Cambios en las Directrices DRIVER para proveedores de contenidos

Juan Corrales
Jordi Prats
Proyecto Recolecta

12 de diciembre de 2008

Directrices de Driver

- Orientan sobre buenas prácticas en políticas y estructuración de los repositorios.
- Gracias a estas buenas prácticas se pueden construir servicios avanzados sobre los documentos.
- Por ahora sólo sirven para documentos textuales.

¿Cómo cumplir las directrices de Driver?

- Leer las directrices: http://www.driver-support.eu/managers.html
- Usar el validador: http://validator.driver.research-infrastructures.eu/validatorWeb/
- Utilizar el portal de soporte: http://www.driver-support.eu/

Proveedor Driver validado. Recursos textuales

Obligatorio

- El repositorio contiene recursos textuales digitales
- Los recursos textuales están en formatos ampliamente utilizados y extendidos (PDF, TXT, RTF, DOC, TeX, etc.).
- Los recursos textuales son de acceso abierto, están disponibles directamente en el repositorio para cualquier usuario del mundo, y sin ninguna restricción (autorización, pago).
- Los recursos textuales están descritos mediante registros de metadatos.
- Los recursos textuales y de metadatos se vinculan entre sí de tal modo, que un usuario final pueda acceder al recurso textual mediante el identificador (normalmente una URL) del registro de metadatos.
- La URL de un recurso codificada en el registro de metadatos siempre se puede localizar y nunca se cambia ni se reasigna.
- Un identificador único identifica el registro de metadatos y el recurso textual (no hay punteros a sistemas externos, como un sistema bibliotecario nacional o un editor).

Proveedor Driver validado. Recursos textuales

Recomendado

- Verificación transparente de la integridad de un recurso textual.
- Medidas de control de calidad (del contenido científico) de los recursos textuales expuestos para limitarlos a, por ejemplo, los recursos textuales incluidos en el informe científico anual (o equivalente).
- La URL de un recurso codificada en el registro de metadatos se basa en un esquema de identificadores persistentes (como DOI, URN, ARK, etc.).

Proveedor Driver validado. Metadatos

Obligatorio

- Los metadatos se estructuran según la norma Unqualified Dublin Core (ISO 15836:2003).
- Los elementos individuales de DC se utilizan según lo dispuesto en este documento
- Es obligatorio el uso de la codificación UNICODE.
- Los valores de las etiquetas DC no deben contener marcado HTML o XML. Puden contener marcado LaTeX.

.

Recomendado

- Los metadatos se estructuran según esquemas más completos, como Dublin Core Qualified o MODS.
- Las directrices DRIVER sobre metadatos solo hacen referencia al formato de intercambio de datos.
- El idioma recomendado para los resúmenes (los resúmenes son optativos) del artículo es el inglés.

Elementos DC

Etiqueta D.C	Versión 1	Version 2
Title	Obligatorio	Obligatorio
Creator	Obligatorio	Obligatorio
Subject	Obligatorio si procede	Obligatorio si procede
Description	Obligatorio si procede	Obligatorio si procede
Publisher	Obligatorio si procede	Recomendado
Contributor	Optativo	Optativo
Date	Obligatorio	Obligatorio
Type	Obligatorio	Obligatorio
Format	Recomendado	Recomendado
Identifier	Obligatorio	Obligatorio
Source	Optativo	Optativo
Language	Recomendado	Recomendado
Relation	Optativo	Optativo
Coverage	Optativo	Optativo
Rights	Recomendado	Recomendado
Audience	Optativo	Optativo

Elemento Type

- Tres posibles usos:
  - Obligatorio: tipo de publicación a partir de un vocabulario controlado
  - Optativo: tipo de publicación a partir de texto libre
  - Recomendado: versión del documento a partir de un vocabulario controlado

Elemento Type

Type DRIVER v 1	Tiype Driver V 2
Article	article
Book	book
Conference lecture	conferenceObject
Conference report	conferenceObject
Contribution for newspaper or weekly	contributionPeriodical
Doctoral thesis	doctoralThesis
Master thesis	masterThesis
Bachelor thesis	bacheloThesis
External research report	Report
Lecture	Lecture
Internal report	Report
Newsletter	contributionPeriodical
Part of book or chapter of book	bookPart
Research paper	workingPaper
	review
	Patent
	annotation
	other

Elemento Type

Versión
Draft
submittedVersion
acceptedVersion
publisherVersion
updatedVersion

Cambios en las directrices OAI-PMH

- Mejores explicaciones en los apartados:
  - Registros sin texto completo
  - Repositorios de tesis
  - Estrategia de borrado de documentos del repositorio
  - Vida del “Resumption Token”
- Aumento del número máximo de elementos de recolección de 200 a 500 registros.

didl

- Wrapper para otros esquemas de metadatos. Describe la naturaleza de las diferentes partes de un registro: metadatos, objetos y páginas de salto.

```
<metadata>
  <didl:DIDL ...>
    <DIDLDocumentId="Identificador del objeto didl">
    <didl:Item>
      <didl:Descriptor>Identificador del objeto (DOI)</didl:Descriptor>
      <didl:Descriptor>Fecha de modificación del objeto</didl:Descriptor>
      <didl:Item>...</didl:Item> <!-- Metadatos (Tipo descriptiveMetadata. 1 o más veces. la
 primera en DC) -->
      <didl:Item>...</didl:Item> <!-- Objetos (Tipo objectFile.  1 o más veces)-->
      <didl:Item>...</didl:Item> <!-- Página de salto (Tipo humanStartPage. 0 o 1 vez -->
    </didl:Item>
  </didl:DIDL>
</metadata>
```

Cambios en las directrices didl

- Arreglo de bugs
- Uso del espacio de nombres eu-repo para definir los tipos de items

Conclusiones

- Directrices orientadas a la transformación de otros esquemas utilizados.
- Buenas prácticas: mapeo entre otros modelos y directrices Driver.
- Implementación de modelos ya desarrollados: Using simple Dublin Core to describe Eprints.
- Ampliación de los tipos de contenidos a información que puede enriquecer los objetos alojados.
- Importancia de identificar los esquemas de codificación utilizados.

Conclusiones

- Más orientada a los servicios de valor añadido (Estandarización de las materias, nombres de autor).
- Gana mucha importancia el uso de espacios de nombres y vocabularios definidos.
- Dan información sobre versiones preliminares de futuras estandarizaciones (Identificador de autor, materias, estadísticas, mods).
- Algún cambio complejo: Elementos date y creator.