A stylized map of Latin America is shown in dark blue. It is overlaid with a network of white dashed lines and colorful nodes (yellow, orange, pink, green, blue). The map is partially covered by large, semi-transparent colored shapes: green for Central America, orange for the Caribbean, and pink for South America. The background is a light blue with abstract white lines and squares.

# Hacia una red latinoamericana de recursos educativos abiertos (REA)

Autores: Diego Ferreyra / Relpe

 relpe

red latinoamericana  
portales educativos 



## Prólogo

---

*En los últimos años pueden observarse dos grandes transformaciones en la dinámica de los contenidos en línea; por una lado el creciente predominio de contenidos generados por usuarios (User-generated content) a través de diferentes plataformas y ambientes en línea; y por el otro, la estabilización y crecimiento lento de fuentes de datos semi-estructurados en línea.*

*Este proceso de articulación nos permite identificar un horizonte comunicacional caracterizado por la generalización de nuevas formas de socialización en ambientes digitales y la consolidación de fuentes estables de datos y recursos en la red. De alguna manera, estos dos fenómenos interpelan dos importantes ejes en la historia de RELPE, a saber: la construcción de los portales educativos como espacios de mediación y participación para las comunidades educativas de cada uno de los países miembro; y en segundo término, la necesidad de garantizar condiciones para la difusión, permanencia, reúso y apropiación de los recursos educativos generados por los portales.*

*Considerando lo antedicho, se presentan en esta publicación dos documentos orientados a mejorar y extender las capacidades de los portales miembro de RELPE en sus estrategias para el reúso y exposición de recursos educativos a través de protocolos abiertos en el contexto de la web.*

*En primer término se presenta un documento orientado a realizar una presentación del marco de datos abiertos enlazados en general y sus alternativas en el sector educativo en particular. El documento contextualiza históricamente la emergencia de la agenda de datos abiertos enlazados y describe su significado en términos de propuesta de infraestructura técnica de servicios como así también en tanto propuesta conceptual, en tanto apuesta proyectual para una Web futura. Finalmente, el documento describe las diversas vías y grados de implementación que prevé el marco de datos abiertos enlazados.*



*En segundo término se presenta la propuesta de RELPE para la explotación cooperativa y federada de recursos educativos. Dicha propuesta, publicada ya con antelación, fue elaborada y evaluada a través de una serie de experiencias impulsadas por RELPE durante los años 2010 y 2011 y en las que participaron un conjunto acotado de portales con el fin de evaluar la factibilidad técnica de la arquitectura propuesta, seleccionar esquemas de solución tecnológica acordes a la diversidad estructural y tecnológica de los portales miembro de RELPE y elaborar un marco metodológico viable para la implementación de la propuesta. El documento presenta la fundamentación de la propuesta, la descripción técnica del modelo y describe tres esquemas de solución según los posibles escenarios relevados en el contexto de RELPE.*

*Presentamos en conjunto estos documentos con el objetivo de acercar herramientas conceptuales y técnicas que apoyen las estrategias y políticas de gestión de recursos educativos y provean marcos de análisis útiles para la toma de decisiones socio-técnicas en el contexto de la comunidad de prácticas de portales educativos miembros de RELPE. Esperamos que se multipliquen las experiencias y los aprendizajes compartidos.*

## El marco de dato abiertos

Consultor:

**Diego Ferreyra**



Esta publicación se basa en el informe realizado a solicitud de RELPE, Red Latinoamericana de Portales Educativos, con financiamiento de OEI, Organización de Estados Iberoamericanos.

Marzo 2015





## INDICE

<b>1</b>	<b>Los datos abiertos enlazados</b> .....	<b>7</b>
1.1	Datos estructurados.....	9
1.2	Datos no estructurados.....	9
1.3	Datos semi-estructurados.....	12
<b>2</b>	<b>La forma de organizar los datos</b> .....	<b>14</b>
2.1	SPARQL: Un camino para la explotación y reutilización de datos.....	17
2.2	Condiciones óptimas para una web de datos.....	19
<b>3</b>	<b>Datos para gobierno abierto</b> .....	<b>22</b>
<b>4</b>	<b>El marco de datos abiertos en educación</b> .....	<b>24</b>
<b>5</b>	<b>Escenarios diversos requieren soluciones diversas</b> .....	<b>26</b>
5.1	Los datos abiertos enlazados son un camino.....	27
5.2	Impacto tecnológico acotado.....	27
5.3	Una red de datos abiertos enlazados.....	28
5.4	Del banco de recursos a las herramientas de descubrimiento.....	28
<b>6</b>	<b>Anexos</b> .....	<b>30</b>
6.1	Sitios web de interés.....	30
6.2	Documentos y marcos elaborados por el W3C.....	31
6.3	Colecciones de datos en educación.....	31
6.4	Algunas lecturas recomendadas.....	33





## Los datos abiertos enlazados

Los datos abiertos enlazados (en adelante LOD, siglas en inglés de Linked Open Data) constituyen una propuesta iniciada y elaborada por el Consorcio WWW (W3C) hace poco menos que una década, y que está orientada a consolidar un ecosistema en el que los datos puedan ser consultados y utilizados globalmente<sup>1</sup>. Así como la WWW puede ser vista como un ambiente digital de información abierta en el que es posible consultar contenidos que se encuentran enlazados a través de hipervínculos, la propuesta de LOD pretende llevar estas capacidades y funcionalidades al mundo de los datos estructurados y semiestructurados.

