




Configuración de **DSpace**

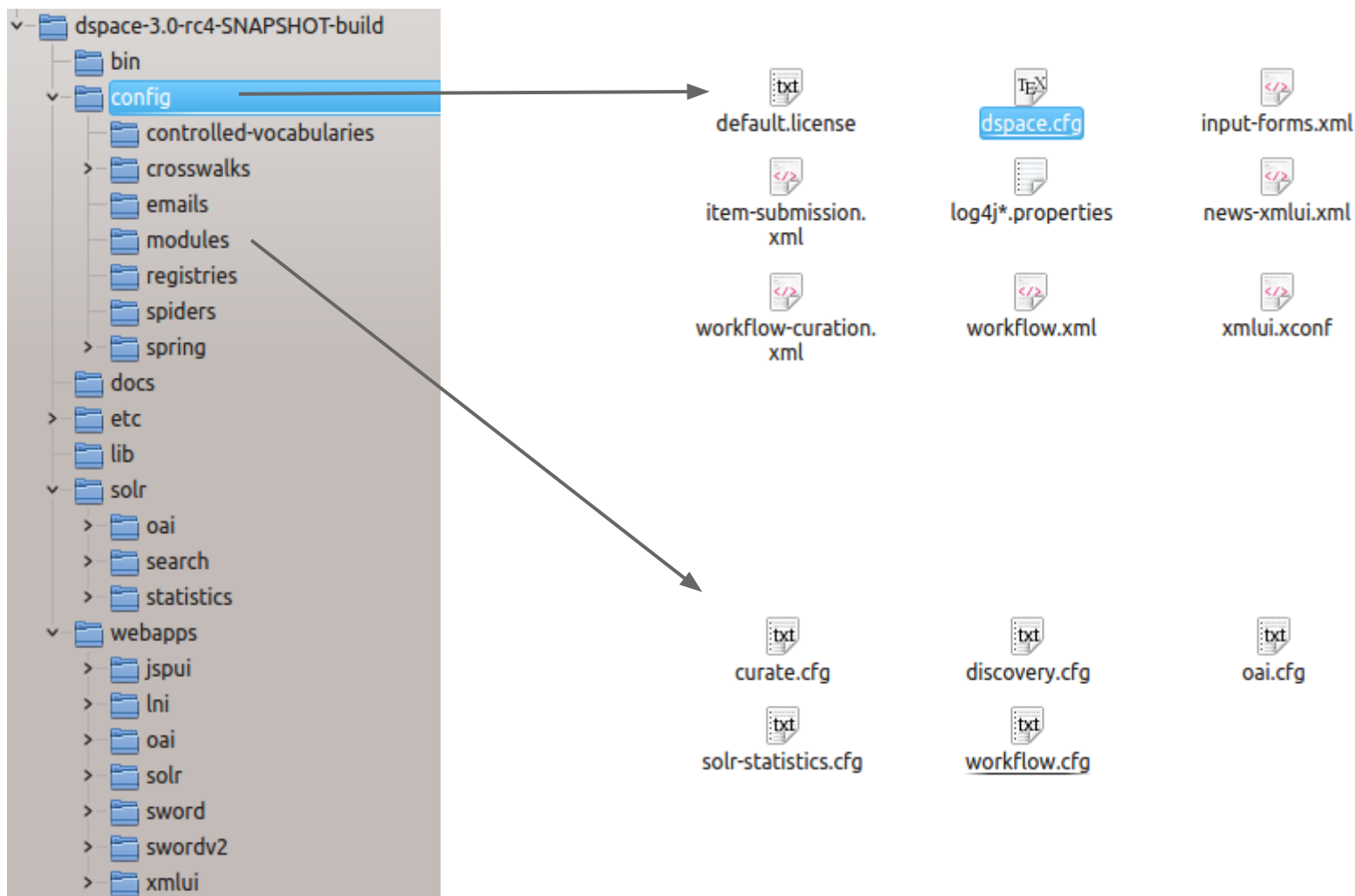


Bloque 3.1: contenido

- Distribución de configuración, sintaxis
- Configuraciones generales, BD
- Envío de correo, plantillas.
- Browse indexes, sort options
- Metadatos, procesos de carga y workflow de revisión.
- Licencias
- MediaFilters
- Embargo
- Internacionalización
- Handle



