

# DRIVER

## Directrices para el Repositorio Institucional del SSPA.

Pilar Toro Sánchez-Blanco

III Jornadas de Bibliotecas del SSPA  
Ronda, a 26 de mayo de 2009

# DRIVER

DRIVER, Digital Repository Infrastructure Vision for European Research, es un proyecto financiado por la Unión Europea dentro del 7º Programa Marco, a fin de crear una **infraestructura de servicios y de datos de investigación** a partir de la agregación de los contenidos de los repositorios abiertos europeos.

DRIVER en el ámbito de la documentación científica, viene a ejercer un papel similar al que hace Europeana en relación al patrimonio documental europeo.

Si el objetivo de Europeana es convertirse en el gran portal común de los países de la Unión Europea en el que se muestra su patrimonio documental y cultural digitalizado, DRIVER también quiere convertirse en un portal para mostrar de forma centralizada la **producción científica europea de acceso abierto.**



# Objetivos de la iniciativa DRIVER

**Recolectar** los contenidos digitales de los repositorios abiertos de investigación europeos que se adhieran a su red, posibilitando así la creación de servicios globales de búsqueda y localización de los contenidos:

- 3 aspectos fundamentales:
  - Es de **acceso abierto**: DRIVER tiene el objetivo explícito de hacer más accesibles los resultados de la investigación.
  - **Basado en estándares ya existentes**: DRIVER se basa en protocolo OAI-PMH para recolectar recursos almacenados en repositorios abiertos distribuidos; estos repositorios han de tener descritos sus recursos con metadatos Dublin Core y permitir ser legibles vía OAI-PMH.
  - **De alcance europeo**: DRIVER quiere aportar un nivel superior de visibilidad y accesibilidad a los fondos contenidos en repositorios abiertos de documentación científica de ámbito institucional, de un grupo o consorcio de entidades, o nacional: el de la Unión Europea.



# INTEROPERABILIDAD

Para un entendimiento entre repositorios es necesario que hablen un mismo lenguaje y por tanto es necesario crear un terreno común.

La estandarización de la construcción de los servicios locales de información permiten utilizar de manera especialmente eficiente las inversiones de cada institución en infraestructura, con una alta utilidad adicional.

En las Directrices DRIVER se intenta alcanzar la interoperabilidad en dos niveles:

**Sintáctico:** uso de **OAI-PMH** y uso de **OAI\_DC**

**Semántico:** Uso de **vocabularios**

DRIVER utiliza el **protocolo OAI-PMH (Open Archives Initiative-Protocolo for Metadata Harvesting)** para recolectar los contenidos digitales de los repositorios abiertos europeos que se adhieran\* a su red, posibilitando la creación de servicios globales de búsqueda y localización de los contenidos - orientados al usuario final.

\*Carácter voluntario de la adhesión por parte de repositorios



Desde punto de vista de usuarios, DRIVER es un **metabuscador**

Desde punto de vista técnico, es un **recolector de recursos digitales**

- Los servidores centrales de DRIVER recolectan los metadatos de los repositorios asociados, pero los documentos a texto completo no son copias sino que se mantienen en los repositorios originales.

# Elaboración de directrices de DRIVER

A partir de experiencia práctica y de otras directrices existentes a nivel internacional:

- **Directrices de DARE** (estructura de red distribuida de proveedores de contenido operativas de Países Bajos). Lo toma como base pero como un único documento.
  - Contiene 2 secciones fundamentales:
    - Uso de DRIVER de Dublin Core
    - Uso en DRIVER de las directrices de OAI-PMH



Requisito DARE: utilidad de un 'contenedor XML' para cada recurso:

- permite recolección de recursos con OAI-PMH
- proporciona un vínculo inequívoco con el recurso (no mediante una página de acceso)
- admite indexado de texto completo
- permite representar documentos complejos compuestos de varios pdf.

- **Certificado DINI** (Iniciativa alemana para la información en red) para los servicios de documentación y publicación de las universidades alemanas.

El DINI se fundó para fomentar, en las universidades alemanas regionalmente o más allá de ellas, la mejora de los servicios de Información y comunicación, así como para el desarrollo de la infraestructura informativa necesaria para ello.

El certificado está diseñado como guía de operación local general de un repositorio con unos criterios mínimos, cuyo cumplimiento es indispensable para la conservación de una moderna comunicación científica, y unas recomendaciones adicionales para exigencias deseables.

Las directrices DRIVER no cubren aspectos descritos en el certificado DINI, porque DRIVER estudia los repositorios desde el punto de vista de un agregador, pero sí tiene en cuenta sus criterios para poder trabajar con ellos.



