



La selección de metadatos para repositorios mixtos (datos y literatura) y su apoyo a la ciencia abierta

Minerva Castro Escamilla

Biblioteca Conjunta de Ciencias de la Tierra - UNAM

Miguel Ángel Sosa Jiménez

Servicio Sismológico Nacional, Instituto de Geofísica, UNAM

Presentación

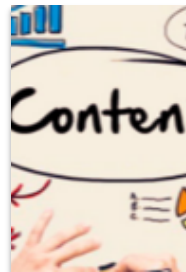
- Los repositorios son herramientas de difusión y consulta tanto de documentos académicos y datos de y para la investigación, que apoyan el desarrollo de la ciencia y contribuyen al proyecto mundial sobre ciencia abierta.



Presentación



Deben discutirse normativas, técnicas y elementos que definirán el registro y la especificidad de cada uno de los objetos digitales albergados.



Indispensables para que los contenidos sean recuperables mediante herramientas de búsqueda.



Deberán estar bajo parámetros internacionales que brinden la posibilidad no solo de realizar una organización técnica, sino además alternativas para el intercambio de registros y datos entre sus pares y la unificación de contenidos.



Variantes para los tipos de información, datos y textos completos que se maneja en los repositorios.

La selección



CIENCIA ABIERTA

Proporciona acceso a texto completo a diversos recursos de información académica, datos científicos, es decir, sin requerimientos de suscripción, registro o pago.