Estructura de directorios





Sintaxis de parámetros de configuración

- Java properties (*.properties, *.cfg)
 - property.name = word1 \${otra.propiedad}
 - property2.name = las propiedades pueden \ separarse en multiples lineas consecutivas \ separadas por barras invertidas
- Spring beans (spring/*)
- El resto usa sintaxis a medida
 - input-forms
 - item-submission
 - registries
 - crosswalks
 - controlled vocabularies
 - etc





Dónde cambiar la configuración? Alternativas

1. En `${dspace}/config/`
 - los cambios se aplican al reiniciar el contenedor
 - los cambios son más volátiles, pueden perderse
2. En `${dspace-src}/dspace/config/`
 - los cambios se aplican luego de ejecutar un "ant update%" y reiniciar el contenedor

El equipo de dspace recomienda:

- 1. Editar en `${dspace-src}/dspace/config/*`
- 2. Reempaquetar: `cd [dspace-source]/dspace/ && mvn package`
- 3. Actualización de configuraciones:
 - `cd target/dspace-<version>-build.dir/ && ant update_configs`
- Los archivos viejos quedan en `[dspace]/config/*.old`. Alternativamente, puede usarse `-Doverwrite=false` para que deje los archivos nuevos modificados en `[dspace]/config/*.new`





Configuración General

Archivo dspace.cfg



Configuración General y directorios

dspace.hostname	http://sedici.unlp.edu.ar	hostname del servidor que contiene las aplicaciones web
dspace.baseUrl	\${dspace.hostname}:puerto	hostname:puerto del servidor que contiene las aplicaciones web
dspace.url	\${dspace.baseUrl}/	Url completa, con path, a la aplicación principal del repositorio. /xmlui, /, /jspui, etc
dspace.oai.url	\${dspace.baseUrl}/oai	idem anterior, para para la webapp oai
dspace.name	SeDiCI - UNLP	

Es posible cambiar la ruta de casi todos los directorios de datos Dspace, los más importantes son:

dspace.dir	/var/dspace	Aloja por default gran parte del repositorio: aplicaciones, logs, recursos
assetstore.dir	\${dspace.dir}/assetstore	Contiene todos los archivos de los bitstreams
log.dir	\${dspace.dir}/log	Logs de aplicación: dspace, cocoon, checker, handle-server

Además, es posible cambiar directorios de reportes, exports, uploads, entre otros.



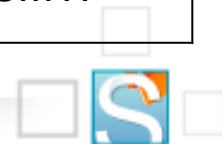
Base de Datos y proxy saliente


db.name	postgres	oracle o postgres
db.url	jdbc:postgresql://localhost:5432/dspace3	hostname:puerto del servidor de base de datos y nombre de base de datos
db.username	dspace	Url completa, con path, a la aplicacion principal del repositorio. /xmlui, /, /jspui, etc
db.password	dspace	idem anterior, para para la webapp oai
db.maxconnections	30	Máximo de conexiones activas en el pool
db.maxwait	5000	
db.maxidle	-1	Máximo de conexiones libres en el pool
http.proxy.host		
http.proxy.port		



Emails - SMTP

mail.server.disabled	false	Permite deshabilitar el envío de correos. (Para desarrollo)
mail.server	smtp.gmail.com	localhost es default
mail.server.username	dspace@sedici.unlp.edu.ar	Usuario de conexión, sólo si se usa autenticación.
mail.server.password	xyz	Password de conexión, idem anterior.
mail.server.port	465	idem anterior, para para la webapp oai
mail.from.address	dspace@sedici.unlp.edu.ar	Máximo de conexiones activas en el pool
mail.extraproperties	mail.smtp.socketFactory.port=465, \ mail.smtp.socketFactory.class=javax.net.ssl.SSLSocketFactory, \ mail.smtp.socketFactory.fallback=false	Indican que debe usar conexión segura en las comunicaciones SMTP





Emails (2) - Cuentas

mail.from.address	dspace@sedici.unlp.edu.ar	Máximo de conexiones activas en el pool
feedback.recipient	info@sedici.unlp.edu.ar	Destinatario del form de feedback de xmlui
registration.notify	info@sedici.unlp.edu.ar	Aviso de nuevos usuarios registrados
mail.admin	admin@sedici.unlp.edu.ar	webmaster email
alert.recipient	admin@sedici.unlp.edu.ar	Destinatario de mails de alertas y excepciones

