds)

1 2 3 4 ... 23 Next Page

[How to improve public health systems : lessons from Tamil Nadu](#)  
Das Gupta, Monica; Desikachari, B.R.; Somanathan, T.V.; Padmanaban, P. (2009-10-01)

[How might India's public health systems be strengthened ?](#)  
Das Gupta, Monica; Shukla, Rajendra; Somanathan, T.V.; Datta, K.K. (2009-11-01)



# Configuración - discovery.xml.

## Agregar un Facet

Defines which metadata fields should be offered as a contextual sidebar browse options, each of these facets has also got to be a search filter

```
<bean id="searchFilterAuthor"  
class="org.dspace.discovery.configuration.DiscoverySearchFilterFacet">  
  <property name="indexFieldName" value="author"/>  
  <property name="metadataFields">  
    <list>  
      <value>dcterms.creator.author</value>  
      <value>dcterms.creator.compiler</value>  
      <value>dcterms.creator.corporate</value>  
      <value>dcterms.creator.editor</value>  
    </list>  
  </property>  
  <property name="facetLimit" value="10"/>  
  <property name="sortOrderSidebar" value="COUNT"/>  
  <property name="sortOrderFilterPage" value="COUNT"/>  
</bean>
```

Autor
Conti, Alfredo Luis (Laboratorio de Investigaciones del Territorio y el Ambiente) (34)
Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CICBA) (21)
Contin, Mabel I. (Laboratorio de Investigaciones del Territorio y el Ambiente) (20)
Morosi, Julio A. (Laboratorio de Investigaciones del Territorio y el Ambiente) (17)
Mariñelarena, Patricia Ines Maria (Laboratorio de Investigaciones del Territorio y el Ambiente) (16)
Pacheco, Mara (Laboratorio de Investigaciones del Territorio y el Ambiente) (15)
Amarilla, Beatriz (Laboratorio de Investigaciones del Territorio y el Ambiente) (14)
Coletti, Renaldo (Laboratorio de Investigaciones del Territorio y el Ambiente) (14)
De Giusti, Marisa Raquel (12)
Molinari, Graciela Alicia (Laboratorio de Investigaciones del Territorio y el Ambiente) (11)
... ver más



## Vías de ingesta de ítems



## Vías de ingesta

### Cómo ingresan los recursos al repositorio?

- Submission
  - Autoarchivo ⇒ asistente sencillo para los usuarios
  - Carga desde la administración
- Importación ⇒ carga masiva de ítems en la cola de procesamiento
- SWORD ⇒ depósito de ítems desde otros sistemas
- Cliente OAI ⇒ importación desde otros repositorios

Independientemente de la vía de ingesta, se puede hacer que los ítems pasen por un WORKFLOW de revisión.



## Submission

Permite definir circuitos de carga de acuerdo al lugar dónde se desea subir el material, es decir, una comunidad, una colección o en cualquier lugar.

Cada circuito o *submission process* contiene una serie de pasos que incluyen típicamente

- carga de campos en un formulario (DescribeStep)
- subida de archivos
- selección de licencia de uso
- confirmación de datos
- otros pasos customizados, con o sin interfaz.



## Submission

El step Describe define los metadatos del ítem

```
<submission-process name="autoarchiveSubmission">
  <step>
    <processing-
class>org.dspace.submit.step.SkipInitialQuestionsStep</process
ing-class>
  </step>
  <step id="sediciDescribeItem"/>
  <step id="sediciUploadItem"/>
  <step id="sediciVerifyItem"/>
  <step id="sediciCCLicense"/>
  <step id="sediciSignOffLicense"/>
  <step id="sediciWorkflowEditedControl"/> <!-- setea el
flag indicando que se completó el proceso de edición desde el
workflow. Sin interface. -->
</submission-process>
```



## Submission > DescribeStep

El archivo *config/input-forms.xml* define los formularios de carga de metadatos que deberán usarse durante la carga manual de ítems.

Dspace permite definir un formulario de acuerdo al lugar dónde se desea subir el material, es decir, una comunidad, una colección o en cualquier lugar.