# DRIVER

## Desarrolla

- una infraestructura de **servicios** (para el usuario final)
- una infraestructura de **datos**:  
sustentada sobre los recursos alojados localmente, como publicaciones científicas recopiladas en repositorios digitales de instituciones y organismos de investigación

# Objetivos de las Directrices

- DRIVER 1:
  - Recuperación de texto completo con datos bibliográficos: paso fundamental y necesario para conseguir servicios ricos en información basados en repositorios digitales

- DRIVER 2.0:

- Incluye esfuerzos para expandir, enriquecer y fortalecer los resultados de DRIVER, en las siguientes áreas:

- Estrategia geográfica y expansión de la comunidad por medio de la confederación DRIVER.
    - Establecimiento de una robusta infraestructura escalable de repositorio acompañada por el paquete de software abierto D-Net
    - Cobertura más amplia de contenidos a través del uso de publicaciones mejoradas
    - Mayor difusión y programas de promoción
    - Soporte continuado de repositorios
    - Guías para interoperabilidad en la mayor comunidad europea de bibliotecas digitales
    - Acceso abierto para los materiales de investigación europeos

# Recursos textuales en DRIVER

DRIVER se centra en recursos textuales en formato digital de acceso abierto

- ‘**recurso textual**’: artículos científicos, tesis doctorales, ensayos, libros electrónicos y formatos similares para actividades de investigación científica
- ‘**acceso abierto**’: acceso sin necesidad de ningún medio de pago, licencias, control de acceso con contraseña, control de acceso mediante ip, etc.

# Necesidad de las directrices

- Sirven de guía:
  - a los **administradores de nuevos repositorios** en la definición de las políticas de administración de datos
  - A los **administradores de repositorios existentes** en pasos a seguir para conseguir servicios mejorados
  - A los **desarrolladores de plataformas de repositorios** para la incorporación de funciones adicionales en versiones futuras

# Validación de directrices a través de interfaz web

- Según grado de conformidad con directrices, estado de repositorio:
  - **Validado** (cumple puntos obligatorios): pasan a formar parte de la red DRIVER de proveedores de contenido. Pueden reutilizar datos de DRIVER para desarrollar servicios locales.
  - **Con futuro** (cumple además puntos recomendados)

La no conformidad con todos los puntos obligatorios o recomendados de las directrices no significa, que el contenido no sea recolectado o agregado por DRIVER, pero sí que en función de sus servicios específicos, el contenido no sea recuperable.



# Componentes principales de las directrices DRIVER

- Colecciones

obligación de utilizar 'sets' (agrupaciones) que definan las colecciones accesibles a texto completo

- Protocolo OAI-PMH

Definición de algunas características obligatorias y otras recomendadas para solucionar problemas en distintas implementaciones de repositorio local

- Metadatos

Definición de algunas características obligatorias y otras recomendadas para solución de dificultades semánticas por diferentes interpretaciones de DUBLIN CORE

# DIRECTRICES DRIVER

- Apartado A: RECURSOS TEXTUALES
- Apartado B: METADATOS
- Apartado C: IMPLEMENTACIÓN DE OAI-PMH

# Directrices DRIVER 2.0

## Apartado A - Recursos textuales

### Obligatorias

El repositorio contiene **recursos textuales digitales**

Los recursos textuales están en formatos ampliamente utilizados y extendidos (PDF, TXT, RTF, DOC, TeX,...)

Los recursos textuales son de **acceso abierto**, están disponibles directamente en el repositorio para cualquier usuario del mundo sin restricciones tales como autorización o pago

Los recursos textuales están **descritos mediante registros de metadatos**

Los metadatos y los recursos textuales se vinculan entre sí de tal modo, que un usuario final pueda **acceder** al recurso textual **mediante un identificador** (normalmente una URL) en el 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 señala a sistemas externos tales como una biblioteca nacional o un editor**)

### Recomendadas

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 tales como DOIs, URNs, ARKs

El uso del contenedor-XML DIDL para la exposición de recursos textuales (capítulo “Use of MPEG-21 DIDL (xml-container) – Compound object wrapping



Apartado B – Metadatos	Obligatorios	Los metadatos se estructuran según la norma <b>Unqualified Dublin Core</b> (ISO 15836:2003)
		Los elementos individuales de DC se utilizan según lo dispuesto en el documento “Use of Metadata OAI_DC”
	Recomendados	Los metadatos se estructuran preferiblemente según esquemas más completos, como Dublin Core Qualified o MODS*
		Se recomienda utilizar el inglés
		El idioma recomendado para los resúmenes (los resúmenes son optativos) del artículo es el inglés