Pruebas

- *./bin/dspace test-email*

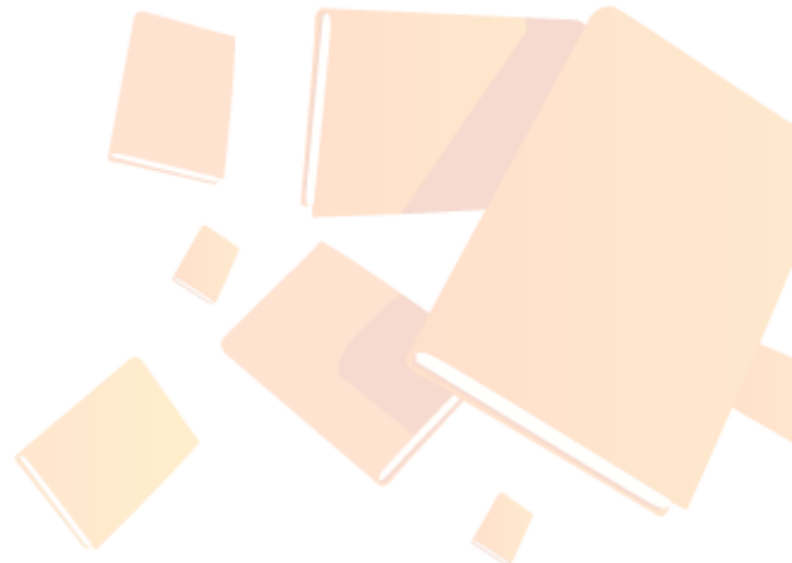


Emails - Uso de Plantillas

- 14 plantillas predefinidas (config/emails/*)
 - usuarios: change_password, **register**, **submit_archive**, submit_reject, subscription
 - gestores/administradores de colecciones: feedback, flowtask_notify, registration_notify, suggest, **submit_task**
 - Administrador: harvesting_error, internal_error, export_error, export_success
- Sintaxis
 - #Comentarios
 - Subject: asunto {}
 - texto con placeholders {}
 - parámetros posicionales y obligatorios



Exploración de recursos





Visualización de campos

- Ocultar campos

- `metadata.hide.schema.element.qualifier = true;`
- visible sólo para usuarios con rol ADMINISTRATOR
- ejemplos:
 - `dc.description.provenance`
 - datos de importación
 - datos privados de usuarios
- Funcionan en OAI, XMLUI y JSPUI





Browse

- Selección de Engine (Solr, Postgres, Oracle)
browseDAO.class = org.dspace.browse.SolrBrowseDAO
browseCreateDAO.class = org.dspace.browse.SolrBrowseCreateDAO
#SolrEngine requiere discovery habilitado
- Declaración de índices:
webui.browse.index.<n>
- Declaración de ordenamientos:
webui.itemlist.sort-option.
- Mostrar Frecuencia de valores / ocurrencias
webui.browse.metadata.show-freq.<n> = true | false
- Case sensitivity
webui.browse.metadata.case-insensitive = true



Browse - browse por items

webui.browse.index.<n>=

1. <index name>:**item**:<sort-option-name>:<sort-option>

<n>	int	Número de índice creciente. <u>No reutilizar n</u> . Es posible deshabilitarlos y rehabilitarlos.
<index name>	string	Código del índice que se usa para referenciar el índice desde los messages.
<sort-option-name>	código de sort-option	código de un sort-option declarado con webui.itemlist.sort-option.
<sort-option>	asc, desc	





Browse - Browse por metadatos o authorities

webui.browse.index.<n>=

1. <index-name>:**metadata**:<schema-prefix>.<element>.<qualifier>:<data-type field>:<sort-option>

<metadata>	metadata, metadataAuthority	Indica qué es lo que se explora
<schema prefix>	schema	
<element>	element	
<qualifier>	qualifier, *	un calificador o todos.
<datatype field>	date, title, text, authority	fecha, titulo con link , texto plano
<sort-option>	asc, desc	