Para conceptualizar el horizonte de expectativas concebido por el W3C a la hora de encarar el proyecto de LOD resulta útil considerar que, ya en los lineamientos iniciales de diseño de la web, la arquitectura propuesta estaba orientada a resolver problemas de articulación entre marcos tecnológicos y organizacionales en el

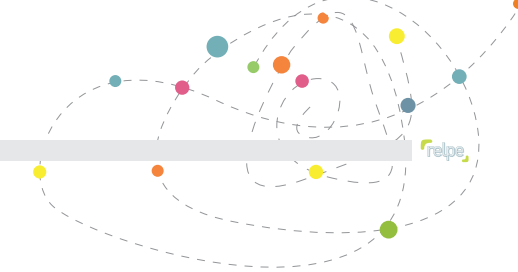
contexto de la gestión de bases de datos<sup>2</sup>. En términos más amplios, se proponía crear un “mundo interactivo de la información compartida a través del cual las personas pudieran comunicarse entre sí y con las máquinas”. Desde sus inicios, el proyecto de la web, en tanto proyecto tecnológico, contenía los conceptos que luego desbordarían lo meramente técnico atravesando y transformando las prácticas sociales en su conjunto: interacción, información compartida, comunicación social y comunicación mediada por máquinas. En este sentido, es posible considerar la propuesta de LOD, en tanto propuesta técnica, como la expresión de continuidad del proyecto inicial de la WWW.

Para observar de manera más concreta esta continuidad, consideremos por un momento algunas características de la WWW como tal:

1 Tim Berners-Lee (2006). Linked Data Date: 2006-07-27, last change: 2009/06/18. <http://www.w3.org/DesignIssues/LinkedData.html>

2 Tim Berners-Lee (1996). The World Wide Web: Past, Present and Future. <http://www.w3.org/People/Berners-Lee/1996/ppf.html>





*“The Web is a space in which resources are identified by Uniform Resource Identifiers (URIs). There are protocols to support interaction between agents, and formats to represent the information resources. These are the basic ingredients of the Web.”<sup>3</sup>*

Como vemos, la web es un espacio digital en donde hay recursos de todo tipo, que son identificados por identificadores uniformes, es decir, que siguen reglas definidas para su redacción, al igual que los números telefónicos (código de país, código de área, número telefónico). Con el fin de permitir el acceso a estos recursos, existen protocolos que dinamizan y brindan operatividad a la localización de estos recursos a través de los identificadores. Siguiendo el ejemplo de los teléfonos, serían como los protocolos telefónicos que permiten que, al digitar un número telefónico, la comunicación se establezca. Por último, los .....

<sup>3</sup> Berners-Lee, Tim, Hall, Wendy, Hendler, James A., O'Hara, Kieron, Shadbolt, Nigel and Weitzner, Daniel]. (2006) A Framework for Web Science. Foundations and Trends in Web Science, 1, (1), 1-130.

recursos, en el contexto de la web, tienen distintos formatos de representación: identificadores uniformes de recursos (URI), como por ejemplos las URL de cada sitio miembro de RELPE, el protocolo de transferencia de hipertexto (HTTP) y los distintos esquemas de representación de recursos en uso actualmente en la web, como ser el HTML (o también XML, PDF, etc). Al decir de Tim Berners-Lee: “Estos son los ingredientes básicos de la Web”.

Ahora bien, estos ingredientes favorecieron el desarrollo de un ambiente global de comunicación al permitir no solo la disponibilidad de un sin fin de recursos en línea, ya que la propuesta de LOD pretende llevar las capacidades y funcionalidades de los datos no estructurados al mundo de los datos estructurados y semiestructurados. Para poder abordar este procedimiento resulta conveniente formalizar la distinción entre estos tipos de datos:

### 1.1 DATOS ESTRUCTURADOS

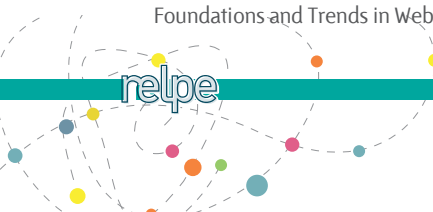
Son datos almacenados de acuerdo con un esquema definido y en general orientados o involucrados en un proceso de gestión. Son datos consolidados y provistos de un esquema formal. Son ejemplos comunes de este tipo de datos los archivos Excel, las bases de datos, los reportes de un sistema, datos analíticos, etc. Desde el punto de vista de su implicancia para el desarrollo de servicios en línea, podemos identificar los siguientes aspectos relevantes:

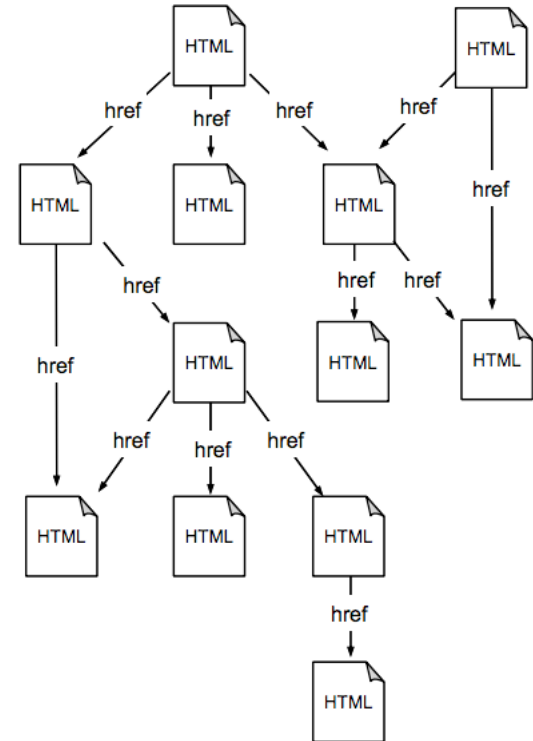
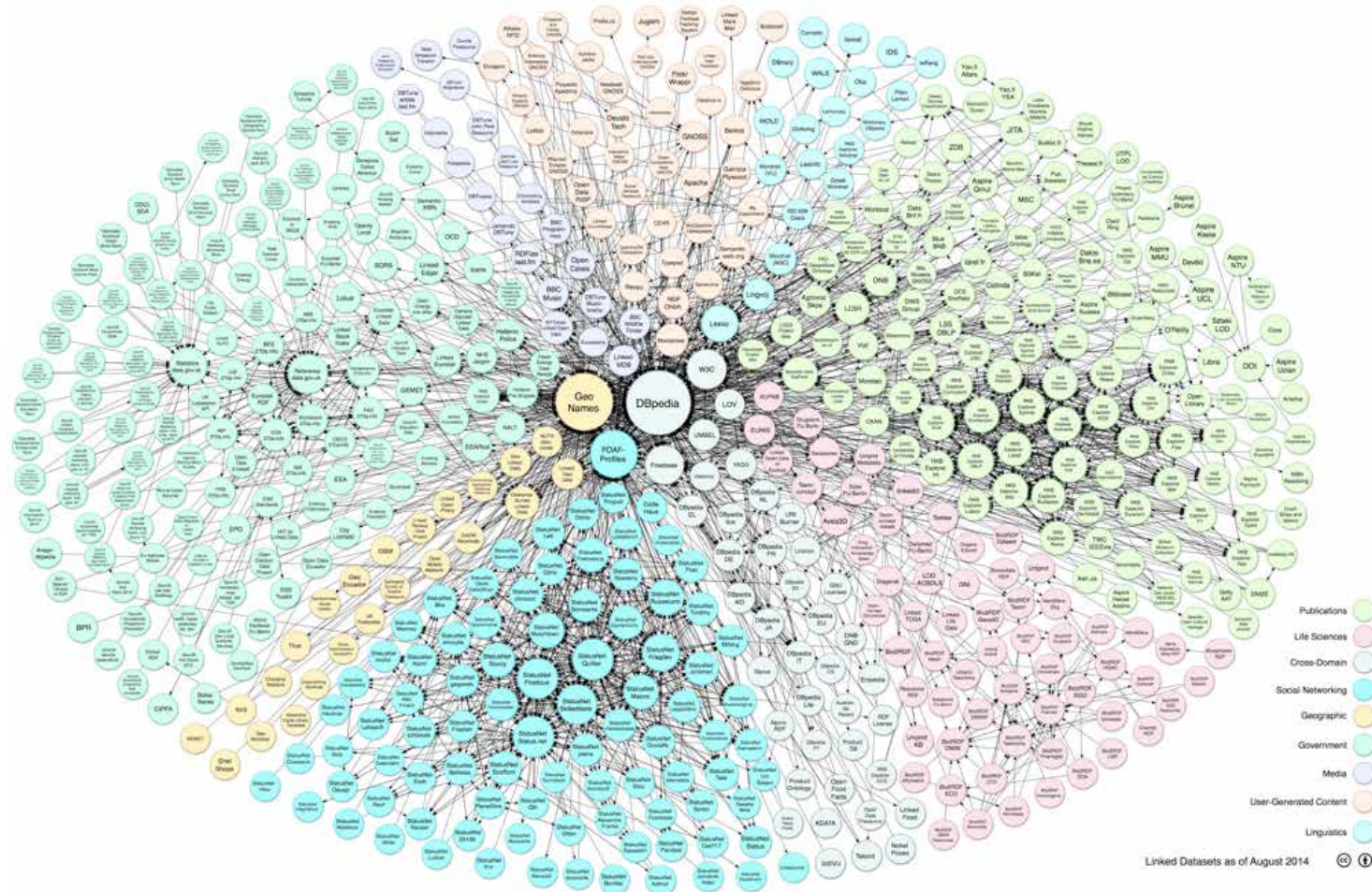
- La forma en que se estructuran los datos es conocida y en general explícita, esto facilita su manipulación y análisis por medios automatizados.
- Tienen poca flexibilidad. Al disponer de estructuras formales, es complejo incorporar nuevos atributos o tipos de datos al modelo.
- Facilitan los procesos de control de calidad de datos. Es sencillo realizar consultas formales o utilizar herramientas informáticas para detectar problemas de normalización, integralidad o consistencia.

- Ciclo de vida. A medida que los datos se encuentran más ligados a un proceso de gestión en particular, tienen un alto nivel de estructuración y normalización; pero su vida media útil resulta más acotada al proceso mismo en el que están involucrado. Por ejemplo, un proceso de pago de honorarios o el recuento de inventario de una institución.

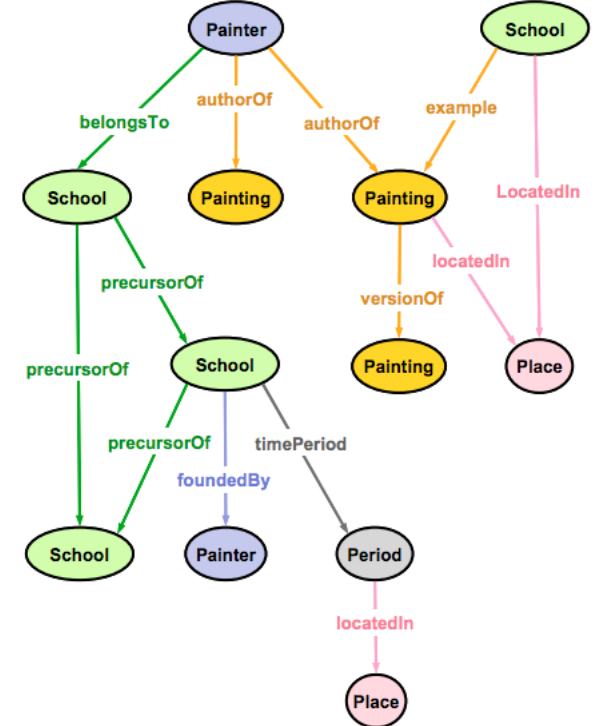
### 1.2 DATOS NO ESTRUCTURADOS

Son datos que no se encuentran organizados de acuerdo con esquemas formales específicos o cuyos esquemas se consideran implícitos. Un ejemplo común podría ser un documento académico: si bien el texto sigue un plan textual relativamente formal, el mismo no resulta explícito en el mismo documento. Otros ejemplos comunes son las páginas HTML, los documentos en general, imágenes, etc. Constituyen el tipo de dato más común en los portales de contenidos. En relación con los aspectos antes analizados:





a) Web actual



b) Web semántica

Figura 2. La web actual vs. la web semántica.