\* Las guías para esos esquemas más completos se seguirán en la futura versión de las Directrices DRIVER



<p>Apartado C - Implementación de OAI-PMH</p>	Obligatorios	El repositorio debe ser conforme con <b>OAI-2.0</b> y con la especificación del capítulo “Use of OAI-PMH”
		Debe existir un <b>identificador de repositorio</b> y debe utilizarse el esquema de <b>identificador OAI</b>
		Si (y sólo si) el repositorio contiene recursos que no sean obligatorios en el “APARTADO A – Recursos textuales”, debe definirse un set OAI que identifique la colección de recursos textuales digitales accesibles en Open Access*
	Recomendados	Previsiones de modificación de URL base
		Respuesta completa a Identify, incluido el uso de la declaración Descripción
		Uso del contenedor XML de DIDL
		Uso persistente de la estrategia de borrado temporal (Transient deleting strategy)
		Uso de un tiempo de conservación de los testigos de reanudación (Resumption Token)

\* Ver explicaciones “DRIVER set naming”, “DRIVER Set Content definitions” y “Set Location”



# Los elementos descriptivos

En estas Directrices DRIVER, los gestores de repositorios tienen que aceptar el hecho de que no todo se puede expresar con Dublin Core sin calificar y que además, estas Directrices concentran la información más importante en la perspectiva del usuario final, que no es bibliotecario.

- En DRIVER el uso de los elementos puede ser:
  - **Obligatorio (M)**: el elemento debe estar siempre presente en el registro de metadatos
  - **Obligatorio cuando aplicable (MA)**: cuando el elemento se pueda obtener, debe estar presente en el registro de metadatos
  - **Recomendado (R)**: el uso del elemento está recomendado
  - **Opcional (O)**: No es importante si el elemento se usa o no.

Basic element	Status	Encoding schemes
Title	M	None, free text
Creator	M	<a href="#">APA bibliographic writing style as in a reference list</a> . Syntax: surname, initials (first name)
Subject	MA	Choice of keywords and classifications can be free text (preferably in English) and defined by an URI scheme (preferably info:eu-repo/classification)
Description	MA	None, free text. Recommended practice is to include an abstract in English. “Abstract” is the default interpretation to the value for dc:description
Publisher	R	None
Contributor	O	<a href="#">APA bibliographic writing style as in a reference list</a> . Syntax: surname, initials (first name)
Date	M	Date   ISO 8601 W3C-DTF - “Published” is the default interpretation to the value for dc:date
Type	M	Publication type and Version type can be free text (preferably in English) and defined by an URI scheme (preferably info:eu-repo/semantics).
Format	R	<a href="#">IANA registered list of Internet Media Types</a> (MIME types)
Identifier	M	URI scheme, linking to persistent identifier (URN, handle, DOI), full text document or human start page.
Source	O	<a href="#">Guidelines for Encoding Bibliographic Citation Information in Dublin Core Metadata</a> as in dcterms:bibliographicCitation
Language	R	ISO 639-3
Relation	O	None
Coverage	O	“Period” is the default interpretation to the value for dc:coverage Encoding: DCMI Period [http://dublincore.org/documents/2000/07/28/dcmi-period/] For more encoding schemas see Chapter 5 Use of vocabularies and semantics.
Rights	R	None
Audience	O	None. “Education level” is the default value for dc:audience.



# Normalización de descripción: Interoperabilidad semántica

- **Creator\*** y **Contributor**: sintaxis: Apellido, iniciales (primer nombre) APA bibliographic writing style as in a reference list.
- **Subject\***: en inglés. Para esquemas de clasificación o vocabularios controlados se recomienda uso de un URI
- **Type**: **Vocabularios controlados** para *Tipo de publicación* (DRIVER-type vocabulary) y *Versión* (DRIVER-version vocabulary)
- **Format**: IANA registered list of Internet Media Types (MIME types)
- **Source**: Sólo para recursos resultado de digitalización de originales no digitales. Guidelines for Encoding Bibliographic Citation Information in Dublin Core Metadata

\*Identificadores persistentes (personas, temas)



- Type:

article, BachelorThesis, masterThesis, doctoralThesis,  
book, bookPart, review, conferenceObject, lecture,  
workingPaper, preprint, report, annotation,  
contributionToPeriodical, patent, other

- Version:

Draft, submittedVersion, acceptedVersion,  
publishedVersion, updatedVersion

# Conclusiones

¿Por qué seguir las directrices DRIVER?

Porque el fin del RI del SSPA es **alojar**, gestionar, preservar y **facilitar el acceso** a la producción científica del SSPA para **ponerla al servicio** de la sociedad y del resto de los profesionales, haciendo que el contenido sea lo más fácilmente recuperable y **visible** para la propia Institución, la comunidad científica y el resto de la sociedad.