Browse - Sort Options

webui.browse.sort-option.<n>

<option name>:<schema prefix>.<element>.<qualifier>:<datatype>:<sort-visibility>

<n>	int	Número de índice creciente. No es recomendable reutilizarlos. Es posible deshabilitarlos y rehabilitarlos luego.
<option name>	string	Código del sort que se usa para referenciar el índice desde los messages.
<schema prefix>	schema	
<element>	element	
<qualifier>	qualifier, *	indica si debe indexarse sólo los elementos de un calificador o todos.
<datatype>	date, title, text	Normalizacion a aplicar sobre el campo para : fecha, titulo, texto plano
<sort-visibility>	show, hide	Se debe mostrar o no, en la interface web

Ejemplos:

webui.itemlist.sort-option.1 = title:dc.title:title

webui.itemlist.sort-option.2 = dateissued:dc.date.issued:date

webui.itemlist.sort-option.3 = dateaccessioned:dc.date.accessioned:date





Browse - Ejemplos

- Solr Browse Engine
- case-insensitive
- índices mínimos: author, title, date issued y subjects
 - Dspace
 - webui.browse.index.1 = dateissued:metadata:dc.date.issued:date:full
 - webui.browse.index.2 = author:metadata:dc.contributor.*:text
 - webui.browse.index.3 = title:metadata:dc.title:title:full
 - webui.browse.index.4 = subject:metadata:dc.subject.*:text
 - SeDiCI
 - webui.browse.index.1 = dateissued:item:dateissued
 - webui.browse.index.2 = author:metadataAuthority:sedici.creator.person:authority
 - webui.browse.index.3 = title:item:title
 - webui.browse.index.4 = subject:metadataAuthority:sedici.subject.materias:authority



Metadatos, proceso de carga y workflow de revisión





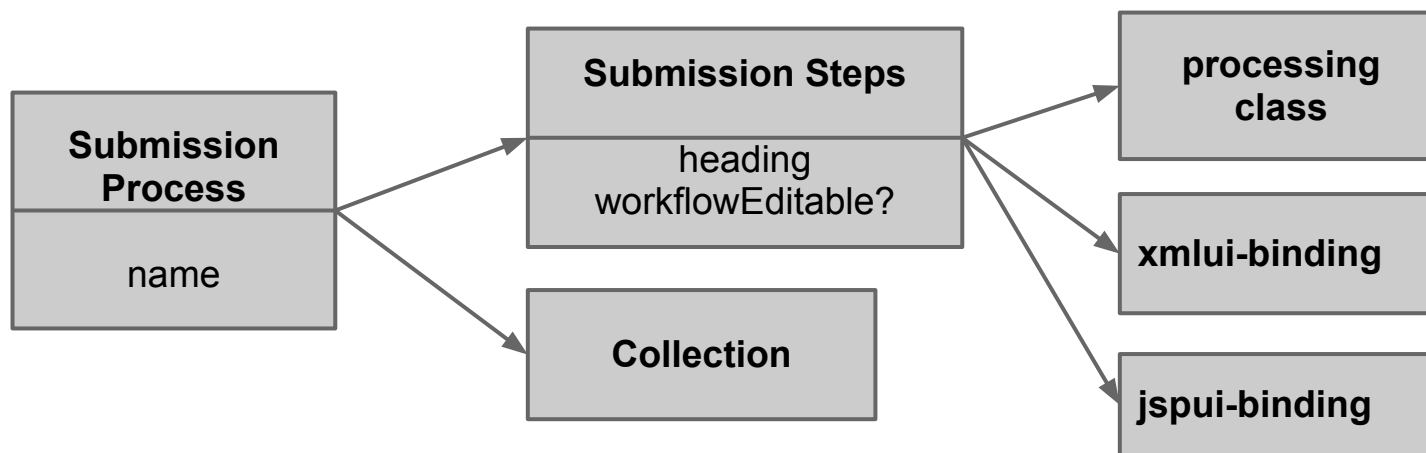
Definición de Formatos de Metadatos

- La creación de metadatos y schemas se hace desde las aplicaciones web
- **Los elementos del profile de metadatos no están directamente conectados con las configuración.**
- Al agregar o quitar metadatos, actualizar siempre
 - formularios de submission (input-forms.xml)
 - crosswalks: oai, google-scholar, package exporters/importers, sword, RSS
 - visualización de ítems
 - propiedades dependientes de campos:
 - Browse Indexes, facets,
 - Authorities plugins
 - etc