Un **form** es un **conjunto de páginas** a través de las cuales los usuarios ingresan sus metadatos



## Submission > DescribeStep



Estructura simplificada del archivo input-forms



## Submission > DescribeStep

```
<form name="traditional">
  <page number="1">
    <!-- Tipo de documento: determina la carga de otros metadatos -->
    <field>
      <dc-schema>dc</dc-schema>
      <dc-element>type</dc-element>
      <label>Tipo de documento (*)</label>
      <input-type value-pairs-name="common_types">dropdown</input-type>
      <hint>Seleccione el Tipo de Documento que desea cargar</hint>
      <required>Debe seleccionar un tipo de documento, ya que el proceso de
    </field>
    <field>
      <dc-schema>sedici</dc-schema>
      <dc-element>subtype</dc-element>
      <label>Especifique el tipo de Artículo (*)</label>
      <input-type value-pairs-name="common_subtypes_articulo">dropdown</input-type>
      <hint>Seleccione el tipo de artículo que desea cargar</hint>
      <required>Debe seleccionar un subtipo de documento</required>
      <visibility>workflow</visibility>
      <type-bind>Articulo</type-bind>
    </field>
    <field>
      <dc-schema>sedici</dc-schema>
      <dc-element>subtype</dc-element>
      <label>Especifique el tipo de Tesis (*)</label>
      <input-type value-pairs-name="common_subtypes_tesis">dropdown</input-type>
      <hint>Seleccione el tipo de tesis que desea cargar</hint>
      <required>Debe seleccionar un subtipo de documento</required>
      <visibility>workflow</visibility>
      <type-bind>Tesis</type-bind>
    </field>
  </page>
</form>
```

Es posible especificar las propiedades de cada campo individual de cada form

Algunos campos sólo se muestran en ciertos tipos (dc.type) de contenidos (**type-bind**)



## Submission > DescribeStep

Cada formulario tiene varias etapas/páginas

The image shows a screenshot of a web form titled "Envío de ítems" (Item Submission) with a progress bar at the top. The form is divided into several sections, each with a corresponding annotation box on the right:

- Submission progress bar:** Points to the progress bar at the top of the form, which includes a "Describe" button and a series of steps: Describe, Editar, Ajustar, Revisar, Librada CC, Librada, and Finalizar.
- Selector simple (select):** Points to the "Tipo de documento (\*)" dropdown menu, which is currently set to "- Seleccione un Tipo -".
- Texto de una línea (text):** Points to the "Autor (\*)" text input field, which is currently empty and includes an "Agregar Otro" button.
- Texto multi línea (textarea):** Points to the "Título (\*)" text area, which is currently empty and includes a placeholder text "El título principal de la obra".
- Calendario:** Points to the "Fecha de Publicación:" date input field, which includes a placeholder text "Fecha en la que la obra fue publicada en una revista, libro, etc. No debe confundirse con la fecha de entrega o defensa de una tesis, que debe cargarse en el campo Fecha de Presentación. Los valores posibles para este campo son día/mes/año."

## Web UI en DSpace - Formularios

Distintos conjuntos de metadatos para distintos tipos de contenidos

- ítems de congresos: nombre de la conferencia, lugar, fecha de exposición
- tesis: grado alcanzado, director, subdirector
- libro: capítulos, ISBN, editorial
- artículo de revista: datos de la revista (nombre, ISSN), número, volumen, año

No es viable (ni aceptable) un gran formulario con todos los posibles campos para todos los posibles tipos de contenidos)

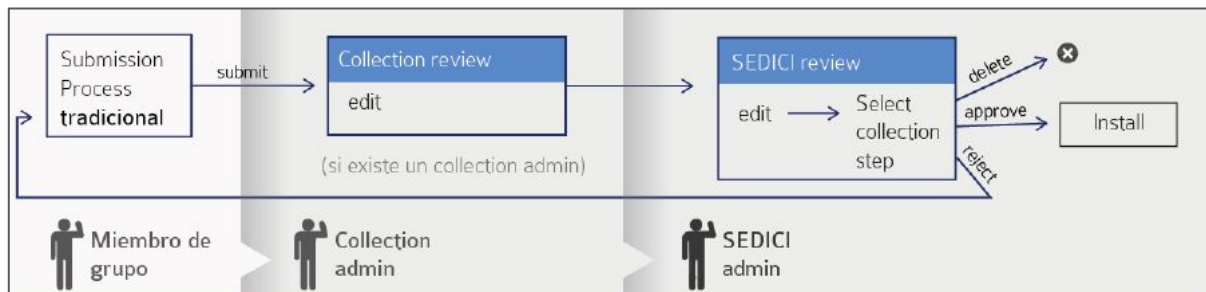
La interfaz de carga debe adaptarse al contenido a cargar...



## Workflow configurable

Cuando se termina de realizar un envío mediante el Submission, el ítem enviado entra al flujo de trabajo (workflow) del sistema, que obligatoriamente deberá recorrer para ser finalmente publicado o descartado.