Fuente: <http://lod-cloud.net/>





- **La forma en que se estructuran los datos es muchas veces desconocida e implícita**, lo que dificulta su análisis por medios automatizados. En general, requieren de tecnologías especializadas en análisis de lenguaje natural o reconocimiento de patrones (Google, modelos estadísticos, algoritmos entrenables, etc.).
- **Tienen alta flexibilidad**, ya que al no disponer de estructuras formales fijas es posible variar el tipo de datos o contenidos. Por ejemplo, podemos insertar una imagen, un sonido o una tabla en una página HTML.
- **Procesos de control de calidad de datos.** Puede ser costoso uniformar y normalizar un banco de recursos no estructurados.
- **Ciclo de vida.** Pueden ofrecer una vida media extensa. Con relativa independencia en relación con la forma en que están estructurados, la vida media de estos recursos depende de las prácticas socio-culturales y de las estrategias de cada comunidad para construir su memoria social. Por ejemplo, las constituciones nacionales o la Biblia.

### 1.3 DATOS SEMI-ESTRUCTURADOS

Finalmente, el tipo de datos predominantes en los portales de contenidos son los que disponen de algún nivel de estructuración y formalización sin llegar al nivel de integralidad estructural de un ciclo de gestión administrativo o datos estadísticos; pero de todos modos, es posible identificar categorías débiles de organización. Es el caso, por ejemplo, de los catálogos y los bancos de recursos en general, es decir, fuentes de datos que disponen de datos estructurados (los metadatos) que refieren a datos no estructurados (recursos educativos, por ejemplo). De manera cada vez omnipresente, también forman parte de este segmento los datos provistos por la interacción con medios sociales y ambientes en línea: son datos en los que es posible identificar la estructuración de ciertos datos (quién, cuándo, dónde, con respecto a qué, etc.) que incluyen otros de carácter no-estructurado (por ejemplo, el contenido del mensaje en un foro).

En virtud de esta perspectiva, consideremos nuevamente algunas particularidades de la web como ecosistema. Imaginemos

por un momento cómo sería una web sin enlaces, como un espacio con la misma cantidad de recursos y contenidos disponibles hoy, pero sin la posibilidad de establecer hipervínculos entre los mismos <sup>4</sup>. La proyección de éste es el tipo de escenario en el mundo de los datos y la posibilidad de retomar la agenda inicial de la web. De alguna manera, convergieron para dar cuerpo al proyecto de LOD: la idea de que es posible recrear la organicidad lograda por el ecosistema de contenidos abiertos a través

hipervínculos de la web, pero con datos estructurados y semi-estructurados.

Como vimos anteriormente, la vinculación entre recursos que permite el marco técnico de la web constituye en gran medida su especificidad diferencial con respecto a otros entornos. Replicar esta capacidad en el universo de los datos estructurados y semi-estructurados implica ciertos cambios. A continuación realizaremos una pequeña revisión de los más relevantes.

.....  
 4 Barabasi, Albert-Laszlo (2003) Linked: The New Science of Networks. Perseus Books Group.







## 2. La forma de organizar los datos

El marco conocido como “web semántica” expresa la estrategia técnica elegida por el W3C para mejorar las condiciones de disponibilidad e interoperabilidad de datos. Este marco, propuesto originalmente en 1998, se encuentra orientado a mejorar la manera en que se codifican los datos utilizando categorías de descripción formales y específicas para cada entidad y para cada relación posible entre entidades. Por ejemplo, si disponemos de un recurso educativo, el mismo tendrá un título, un autor y un mediador didáctico privilegiado al que está orientado. Si enunciamos estas características tal como están redactadas en este mismo documento, será legible e inteligible para cualquier persona que sepa leer; pero si quisiéramos que un proceso automático reconociera en el texto cuál es el autor o a quién está orientado, sería ya bastante más complejo. El marco de solución propuesto por el W3C para lograr que procesos automatizados puedan reconocer estas entidades —autor, recurso, mediador

didáctico y cualquier otra entidad— y las posibles relaciones que pudieran establecerse se denomina Resource Description Framework (RDF, marco de descripción de recurso). Este marco propone describir todas las entidades a partir de sentencias compuestas por un sujeto, un predicado y un objeto, en donde el sujeto es de quién se habla, el predicado es lo que se predica o se dice acerca del sujeto y el objeto es el que recibe la acción o característica expresada por el predicado. Por ejemplo:

SUJETO	PREDICADO	OBJETO
Pedro	Es autor de	recurso
Recurso	Esta orientado a	docente
Docente	Es usuario de	Portal
Portal	Tiene URL	<a href="http://www.relpe.org">http://www.relpe.org</a>
<a href="http://www.relpe.org">http://www.relpe.org</a>	Es una	Red de portales educativos

Observando la tabla, podemos diferenciar tres aspectos centrales del RDF:

**1. Economía estructural:** observemos que con solo tres columnas es posible describir prácticamente cualquier entidad. Si surgiera algún aspecto nuevo con respecto a un recurso (por ejemplo, si quisiéramos registrar quién ha utilizado un recurso), podemos agregar un nuevo tipo de predicado, por ejemplo: “*fue usado por*”:

SUJETO	PREDICADO	OBJETO
El recurso	Fue utilizado por	José María Usuario

Si se incorporara un nuevo miembro a RELPE, se puede agregar una fila. De esta manera se verifica que con una misma estructura sencilla es posible describir un recurso que puede variar ya sea en su cantidad como en sus cualidades. Este tipo de sentencias conformadas por tres componentes se denomina **triple** o **tripleto**.

**2. Datos literales o referenciados:** como vimos en el ejemplo, es posible utilizar el nombre de una entidad como un dato

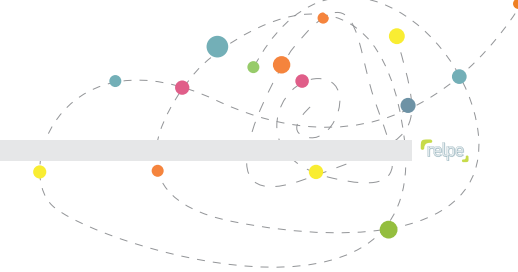
literal (RELPE, José María Usuario, etc.), pero también es posible utilizar un identificador único (URI) para designar o referenciar una entidad (<http://www.relpe.org>, [jmaria.usuario@relpe.org](mailto:jmaria.usuario@relpe.org)). También se puede utilizar una referencia hacia un servicio o fuente de datos capaz de aportar información estructurada y enriquecida acerca de una entidad.