Item Submission

- Se definen submission-process o caminos de carga
 - secuencia de pasos
 - puede tener interfaz asociada
 - puede aplicarse durante el workflow
 - se asocia a una serie colecciones



El Submission Process "traditional" es:

Select Collection -> Initial Questions -> Describe ->
Upload -> Verify -> License -> Complete

<https://wiki.duraspace.org/display/DSDOC3x/Submission+User+Interface> –





Item Submission - Configuración

item-submission.xml

```
<item-submission>
  <submission-map>
    <name-map collection-handle="default" submission-name="traditional" />
    ...
  </submission-map>
  <step-definitions>
    <step id="collection">
      <processing-class>org.dspace.submit.step.SelectCollectionStep</processing-class>
      <workflow-editable>>false</workflow-editable>
    </step>
    ...
  </step-definitions>
  <submission-definitions>
    <submission-process name="traditional">
      <step id="collection" />
      ... <!-- Mas steps -->
    </submission-process>
    ... <!-- Mas submission-process -->
  </submission-definitions>
</item-submission>
```





Submission - DescribeStep (1)

- conjunto de formularios
 - posee una o más páginas
 - con indicaciones de cuáles metadatos deben mostrarse, cómo y si son obligatorios.
 - se asocian a una, varias o todas las colecciones



Submission - DescribeStep (2)

input-forms.xml

```
<input-forms>
  <form-map>
    <name-map collection-handle="default" form-name="traditional" />
  </form-map>
  <form-definitions>
    <form name="traditional">
      <page number="1">
        <field>
          </field>
          ...
        </page>
        ...
      </form>
      ...
    </form-definitions>

    <field>
      <dc-schema>SCHEMA</dc-schema>
      <dc-element>ELEMENT</dc-element>
      [<dc-qualifier>QUALIFIER</dc-qualifier>]
      <label>LABEL</label>
      <input-type [value-pairs-name="common_types"]>FIELD_TYPE</input-type>
      <hint>HINT</hint>
      <required>REQUIRED MESSAGE</required>
      [<visibility [otherwise="readonly"]>(workflow|submit)</visibility>]
      <type-bind>CSV_DC_TYPE_VALUES</type-bind>
    </field>

  <form-value-pairs>
    <value-pairs value-pairs-name="common_types" dc-term="type">
      <pair>
        <displayed-value>Article</displayed-value>
        <stored-value>Article</stored-value>
      </pair>
      ...
    </value-pairs>
    ...
  </form-value-pairs>
  ...
</input-forms>
```





Submission - DescribeStep (3)

- Tipos de campos (input-types)
 - onebox: caja de texto simple
 - twobox: par de onebox, es para facilitar la carga de campos repetibles
 - textarea: texto largo
 - name: 2 inputs, uno para apellido y otro para nombre. Se guardan como 'Apellido, Nombre'
 - date: Fecha. En campos requeridos, obliga a poner al menos el año
 - series: 2 cajas de texto, una para título de la serie y otro para número. Se guardan juntos, con un ';'entre si
 - dropdown: selector que despliega un listado de value-pairs
 - qualdrop_value: caja doble con selector de calificador (en base a un value-pairs) y una caja de texto.
 - list: conjunto de checkboxes (campos repetibles) o radio buttons (campos no repetibles) . Los valores se toman de un value-pairs





Submission - Upload Step

Subida de archivos durante el proceso de envío

- Obligatoriedad: `webui.submit.upload.required`
- Alternativa, permisos por Grupo
 - a. definir nueva propiedad
 - `webui.submit.upload.excluded_group = SeDiCIAdmin`
 - b. Extender `org.dspace.submit.step.UploadStep` para que valide por el grupo del usuario actual
 - c. cambiar el step de upload en el `<submission-process>` de `item-submission.xml`, y usar uno nuevo con `<processing-class>` que apunte a la nueva clase



Otros módulos fundamentales





Licencias

- De depósito y/o distribución
 - Default: {dspace.dir}/config/default.license
 - Cada colección puede tener su propia licencia
- De uso (Creative Commons)

cc.api.rooturl	http://api.creativecommons.org/rest/1.5	Rest Service URL. No cambia
cc.license.uri	dc.rights.uri	Campo que guarda la uri de la licencia
cc.license.name	dc.rights	Campo que guarda el nombre de la licencia
cc.submit.setname	true	Indica si se guarda el campo cc.license.name
cc.submit.addbitstream	true	Indica si se guarda el bitstream con el rdf de la licencia CC
cc.license.classfilter	mark, recomb, zero	http://api.creativecommons.org/rest/1.5/classes
cc.license.jurisdiction	ar	Jurisdicción de la licencia o vacío
cc.license.jurisdiction.description	Argentina	Descripción (o nada)