Metadatos



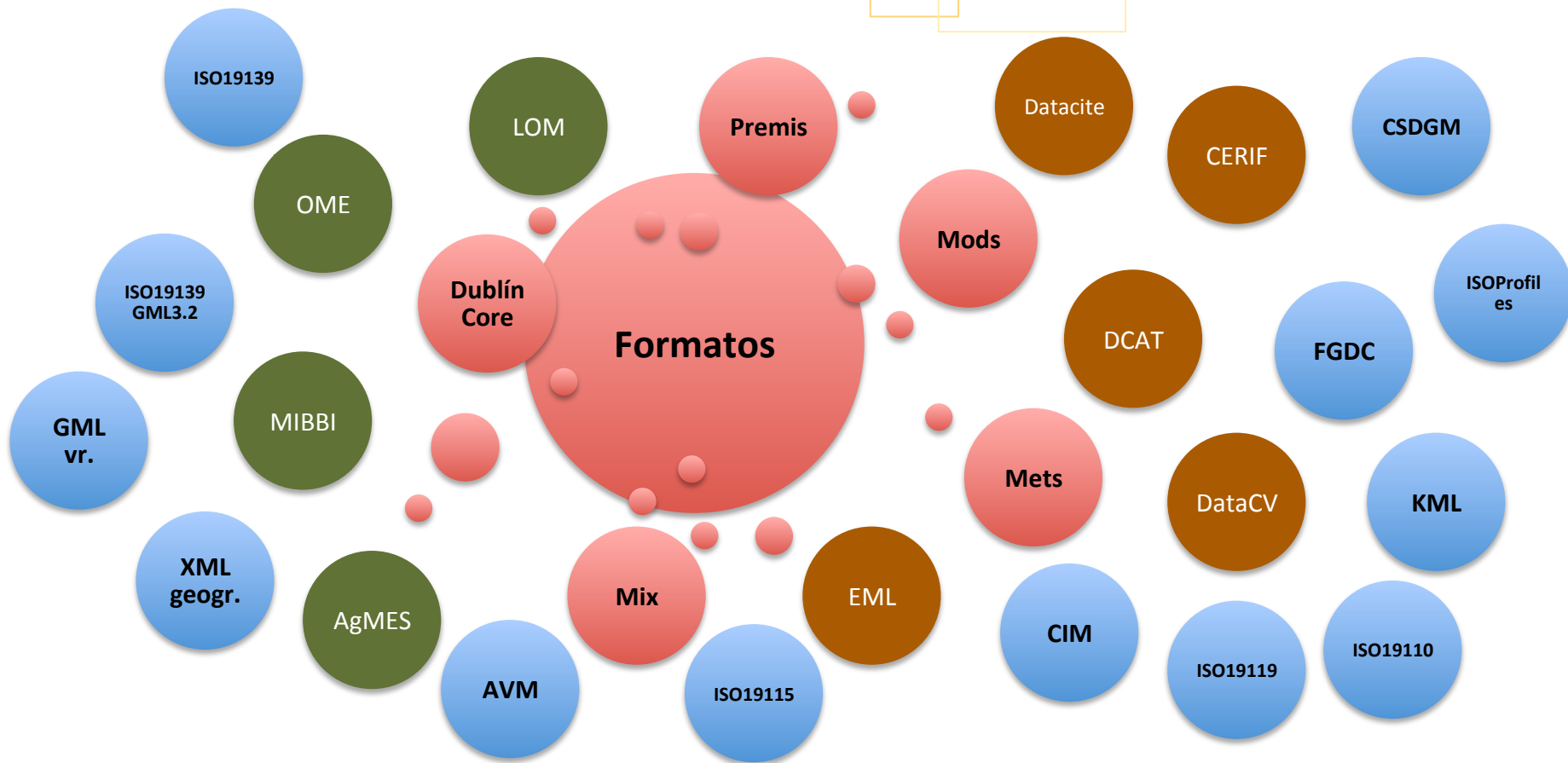
Funciones

- Facilitan la organización de los recursos basados en la audiencia.
- Se pueden organizar mediante una identificación digital (URL, PURL o un DOI) y estructura
- Aseguran su preservación.
- Expresados en XML, lenguaje independiente del software e intercambiable entre distintos sistemas
- Se expresa a partir de una Definición de Tipos de Documentos (DTD), especificando las etiquetas que se emplearán.
- Deberán mantener su capacidad de interpretación y utilidad para que los registros puedan almacenarse y conservarse a lo largo del tiempo.

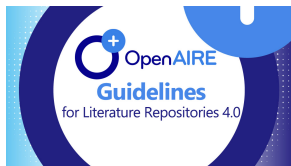
Tipos



Formato de Metadatos

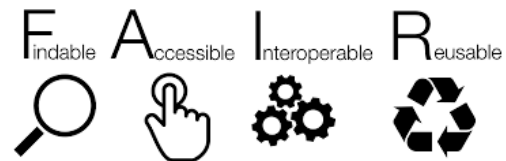
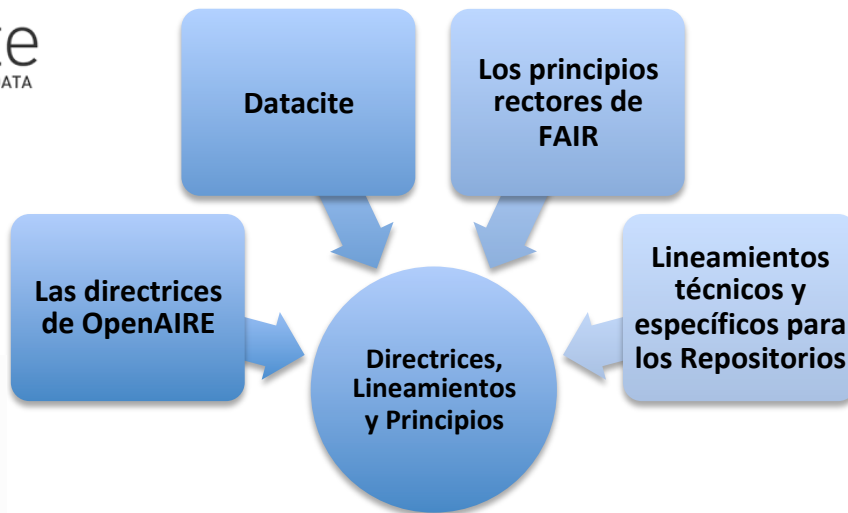


Directrices, Lineamientos y Principios



OpenAIRE Guidelines for Data Archives

- Introduction
- Use of OAI-PMH
- Use of DataCite
- Application Profile Overview



Esquema de metadatos en DataCite: cambios que incluyen en sus versiones

30 de marzo de 2021	16 de agosto de 2019	20 de marzo de 2019	23 de octubre de 2017	19 de septiembre de 2016	16 de octubre de 2014
Versión 4.4.	Versión 4.3.	Versión 4.2.	Versión 4.1.	Versión 4.	Versión 3.1.

Publicado el 24 de julio de 2013.