Por ejemplo, es posible decir: “*Alfredo Bryce Echenique es el autor de Un mundo para Julius*”, pero también es posible decir lo mismo utilizando identificadores únicos:

SUJETO	PREDICADO	OBJETO
<a href="http://viaf.org/viaf/68926324/#Bryce_Echenique,_Alfredo,_1939">http://viaf.org/viaf/68926324/#Bryce_Echenique,_Alfredo,_1939</a>	<a href="http://purl.org/dc/elements/1.1/creator">http://purl.org/dc/elements/1.1/creator</a>	<a href="https://openlibrary.org/books/OL9200458M/Un_Mundo_Para_Julius">https://openlibrary.org/books/OL9200458M/Un_Mundo_Para_Julius</a>
Alfredo Bryce Echenique	es el autor de	Un mundo para Julius

**3. Definición de vocabularios:** como seguramente a esta altura resultará evidente, es posible definir un conjunto específico de predicados para cada tipo de entidad, por ejemplo: las rece-





tas tendrán un predicado que del estilo “ser ingrediente de” o un recurso educativo podrá tener un predicado que podría ser “orientado al rol educativo...” para definir la rol dentro del sistema educativo al que está orientado un recurso.

En el marco de la web semántica, la definición del conjunto de predicados que pueden ocurrir en un esquema específico se denomina vocabularios, ya que se considera que esas son las categorías que modelan, organizan y definen una determinada concepción de un fenómeno o entidad. El vocabulario “Dublin Core” expresa las categorías posibles para describir genéricamente un recurso web. El vocabulario LRMI está orientado a la descripción de recursos educativos <sup>5</sup>.

A través del uso de vocabularios e identificadores uniformes para la tipificación de entidades, es posible lograr en un mismo procedimiento normalizar la denominación y designación de entidades (a través de los identificadores uniformes) y la de las relaciones posibles entre entidades (a través de los vocabularios).

5 (<http://dublincore.org/dcx/lrmi-terms/2014-10-24/>)

Estas son las razones por las cuales el W3C considera que sería posible lograr un ecosistema orgánico similar al logrado en la web de los contenidos no-estructurados a través de las páginas HTML y sus hipervínculos; siempre que los actores que participan de la web expresen sus recursos identificándolos y referenciándolos con identificadores uniformes persistentes y utilizaran vocabularios basados en el marco RDF para expresar sus relaciones. En el caso de la web de datos, sin embargo, las páginas HTML serían reemplazadas por datos identificados unívocamente y las relaciones entre ellos serían designadas por vínculos categorizados y normalizados.

Es interesante notar que en la web los hipervínculos son unidireccionales, es decir, van de una página “A” hacia la “B”, pero la página “B” puede no saber siquiera que existe el enlace. A su vez, es complejo asignarle un sentido a un enlace en el contexto de la web. Un enlace, por ejemplo, puede referenciar una procedencia, autoría o hasta un repudio. El marco de la web semántica pretende justamente resolver estos fenómenos percibidos como un problema para el proyecto inicial de la web, a través del uso

de referencias basadas en identificadores únicos y vocabularios que definen el tipo de vinculación entre entidades:

*Al igual que la web de hipertextos, la web de los datos se construye con documentos en la web. Sin embargo, a diferencia de la web de hipertextos, donde los vínculos son relaciones marcadas en documentos hipertextuales escritas en HTML, para los datos los vínculos entre cosas arbitrarias están descritos en RDF. Los URI identifican cualquier tipo de objeto o concepto. <sup>6</sup>*

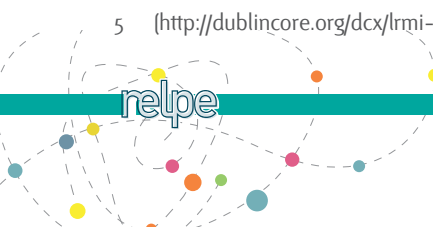
## 2.1 SPARQL: UN CAMINO PARA LA EXPLOTACIÓN Y REUTILIZACIÓN DE DATOS

Una vez codificados los datos en RDF es posible imaginar diversas vías para su exposición y diseminación, y crear un archivo estático descargable y publicarlo o exponer servicios a través de un servidor OAI-PMH, por ejemplo. Todas estas son estrategias posibles ampliamente utilizadas en el mundo en los diferentes

6 Tim Berners-Lee (2006). Linked Data Date: 2006-07-27, last change: 2009/06/18. <http://www.w3.org/DesignIssues/LinkedData.html>

catálogos de datos abiertos. Sin embargo, si más allá de exponer los datos se desea permitir que otros actores puedan realizar operaciones de búsqueda e interrogación sobre nuestros datos —sin necesidad de realizar una copia de los mismos en su propia infraestructura—, se puede recurrir al protocolo SPARQL.

SPARQL es el acrónimo de SPARQL Protocol and RDF Query Language (es un acrónimo recursivo, SPARQL es la partícula que se repite recursivamente), y se trata de un protocolo que permite realizar consultas utilizando los vocabularios y los valores de los sujetos, predicados y objetos como si fueran filtros de una planilla de cálculo. De esta manera, es posible interrogar a un servidor que disponga de un punto de consulta SPARQL (SPARQL endpoint) tanto acerca de recursos específicos como así también consultas agregadas que permitan sacar conclusiones que describan la composición de un conjunto de datos. Por ejemplo, se puede realizar una sentencia como: “Los recursos educativos creados en el 2014”, donde el sujeto son los “recursos educativos”, el predicado es “creados en” y el objeto es “el año 2014”. Así se obtiene una lista de recursos educativos, pero que al ser un





lenguaje de consultas también permite realizar preguntas para descubrir características del set de datos, como ser: cuántos recursos fueron creados por el mismo autor o la cantidad de recursos agrupados por tipo de licencia.