- Consiste en un conjunto de pasos ejecutar
  - por responsables del repositorio.
  - por el sistema
- Define un flujo de ejecución no secuencial basado en los resultados de ejecución de los pasos



## Workflow configurable > Configuración

En el archivo `config/workflow.xml` DSpace define un *workflow general* denominado “traditional” y permite definir un *workflow específico* para una colección o comunidad específica.

```
<?xml version="1.0" encoding="UTF-8"?>
<wf-config>
  <workflow-map>
    <name-map collection="default" workflow="default"/>
    <name-map collection="11746/2" workflow="autoarchive"/>
  </workflow-map>
  <workflow start="CIC-ADMIN_review" id="autoarchive">
    <roles>
      <role id="Administrador_CIC" name="CIC-ADMIN" scope="repository"
        description="Administradores de CIC-Digital."/>
    </roles>
    <step id="CIC-ADMIN_review" role="Administrador_CIC" userSelectionMethod="claimaction">
      <actions>
        <action id="changecollection"/>
        <action id="editaction"/>
      </actions>
    </step>
  </workflow>
  <workflow start="CIC-ADMIN_review" id="default">
    <roles>
      <role id="Administrador_CIC" name="CIC-ADMIN" scope="repository"
        description="Administradores de CIC-Digital."/>
    </roles>
    <step id="CIC-ADMIN_review" role="Administrador_CIC" userSelectionMethod="rejectrole">
      <actions>
        <action id="editaction"/>
      </actions>
    </step>
  </workflow>
</wf-config>
```

→ Mapeo de workflows

→ Definición de workflow

→ Definición de roles de usuario

→ Definición de un paso:

- identificador (*id*)
- responsables asignados (*role*)
- método de selección de usuarios responsables (*userSelectionMethod*).



# XMLUI

## *Frontend Web*



# Frontend web

**CIC DIGITAL** Repositorio Institucional  
Comisión de Investigaciones Científicas

Inicio Explorar Aportar Material Mas información Contacto Mi cuenta Estadísticas ES

CIC-Digital es el Repositorio Institucional de la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires, creado con el fin de preservar y dejar accesible en abierto toda la producción científico-tecnológica generada en el ámbito de las Instituciones CIC.

### BUSCAR EN EL REPOSITORIO

Ingrese su búsqueda ...

EXPLORAR AUTORES EXPLORAR CENTROS TODO EL REPOSITORIO

#### Añadido Recientemente

Estudio de la dinámica del crecimiento de colonias de células transformadas  
Huelgo, M. A. C.; Pasquale, Miguel Angel; Bolzán, A. E.; Arvia, Alejandro Jorge; Gonzalez, Pedro Horacio

Evolución del estado trófico y estructura del fitoplancton de un lago somero de la Región Pampeana: laguna Sauce Grande (provincia de Buenos Aires, Argentina)  
Rory, Noelia Leonor; Ferrer, Nora Cristina; Cáceres, Eduardo Jorge

Morfología microscópica del pabellón auricular de dos especies de marsupiales de la Argentina

**Subir Material**

A través del proceso de autoarchivo usted podrá contribuir con sus investigaciones y desarrollos científicos propios, al contenido dentro del Repositorio Institucional de la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires.

## DSpace@MIT

Search (Ex: crystalline silicon solar) Búsqueda avanzada

DSpace Principal

- Listar

**Todo DSpace**  
Comunidades y colecciones  
Por fecha de publicación  
Autores  
Títulos  
Temas

• Mi cuenta

**Download statistics now available for MIT Open Access Articles**  
MIT authors can now log in to see how many times each of their own papers has been downloaded, and from which countries. Aggregated data is available by MIT department, lab, or center.

**Deposit Your Work**

- Establish a New Community or Collection
- Deposit to an Existing Collection
- Deposit to the MIT Open Access Collection
- Learn more about DSpace@MIT

## SEDICI

REPOSITORIO INSTITUCIONAL DE LA UNLP

Inicio Buscar material Subir material Institucional Preguntas

**Buscar material**  
Busque entre los mas de 43000 recursos disponibles en el repositorio

**Subir material**  
Suba sus trabajos a SEDICI, para mejorar notoriamente su visibilidad e impacto

**Recursos para su Investigación**  
Acceda a miles de recursos científicos y académicos producidos dentro de la UNLP

**Visibilidad**  
Al publicar o la visibilidad

**Preservación digital**  
Asigne el acceso a su obra a largo plazo, aun frente a los cambios tecnológicos

**Contacto**  
Contactese encuentren a la suya

El Servicio de Difusión de la Creación Intelectual es el Rep. La Plata, un servicio libre y gratuito creado para albergar, p. Unidades Académicas de la Universidad.