Características a considerar



Características	Descripción
Estructura de datos	<ul style="list-style-type: none">-Material digital que constituye el recurso documento-Información sobre los datos metadatos-Identificador único y persistente del objeto digital
Origen	<ul style="list-style-type: none">-Creados en forma digital (born digital)-Digitalizados
Número de documentos	<ul style="list-style-type: none">-Sencillos: un solo documento-Compuestos: paquetes de documentos
Almacenamiento	<ul style="list-style-type: none">-Local (servidor propio)-Externo (nube)
Acceso	<ul style="list-style-type: none">-Público-Forma restringida-Contenidos descritos a través de estándares de metadatos para que puedan ser reutilizados por otros sistemas
Formato	<ul style="list-style-type: none">-Texto, imagen, sonido, multimedia, datos
Usos	<ul style="list-style-type: none">-Establecer una comunicación y colaboración académica-Publicación digital de los resultados académicos y de investigación

Metadatos

Dublín Core



Elementos se clasifican en 3 grupos, los cuales indican la clase o el ámbito de la información

LITERATURA

Datacite

Mandatory	Recommended	Optional
Identifier	Subject	Language
Creator	Contributor	Alternate ID
Title	Date	Size
Publisher	Resource Type	Format
Publication year	Related identifier	Version
	Description	Rights
	GeoLocation	

Niveles obligatorios para las propiedades de metadatos DataCite

DATOS



Metadatos en los Datos



Descripción de los datos,
formato y versión
(metodologías)

Cómo se adquieren y
procesan (recolección)

Relación con otros datos

Metadatos para que los
datos posean sentido y
guarden su contexto

Asignación de temas
especializados y palabras
clave

Emplear URIs con otras
fuentes de datos y
aplicaciones, facilitando la
identificación

Políticas para compartir,
acceder y reusar los datos

Licencia abierta para
acceder libremente a la
información

Citación

Responsabilidad y roles

Derechos de autor

Almacenamiento

Los datos deberán estar
estructurados en un
formato legible por
máquina

Se debe depositar en
formatos no propietarios,
es decir, datos abiertos,
estructurados y sin
propietarios y con
nombres



geofísica
UNAM

BCCT
BIBLIOTECA
CONJUNTA DE
CIENCIAS DE LA
TIERRA

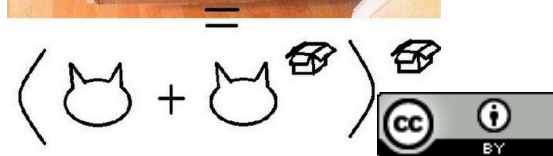
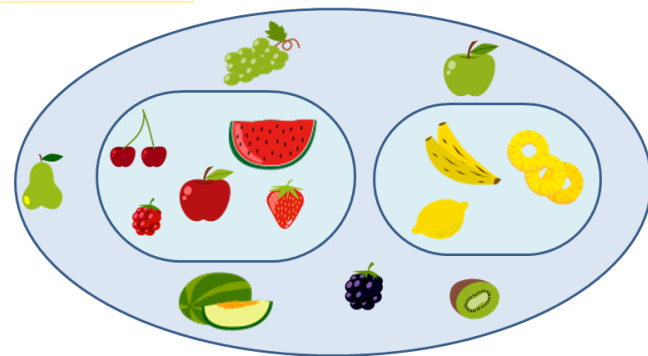
Interoperabilidad

Interoperabilidad de un RI



La interoperabilidad de todo RI institucional basado en la plataforma DSpace, depende de tres elementos fundamentales.

- (Esquema de metadatos): Formato en que deberá ser entregado el conjunto de metadatos: dc, datacite, ore, mets, qdc, etc.
- Filtro: Permite seleccionar la información que debería estar disponible en cada contexto (Datos, Literatura)
- Transformador: Transformar los metadatos solicitados para ser desplegados correctamente en la interfaz de OAI-PMH.
- Contexto: Se puede interpretar como una interfaz virtual de OAI-PMH Server.



¿Cómo puedo personalizar un contexto en Dspace?



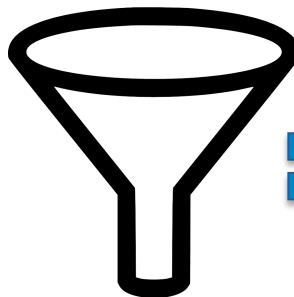
- Consultar la [documentación oficial de Dspace](#). (V6.2)
- Definición de las esquemas de metadatos
[dspace]/config/crosswalks/oai/metadaFormats (XSL)
- Definición de los transformadores de metadatos
[dspace]/config/crosswalks/oai/transformers (XSL)
- Configuración de los filtros, tranformadores y contextos
[dspace]/config/crosswalks/oai/xoai.xml (XML)