Este tipo de puntos de consulta son implementados por sistemas que genéricamente se denominados triplestores, herramientas que permiten almacenar los tripletes y soportan métodos de consulta de los mismos justamente a través de protocolos como SPARQL. Estas herramientas disponen también de mecanismos de ingesta e importación desde otras fuentes.

En este punto, es muy importante resaltar que el marco de LOD supone el uso de al menos dos esquemas para la gestión de datos:

- Esquema primario de datos: se trata del entorno encargado de administrar los datos y recursos para el desarrollo de la infraestructura pública de servicios. En el caso de los portales de RELPE, por ejemplo, serían los distintos CMS utilizados, los repositorios de recursos o incluso los sistemas de gestión educativa como tales.

- Esquema de exposición de datos abiertos (enlazados): más allá de la infraestructura de gestión utilizada, se consolidan set de datos, ya sea estáticos (Excel, XML, RDF, etc.) o a través de la importación o ingesta de datos en bancos de tripletes (triplestores) que dispongan de puntos de consulta SPARQL. Son estos sets de datos los que son expuestos y ofrecidos para su consulta y reutilización.

Como es posible apreciar, la implementación de un esquema de datos abiertos, en general, y de datos abiertos enlazados, en particular, no implica necesariamente modificar la infraestructura tecnológica en uso por cada uno de los portales, sino más bien implica una política de consolidación de datos y luego de selección de una estrategia de exposición de los mismos.

## 2.2 CONDICIONES ÓPTIMAS PARA UNA WEB DE DATOS

Desde el punto de vista del W3C, las mismas condiciones que promueven y mejoran la web como un ambiente estable, resultan necesarias para el desarrollo y crecimiento de una web de datos. A continuación presentamos las condiciones que el W3C considera que promueven el desarrollo de la web, en general, y de la web de datos, en particular:

1. Utilice *URI* como nombres para las cosas.

Si bien es posible utilizar tanto valores literales como *URI* para referenciar recursos, el uso de *URI* facilita la reutilización y la normalización de los datos.

2. Utilice las *URI* basadas en *HTTP* para facilitar que las personas puedan buscar esos nombres.

Como se presentó en el comienzo del documento, el protocolo encargado de dinamizar la navegación e interacción en el contexto de la web es el HTTP. El W3C considera que este protocolo

debería ser suficiente para la localización de recursos, en detrimento del uso de protocolos o interfaces agregadas (API, interfaz de programación de aplicaciones).

3. Cuando alguien busca una URI, provea información útil utilizando estándares.

En este punto se refiere a la conveniencia de que, una vez que un agente (ya sea un usuario o un programa informático) accede a un recurso, ofrezca una descripción estructurada del recurso a través del uso de esquemas estandarizados (*Dublin Core*, *LOM*, etc.).

4. Incluir enlaces hacia otros URI para que sea posible descubrir más cosas.

Finalmente, tal como se describió anteriormente, con el fin de lograr la vinculación entre datos, el W3C propone que una vez que un agente accede a la descripción formal y estructurada de un recurso, ésta incluya enlaces hacia otros sets de datos o recursos que pudieran resultar de interés.

Con el fin de establecer un orden de prioridad y ofrecer una esca-





la que oriente las políticas de datos, el W3C propuso un escalafón de cinco estrellas según sea la infraestructura de LOD expuesta:

ESTRELLAS	NIVEL	DESCRIPCION	EJEMPLO
★	publica tus datos en la Web (con cualquier formato) y bajo una licencia abierta	Los datos están disponibles en la web. No resulta legibles para herramientas informáticas de análisis o para su reúso. Esta orientada principalmente para ser utilizada como usuario único por personas	Pdf
★★	publicalos como datos estructurados (ej: Excel en vez de una imagen de una tabla escaneada)	Los datos están disponibles en la web y pueden ser analizados por aplicaciones informáticas que soporten los formatos utilizados. Los datos pueden ser reusados en determinados contextos, bajo ciertas condiciones y en un horizonte temporal acotado.	excel
★★★	usa formatos no propietarios (ej: CSV en vez de Excel)	Los datos están disponibles en la web y pueden ser analizados por aplicaciones informáticas que soporten los formatos utilizados. Los datos reúnen condiciones óptimas para su uso en diferentes entornos y condiciones a través del tiempo.	CSV, XML, TXT
★★★★	usa URIs para denotar cosas, así la gente puede apuntar a estas	Permite además la identificación interoperables y persistente de datos y porciones de datos.	RDF
★★★★★	enlaza tus datos a otros datos para proveer contexto	Permite además realizar operaciones de análisis, síntesis o descubrimiento de nuevas propiedades a características. Permite procesos de enriquecimiento de datos.	RDF, punto de consulta SPARQL

La escala presentada puede ser consultada en <http://5stardata.info/es/>. Resulta recomendable consultar la lista completa de ejemplos presentados en el sitio y en especial la enumeración de ventajas y objetivos de cada nivel.

Tal como es posible apreciar, afortunadamente hay varios caminos hacia los LOD y también diferentes instancias.

Con el fin de ofrecer un contexto más amplio al marco de LOD, a continuación se presentará el marco general de datos abiertos y su contexto de emergencia en relación con la noción de gobierno y esfera pública para luego realizar una descripción de las posibilidades de los datos abiertos en el ámbito educativo.





### 3. Datos para gobierno abierto

En el ámbito de gobierno, existe un importante número de iniciativas tendientes a lograr la exposición pública de datos a través de medios digitales. Para la presente síntesis resulta interesante considerar la propuesta de lineamientos elaborada a partir de un taller de reflexión y elaboración realizado en el año 2007 en Sabastopol (Estados Unidos) <sup>7</sup> en el que participaron especialistas y entusiastas del gobierno abierto. Los lineamientos fueron cristalizados en ocho principios que, en la actualidad, se han convertido en un estándar de facto para el análisis y evaluación de servicios basados en datos en contextos de gobierno.