¿CÓMO CARGO MIS TRABAJOS?

**U.PORTO** REPOSITÓRIO ABERTO Percorrer: Ajustes

## Bem-vindo ao Repositório Aberto

O Repositório Aberto coleciona, preserva e disponibiliza a

## Smithsonian Libraries

Smithsonian Research Online

DSpace Principal

### DSpace Repository

DSpace is a digital service that collects, preserves, and distributes digital material. Repositories are important tools for preserving an organization's legacy, they facilitate digital preservation and scholarly communication.

### Comunidades en DSpace

Elija una comunidad para listar sus colecciones

- Atoll Research Bulletin [672]
- Center for Folklife and Cultural Heritage [52]
- CONTENT STAGING [1443]
- Cooper-Hewitt National Design Museum [1]

Investigaciones Científicas. investigación del CSIC.

2014 DSA 2017

## DIGITAL.CSIC

CIENCIA EN ABIERTO

### Noticias destacadas

**Preguntas frecuentes sobre el Módulo de Revisiones Abiertas para Repositorios** [31/09/2019]  
El proyecto europeo "Módulo de Revisiones Abiertas para Repositorios de Acceso Abierto" en que participa DIGITAL.CSIC ha preparado unas Preguntas Frecuentes para aclarar cuestiones relativas al funcionamiento del módulo y a su sistema de métricas y reputaciones. En esta fase piloto, las revisiones se realizan por invitación de los autores con trabajos disponibles en acceso abierto en DIGITAL.CSIC. Más información

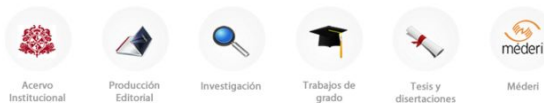
# Frontend web

Además de cumplir con los requerimientos visuales de la institución, la interfaz debe dar claramente acceso a:

- Navegación por el repositorio
- Acceso a funciones básicas (búsqueda, filtrado, autoarchivo, novedades)
- Priorización de contenidos o colecciones

Buscar en EdocUR

Introduzca el texto a buscar en EdocUR



Intellectum es el Repositorio Institucional de la Universidad de La Sabana, se ha creado para administrar, preservar y difundir la producción intelectual, científica, cultural e histórica de la comunidad universitaria.



**Libros**  
Libros digitalizados y e-books



**Red UNCI**  
Artículos y ponencias de la Red UNCI



**Radio Universidad**  
Entrevistas y producciones artísticas, entre otros audios



**Recursos educativos abiertos**  
Material educativo de cursos dictados en la UNLP



**Red de museos**  
Material perteneciente a la red de museos de la UNLP



**Emergencia hídrica**  
Trabajos dentro del Plan de Gestión Integrada de Riesgos de Desastres



## Frontend web - XMLUI

Está construido sobre un Framework Ad-hoc denominado Manakin.

Utiliza el framework Apache Cocoon

- múltiples componentes "simples"
- procesamiento en *pipeline*

Desarrollo maduro y utilizado en la mayoría de las instalaciones de DSpace.