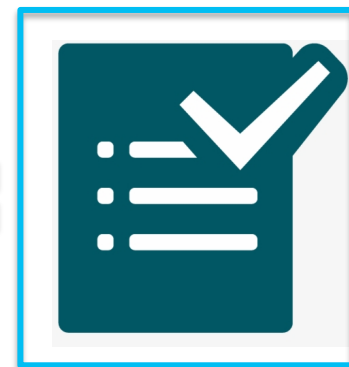
Esquema de metadatos



Transformador



Filtro

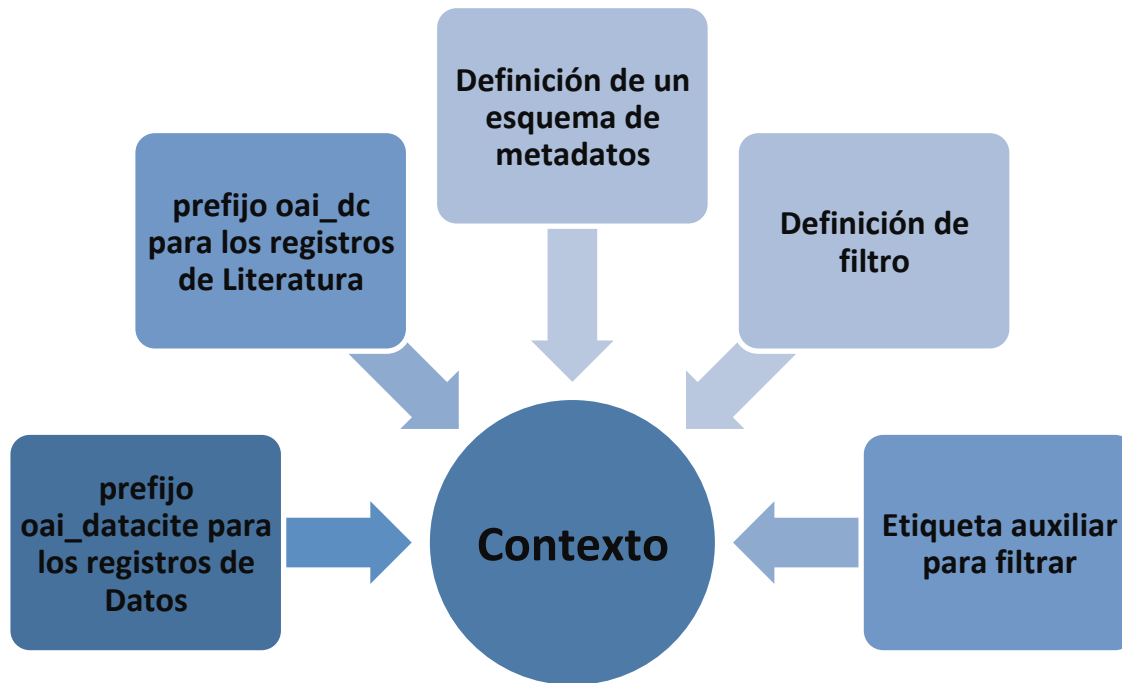


Interfaz OAI-PMH

Contexto

Definición del XML

[dspace]/config/crosswalks/oai/xoai.xml



Definición del XML

[dspace]/config/crosswalks/oai/xoai.xml



Definición de un nuevo contexto utilizando el prefijo **oai_datacite** para los registros de Datos.

```
<!--Contexto para Datos-->
<Context baseurl="datos" name="Contexto de Datos.">
  <Filter ref="DataFilter"/>
  <Format ref="oai_datacite"/>
  <Description>
    Contexto de Datos del RDSSN.
  </Description>
</Context>
```

Definición del esquema de metadatos.

```
<Format id="oai_datacite">
  <Prefix>oai_datacite</Prefix>
  <XSLT>metadataFormats/oai_datacite.xsl</XSLT>
  <Namespace>http://schema.datacite.org/oai/oai-1.1/</Namespace>
  <SchemaLocation>http://schema.datacite.org/oai/oai-1.1/oai.xsd</SchemaLocation>
</Format>
```



Definición del XML

[dspace]/config/crosswalks/oai/xoai.xml



Definición del filtro de Datos

```
<Filter id="DataFilter">
  <Definition>
    <Custom ref="ExcluyeDatos"/>
  </Definition>
</Filter>
```

Con ayuda de una etiqueta auxiliar, se filtran todos aquellos elementos que contengan que cumplan con la siguiente regla.

```
<!-- Para filtrar los elementos de Datos. -->
<CustomCondition id="ExcluyeDatos">
  <Class>org.dspace.xoai.filter.DSpaceAtLeastOneMetadataFilter</Class>
  <Configuration>
    <string name="field">dcterms.format</string>
    <string name="operator">contains</string>
    <string name="value">Datos</string>
  </Configuration>
</CustomCondition>
```





Gracias por su atención

Contacto:

minerva@igeofisica.unam.mx

sosam@sismologico.unam.mx