#### Ocho principios para datos gubernamentales abiertos

**1. Completos:** todos los datos públicos han de estar disponibles. Son públicos todos los datos con excepción de los que tienen limitaciones por cuestiones de seguridad, confidencialidad

o privilegios especiales.

**2. Primarios:** los datos han de estar disponibles con el máximo nivel de desagregación posible, no agrupada o modificada.

**3. A tiempo:** los datos estarán disponibles lo más rápidamente posible tras su obtención para que no pierdan su valor.

**4. Accesibles:** los datos estarán disponibles para la gama más amplia de usuarios y el espectro más amplio de propósitos.

**5. Tratamiento automatizado:** los datos se suministrarán de forma que se permita su tratamiento automatizado. Los datos deben poder ser descargados en forma completa, sin procesamiento, por medio de un mecanismo automatizado.

**6. No discriminación:** los datos estarán disponibles para cualquier persona, sin requisito previo de registro.

**7. Formato no propietario:** los datos estarán disponibles en dicho formato.

**8. Sin licencia:** los datos no estarán sometidos a ninguna regulación de patente, marca registrada o secreto comercial.

Es posible apreciar cierta convergencia entre la agenda temprana de la web como proyecto (expresada a mediados de las década de los 90) y la búsqueda de vías para la expansión y consolidación de procesos de institucionalización y gobierno en ambientes digitales en red. En tanto la agenda del W3C se expresa a través de propuestas de orden técnico, la agenda de datos abiertos se encuentra orientada a promover políticas y prácticas sociales que asignan un rol relevante para los estados en la consolidación de infraestructuras de servicios basados en datos. Se trata de un procedimiento extender las capacidades del estado hacia el territorio digital a través de la provisión de infraestructuras en condiciones de transparencia, equidad y uniformidad. En el caso de las infraestructuras de servicios basados en datos, podría sintetizarse en la capacidad de ofrecer datos codificados de manera tal que su uso, tratamiento y gestión pueda ser realizado:

- utilizando diferentes infraestructuras informáticas;
- según diferentes comunidades de prácticas;
- en función de propósitos diversos;
- a través del tiempo.

Al igual que las autopistas o los trenes, para alcanzar un modelo de explotación equiparable a la noción de infraestructura física, los datos deben ser provistos de manera persistente, uniforme y formal. Es decir, deben ofrecerse en condiciones de previsibilidad constante para que otros actores sociales puedan diseñar y articular servicios basados en una determinada infraestructura, en este caso, infraestructuras digitales de datos en el ámbito educativo.

- utilizando diferentes aplicaciones informáticas;

<sup>7</sup> Open Government Working Group Disponible en: [https://public.resource.org/open\\_government\\_meeting.html](https://public.resource.org/open_government_meeting.html) Fecha de consulta 2015-03-09





## 4 El marco de datos abiertos en educación

Los sistemas educativos en particular y el conjunto de prácticas sociales involucradas en los procesos educativos en general generan y utilizan un amplio espectro de tipos de datos. La edición 2014 del Open Education Handbook<sup>8</sup> distingue entre los siguientes tipos de datos:

- **Datos abiertos (open data):** datos abiertos son los datos que cualquiera puede utilizar libremente, reutilizar y redistribuir. Eventualmente los datos podrán estar sujetos a la obligación de atribuir autoría o responsabilidad intelectual y/o compartir los eventuales sub-productos resultantes del uso de los datos en las mismas condiciones.

- **Datos abiertos en contextos educativos (Open data in an educational context):** Datos cuyo uso en tanto datos abiertos sólo es válido para propósitos educativos.

- **Datos educativos abiertos (Open education data):** Datos generados en el ciclo de gestión de instituciones educativas y/o prácticas educativas expuestos a través de licencias abiertas o mecanismos congruentes con los lineamientos de datos abiertos. Son datos propios de los procesos de administración en el ámbito educativo que pueden ser utilizados para procesos de análisis, síntesis evaluación, toma de decisiones, etc.

- **Datos educacionales abiertos (Open educational data):** Comprende el conjunto de los Datos educativos pero de manera más amplia ya que incluye datos científicos abiertos relativos al ámbito educativo.

- **Recursos educativos abiertos (Open educational resources):** recursos digitales auto-contenidos explícitamente orientados a procesos didácticos y que incluyan un a mención de licencia abierta de uso explícita.

Muy probablemente, este último tipo de datos abiertos resulte predominante en el caso de los portales miembro de RELPE.

8 LinkedUp Project (2014) Open Education Handbook. Open Education Working Group. Disponible en: <http://education.okfn.org/handbook/>. Consultar también glosario: [http://en.wikibooks.org/wiki/Open\\_Education\\_Handbook/Glossary](http://en.wikibooks.org/wiki/Open_Education_Handbook/Glossary)





## 5 Escenarios diversos requieren soluciones diversas

El ámbito de RELPE está conformado por una red plural de portales que, como tal, participan y conviven un conjunto de diversidades:

- **Diversidad de modelos de gestión**
- **Diversidades en las políticas de descripción de recursos**
- **Diversidad de recursos educativos**
- **Diversidad tecnológica**

Considerando las diversidades antes mencionadas como un atributo inherente a la red de portales, es conveniente considerar que existe también una pluralidad de caminos para la adopción de políticas de datos abiertos. La constante a considerar sigue siendo que la web está basada en mecanismos para identificar

datos (métodos específicos para denominar e identificar de manera unívoca cada dato o conjunto de datos), mecanismos para localizar datos (métodos que permitan localizar y acceder a un dato en base a su anterior identificación) y mecanismos para representar datos (esquemas y herramientas que permitan, una vez identificados y localizados los datos, codificar y representar a los mismos para su uso y comunicación). En virtud de lo anterior, siguen siendo relevantes las recomendaciones del W3C con respecto a las mejores condiciones para el crecimiento y desarrollo de una web de datos, es decir, uso de:

- uso de identificadores únicos formales, persistentes y uniformes para mantener condiciones de interoperabilidad en el contexto de la web (ejemplo: URI);

- uso de protocolos abiertos y extendidos para la localización e interacción con recursos (ejemplos: HTTP, URL);
- uso de esquemas formales y explícitos para la codificación de datos y contenidos (ejemplo: XML).