## XMLUI - Web UI

Modularidad, extensibilidad y una organización multi-tier

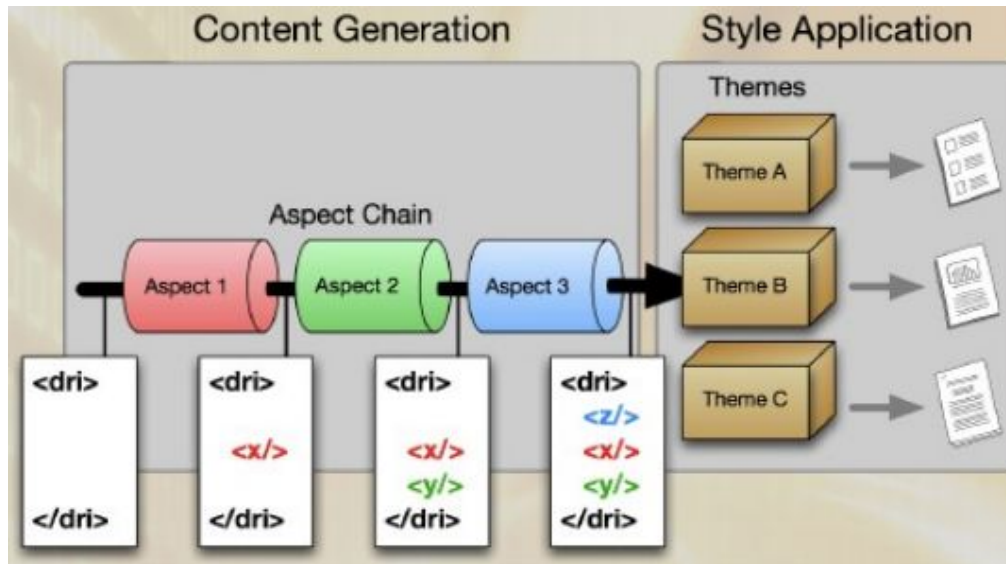
style	CSS y XHTML. Cambios menores sobre temas preexistentes
theme	XSLT, CSS y XHTML. Ajustes avanzados y nuevos temas
aspect	Cocoon, Java, XSLT. Definición de nuevas características, incorporación de contenido a los DRI (Digital Repository Interface)



## XMLUI - Aspect chain

### Aspectos del core de Manakin

- ViewArtifacts
- BrowseArtifacts
- SearchArtifacts
- E-Person
- Administrative
- Submission
- Statistics
- Workflow



Manakin request processing. Fuente: Phillips, S.; Green, C.; Maslov, A.; Mikeal, A. Leggett, J. "Manakin. A New Face for DSpace" D-Lib Magazine, November/December 2007. Volume 13 Number 11/12



## XMLUI - Archivo [xmlui]/sitemap.xmap

### sitemap.xmap:

- Conjunto de documentos XML que describen cómo los componentes de Cocoon se configuran.
- Divide en dos grandes secciones:
  - Definición de componentes, la cual describe cada tipo de componente
  - Sección del pipeline, que define como estos componentes se organizan.





## XMLUI - Internacionalización

XMLUI maneja archivos xml de catálogos de internacionalización.

Cada uno de estos contiene el texto que el usuario visualiza en un idioma determinado

El nombre de los archivos está basado en el lenguaje disponible.

Ejemplo:

*messages\_language\_country\_variant.xml*  
*messages\_language\_country.xml*  
*messages\_language.xml*  
*messages.xml*



## XMLUI - Internacionalización

La interfaz determina automáticamente qué archivo seleccionar basado en el navegador del usuario y la configuración del sistema.

Ejemplo: Si el navegador está configurado con el idioma “español argentina”, el sistema primero busca el archivo *messages\_es\_ar.xml*. De no encontrarlo buscará *messages\_es.xml* y finalmente si este último tampoco está disponible se utilizará *messages.xml*.

DSpace XMLUI tiene como idioma por defecto el inglés y su catálogo está ubicado en: `[dspace]/webapps/xmlui/i18n/messages.xml`

Para agregar nuevos catálogos, ubicarlos en:

`[dspace-source]/dspace/modules/xmlui/src/main/webapp/i18n/`



# XMLUI - Páginas estáticas

Una página estática es:

- un recurso que siempre se muestra de la misma manera, independientemente del contexto actual de la aplicación.
- Son archivos XHTML que contienen código invariable.

```
<?xml version="1.0" encoding="UTF-8" ?>
<article>
<h1>QUÉ ES CIC-DIGITAL?</h1>
El repositorio institucional de la Comisión de Investigaciones Científicas de la pr
Además de cumplir el rol de archivo digital institucional, el repositorio asiste a
En la mencionada resolución se establece que el repositorio dispondrá toda la produ
<ul>
<li>Informes de investigadores, personal de apoyo y becarios</li>
<li>Tesis de grado y posgrado</li>
<li>Artículos científicos y de divulgación</li>
<li>Publicaciones en congresos</li>
<li>Libros y capítulos de libros editados</li>
<li>Proyectos de investigación</li>
<li>Informes técnicos y Reportes</li>
<li>Patentes</li>
<li>Legislación institucional (Reglamentos, convocatorias, actas de directorio,
</ul>
Todos los materiales depositados en el repositorio serán preservados a largo plazo
</article>
```