Para ello debemos hacerlo siguiendo **estándares de compatibilidad e interoperabilidad internacionales**.



# Futuro de los Repositorios Institucionales

## International Repositories Infraestructure Workshop.

Amsterdam, 16-17 marzo 2009

Papel futuro de los RI: no sólo mejorar acceso e información científica a los usuarios, sino también a otros servicios, repositorios y sistemas. Repositorio no es un fin en sí mismo, sino un componente de infraestructura más amplia.



## 4 cuestiones prioritarias:

- **Infraestructura de Identificadores persistentes:** 4 entidades: objetos, personas, Instituciones y repositorios.
  - Necesidad de un meta-servicio que construya puentes entre los diferentes mecanismos (servicio de similaridad/equivalencia)
- **Servicios de citas de los repositorios.** Objetivo- además de compartir los datos de las citas, una vez reunidos, posibilidad de reconocer con exactitud las citas a los resultados de la investigación en los repositorios y en la web: evidencia de impacto de esos resultados en la comunidad
- **“Repository handshake”** (o extendiendo la mano): buscar procedimientos que facilitaran el depósito de los artículos en los repositorios como un medio para aumentar los contenidos de los mismos (introducción de datos o de documentos en un repositorio y transferencia automática a otros.
- **Una organización para la infraestructura internacional de repositorios.** Definir un plan de acción para crear la Confederación o la Organización Internacional de Repositorios. Posibles beneficios:
  - crear una voz colectiva internacional con influencia para promover repositorios de acceso abierto y las infraestructuras apropiadas; hacer visibles a escala global los esfuerzos de repositorios regionales y nacionales; crear una red de experiencia; conectar a las partes implicadas y al público en general; ayudar a desarrollar estándares de interoperabilidad; ofrecer un sello de calidad; apoyar la formación profesional para los gestores de repositorios; facilitar financiación para las infraestructuras de repositorios.

# Bibliografía

DRIVER Guidelines 2.0 Guidelines for content providers: Exposing textual resources with OAI- PMH. November 2008. Consultado en: [http://www.driver-support.eu/documents/DRIVER\\_Guidelines\\_v2\\_Final\\_2008-11-13.pdf](http://www.driver-support.eu/documents/DRIVER_Guidelines_v2_Final_2008-11-13.pdf). Acceso 9 de febrero de 2009. (Archived by WebCite® at <http://www.webcitation.org/5gAsFATHH>)

López Medina A. Informe sobre el "International Repositories Infrastructure", Amsterdam, 16-17 de marzo, 2009. El e-espacio de los contenidos digitales de la Uned (<http://e-spacio.uned.es/fez/view.php?pid=bibliuned:23164>) Consultado en: [http://e-spacio.uned.es/fez/eserv.php?pid=bibliuned:23164&dsID=International\\_Repositories\\_Workshop\\_Amsterdam\\_JISC.pdf](http://e-spacio.uned.es/fez/eserv.php?pid=bibliuned:23164&dsID=International_Repositories_Workshop_Amsterdam_JISC.pdf) Acceso 23 de abril de 2009.

Térmens i Graells M. Compliment de les directrius DRIVER dels repositoris cooperatius RACO, RECERCAT i TDX. Marzo 2009. RECERCAT. Consultado en: <http://hdl.handle.net/2072/13687> Acceso 27 de abril de 2009.

Prats, J. Directrices DRIVER para proveedores de contenidos. Abril 2008. Disponible en: <http://upcommons.upc.edu/e-prints/handle/2117/1998>. Acceso 27 de abril de 2009





# Biblioteca Virtual

del Sistema Sanitario Público de Andalucía

BIBLIOTECA VIRTUAL DEL SISTEMA SANITARIO PÚBLICO DE ANDALUCÍA

C/ Capitán Vigueras, 1 – 1ºE. Mod.141004-Sevilla

Tel.: +34 954 99 49 20 (corp. 394 920)

Fax: +34 954 99 49 35 (corp. 394 935)

<http://www.juntadeandalucia.es/salud/bibliotecavirtual>

[bibliotecavirtualsalud@juntadeandalucia.es](mailto:bibliotecavirtualsalud@juntadeandalucia.es)



Esta obra está bajo una licencia Creative Commons  
[Reconocimiento-Sin obras derivadas 3.0 España](https://creativecommons.org/licenses/by-nc-nd/3.0/es/)

III Jornadas de Bibliotecas del SSPA  
Ronda, 25-27 de mayo de 2009

Organiza: **BV-SSPA**