MediaFilters (1)

- **Uso**

- Extracción automática de texto y miniaturas
 - PDF, HTML, PowerPoint, Word
- Transformación automática de formatos

- **Configuración**

- **Declaración (con alias)**

```
plugin.named.org.dspace.app.mediafilter.FormatFilter = \  
    filterFQCN_1 = FilterAlias_1, ..., filterFQCN_n = FilterAlias_n
```

- **Habilitación (usando el alias):**


```
filter.plugins = FilterAlias_1, FilterAlias_2, ... , FilterAlias_n
```

- **Asociación con formatos (*config/registries/bitstream-formats.xml*)**

```
filter.filterFQCN_i.inputFormats = format1, ..., formatM, donde
```

Ejemplo: filter.org.dspace.app.mediafilter.JPEGFilter.inputFormats = BMP, GIF, JPEG, image/png





MediaFilters (2) - Extracción de texto completo de Archivos PDF

- PDFFilter
 - Preconfigurada con Dspace
 - pdf grandes, usando archivos temporales durante la extracción (más lento)
 - `pdffilter.largepdfs = true`
 - Tolera OOM durante las extracciones y las ignora
 - `pdffilter.skiponmemoryexception = true`
 - Posee muchos bugs reportados sin resolver
- XPDF
 - Licencia no compatible con Duraspace
 - Soporta archivos muy grandes
 - Mejor uso de memoria





Embargo

- Traba temporal sobre un item o bitstream
- La fecha de fin de embargo (lift date) se calcula en base a un período (terms)
- Lo aplica el EmbargoSetter en el InstallItem
- Lo levanta el EmbargoLifter al ejecutar la tarea *embargo-lifter*

embargo.field.terms	sedici.embargo.period	condiciones del embargo. Típicamente es un periodo temporal
embargo.field.lift	sedici.embargo.liftDate	metadato que mantiene la fecha de fin de embargo
embargo.terms.open	forever	Código que representa el embargo ilimitado
#embargo.terms.days	90:90,1 year:365,2 years:730	terms disponibles
plugin.single.org.dspace.embargo.EmbargoSetter	org.dspace.embargo.DefaultEmbargoSetter	Realiza el cálculo de fin de embargo
plugin.single.org.dspace.embargo.EmbargoLifter	org.dspace.embargo.DefaultEmbargoLifter	Levanta el embargo





Advanced Embargo

- Selección de embargo durante el proceso de submission (item-submission.xml)
 - Embargo a nivel de Item --> habilitar AccessStep
 - Embargo a nivel de bitstream --> habilitar UploadWithEmbargoStep
 - 2 tipos de formularios
 - xmlui.submission.restrictstep.enableAdvancedForm*
 - Simple: se permite indicar si el objeto es privado, fecha de fin embargo y causa
 - Advanced: se permite gestionar las políticas completas sobre el item para cualquier grupo
- Hacer un item privado/público desde el Edit





Internacionalización

- **Configuración Global**
 - `dspace.cfg>default.locale = es`
 - `dspace.cfg>webui.supported.locales = es,en,pt`
- **Licencia**
 - `config/default_LOCALE.license`
- **Plantillas de Correo**
 - `emails/plantilla_LOCALE`





Handle Server

handle.canonical.prefix	http://hdl.handle.net/	Prefijo para generar dc.identifier.uri. Se usa: <ul style="list-style-type: none">• <code>\${dspace.uri}/handle/</code>• <code>http://hdl.handle.net/</code> Visibilidad e identidad vs persistencia.
handle.prefix	123456789	CNRI Handle prefix. Se puede mantener el prefijo 123456789, aunque <u>no es conveniente</u> .
handle.dir	<code>\${dspace.dir}/handle-server</code>	Ubicación de archivos del servidor handle: logs, claves, etc

Muchas gracias

Dudas y comentarios?

alira@sedici.unlp.edu.ar

nestor@sedici.unlp.edu.ar

marisa.degiusti@sedici.unlp.edu.ar