**CICDIGITAL** | Repositorio Institucional  
Comisión de Investigaciones Científicas

Inicio Explorar Aportar Material Mas información Contacto ■ Mi cuenta ES

## ¿QUÉ ES CIC-DIGITAL?

El repositorio institucional de la Comisión de Investigaciones Científicas de la provincia de Buenos Aires, denominado **CIC-DIGITAL**, fue creado a fin de 2014 por la resolución 11467/14 con el objeto de reunir, registrar, divulgar, preservar y dar acceso público a toda la producción científico-tecnológica y académica de la institución. Además de cumplir el rol de archivo digital institucional, el repositorio asiste a los investigadores, personal de apoyo y becarios de la CIC para que puedan disponer en Internet su producción, maximizando su difusión e impacto a la vez que protege a largo plazo sus obras. En la mencionada resolución se establece que el repositorio dispondrá toda la producción científico-tecnológica generada en el ámbito de la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires, incluyendo:

- Informes de investigadores, personal de apoyo y becarios
- Tesis de grado y posgrado
- Artículos científicos y de divulgación
- Publicaciones en congresos
- Libros y capítulos de libros editados
- Proyectos de investigación
- Informes técnicos y Reportes
- Patentes
- Legislación institucional (Reglamentos, convocatorias, actas de directorio, resoluciones, acuerdos/convenios, planes anuales de Actividades, etc)

Todos los materiales depositados en el repositorio serán preservados a largo plazo y difundidos de forma libre y gratuita en Internet bajo el modelo de Acceso Abierto.

**SEDICI**

**Comisión de Investigaciones Científicas**  
Calle 526 entre 10 y 11  
CP: 1900 - La Plata - Buenos Aires - Argentina

**CIC** COMISIÓN DE INVESTIGACIONES CIENTÍFICAS

# Máquina virtual vagrant-dspace



## Máquina virtual vagrant-dspace

**vagrant-dspace** es una máquina virtual con un *Ubuntu 16.04* como SO, que se puede montar rápidamente (**vagrant up**) y se encarga de :

- instalar todas las **aplicaciones** necesarias por DSpace,
  - PostgreSQL
  - Tomcat
  - JRE
  - Maven
  - Ant
- y de configurar, compilar, instalar y “levantar” **DSpace**.



Sólo debe utilizarse para probar la aplicación (y no tener que lidiar con cuestiones técnicas de instalación). **Nunca se debería utilizar para un entorno de producción.**



## Máquina virtual vagrant-dspace - Requisitos

Para poder instalar vagrant-dspace es necesario tener instalado las siguientes dependencias:

- **VirtualBox** (en su última versión) - <https://www.virtualbox.org/>
- **Vagrant** (en su última versión) - <https://www.vagrantup.com/>

Debido a [problemas](#) que tiene Vagrant de comunicarse mediante SSH con la máquina virtual, instalar

- **GitHub Desktop** - <https://desktop.github.com/>

Es recomendado para los usuarios en Windows instalarse:

- **Git for Windows** - <https://git-for-windows.github.io/>
  - Provee un entorno de consola que simula una terminal BASH, usado para ejecutar comandos BASH desde la consola.



## Máquina virtual vagrant-dspace - Inicialización

1. Descargar y desempaquetar el paquete de vagrant-dspace
  - la versión oficial se encuentra en <https://github.com/Dspace/vagrant-dspace>
2. Abrir virtualbox
3. En una terminal, posicionarse en el directorio donde se desempaquetó vagrant-dspace.
4. Ejecutar el comando **'vagrant up'** para inicializar la máquina.
  - Si es la primera vez que se ejecuta, se instalará DSpace en la máquina virtual.
5. Si el comando anterior finaliza exitosamente, podremos acceder a la máquina virtual abriendo la siguiente URL en un browser:

**<http://localhost:8080/xmlui>**



## Máquina virtual vagrant-dspace - Eliminación

Si por alguna razón es necesario volver a instalar la máquina virtual, o simplemente queremos desinstalarla hay que hacer lo siguiente:

1. En una terminal, posicionarse en el directorio donde se desempaqueto vagrant-dspace.
2. Ejecutar el comando `'vagrant destroy'`.
3. Confirmar la eliminación apretando la tecla `'y'`.





Universidad Nacional  
de **Entre Ríos**



UNIVERSIDAD  
NACIONAL  
DE LA PLATA

**SEDICI**

REPOSITORIO INSTITUCIONAL DE LA UNLP



**PREBI**  
prebi.unlp.edu.ar

**Muchas gracias!**

Ariel Lira, Facundo Adorno, Pablo de Albuquerque  
{alira, facundo, pablo}@sedici.unlp.edu.ar

*Presentado el 27/09/2016 en el marco del convenio de colaboración entre la UNLP y UNER*