### 5.1 LOS DATOS ABIERTOS ENLAZADOS SON UN CAMINO

En la propuesta técnica inicial de datos enlazados ya estaba presente la idea de que es posible establecer una gradualidad en las políticas de codificación y exposición de datos, es decir, existen varios niveles y métodos posibles para exponer LOD. En este sentido, es posible considerar que los LOD son una camino, una direccionalidad a seguir y no necesariamente un punto de llegada. En general, los portales miembro de RELPE se encuentran en condiciones de exponer datos abiertos, quizás es el momento de establecer planes a mediano plazo que contemplen hitos graduales.

### 5.2 IMPACTO TECNOLÓGICO ACOTADO

A diferencia de otros esquemas que amplían o mejoran la infraestructura de servicios de un portal, la implementación de un servicio de LOD no implica necesariamente modificar la infraestructura tecnológica en uso, pero sí requiere la implementación de una política de consolidación de datos y luego la selección de una estrategia de exposición de los mismos. Como se planteó anteriormente, la exposición de LOD supone el uso de al menos dos esquemas para la gestión de datos: uno primario, a través de la herramienta en uso para la gestión de datos y contenidos del portal; y otro secundario, que se alimenta a través de procesos de importación asincrónica de los datos generados y provistos por el esquema primario.





### 5.3 UNA RED DE DATOS ABIERTOS ENLAZADOS

Son muchos los factores que permitieron y coadyuvaron para que el espacio digital abierto de publicación de contenidos científicos se transformara en un ambiente de comunicación global y una herramienta radical de transformación de las prácticas sociales. Uno de los aspectos centrales de la web está dado por una dinámica de ciertos rasgos propios de un ecosistema: ningún actor por sí mismo puede cambiar la dinámica de la web y todas las acciones eventualmente pueden cambiar en algún grado las propiedades de la web como red. Siguiendo esta línea de pensamiento, cabe señalar que la iniciativa de LOD constituye una apuesta por un ecosistema de datos, es decir, depende de que una multitud de actores no coordinados adopten prácticas similares y su impacto debe ser evaluado en términos sistémicos y no en términos de objetivos y logros propios. Ofrecer LOD en el campo educativo es una apuesta para aportar a un coro polifónico de actores que aspira a ofrecer una melodía.

### 5.4 DEL BANCO DE RECURSOS A LAS HERRAMIENTAS DE DESCUBRIMIENTO

Una de las agendas más prometedoras de los LOD está dada por las facilidades que ofrece para realizar consultas de todo tipo a un proveedor de datos dado, como se describió anteriormente. Por ejemplo, a través de un punto de consulta SPARQL es posible pedir una lista detallada de los recursos educativos disponibles para una determinada temática, pero también pedir una síntesis que describa cómo están compuestos los recursos educativos orientados a un nivel educativo específico agrupados según cantidad de visitas o que hubieran sido financiado por una determinada institución, y así en adelante. Los resultados de estas consultas se pueden integrar dinámicamente y en tiempo real con los resultados propios de cada portal. Este tipo de capacidades permite acercar las funcionalidades de los bancos de recursos hacia las provistas por las denominadas herramientas de descubrimiento, que se tratan de sitios o aplicaciones que exploran ambientes abiertos de datos buscando descubrir recursos que

puedan ser significativos a partir de una semilla de búsqueda dada; por ejemplo, qué recursos hay acerca del “dulce de leche”, considerando sus distintas denominaciones regionales (manjar, manjar blanco, arequipe o cajeta). De esta manera, si un catálogo de recursos está basado en su capacidad para describir, repre-

sentar y buscar en un universo propio y acotado, las herramientas de descubrimiento permiten replicar estas funcionalidades pero en ambientes abiertos y diversos, facilitando el descubrimiento de nuevas asociaciones conceptuales, propiedades, aspectos, puntos de vista y recursos.







## 6 Anexos

### 6.1 Sitios web de interés

- **5 ★ Datos Abiertos:** Sitio que documenta la capa de 5 estrellas propuesta por Tim Berners-Lee. Además ofrece acceso a ejemplos y modelos de políticas para organizaciones y tomadores de decisión. <http://5stardata.info/es/>
- **Dbpedia:** Proyecto orientado a estructurar y exponer los contenidos de la Wikipedia como datos enlazados. <http://dbpedia.org/>
- **Europeana:** Espacio que expone los catálogos de datos reunidos en el proyecto Europeana. <http://labs.europeana.eu/api/linked-open-data/data-downloads/>
- **Iniciativa de datos abiertos en América Latina:** Proyecto liderado por CEPAL, OEA, Fundación AVINA orientado a fomentar la investigación, el uso y la apropiación de los datos abiertos

en América Latina. <http://idatosabiertos.org/>

- **Linked Data - Connect Distributed Data across the Web:** Sitio que reúne enlaces, recursos, manuales y recursos en general acerca de datos enlazados. <http://linkeddata.org/>
- **Linked Education:** Blog dedicado a las iniciativas de datos abiertos enlazados en el ámbito educativo. <https://linkededucation.wordpress.com/>
- **LOD cloud diagram:** Página mantenida por Richard Cyganiak y Anja Jentzsch que ofrece un gráfico actualizado del estado de demografía y grado de vinculación de los datos abiertos enlazados. <http://lod-cloud.net/>
- **LODStats:** Proyecto orientado a reunir y mantener estadísticas acerca del estado de la iniciativa de datos abiertos enlazados. <http://stats.lod2.eu/>

- **Proyecto Linkedup:** Proyecto de la Unión Europea orientado a la coordinación y dar soporte a iniciativas de datos abiertos enlazados en el ámbito educativo. <http://linkedup-project.eu/>
- **State of the LOD Cloud 2014:** Informe desarrollado por Max Schmachtenberg, Christian Bizer y Heiko Paulheim orientado a describir el estado y alcance de los datos abiertos enlazados. <http://linkeddatacatalog.dws.informatik.uni-mannheim.de/state/>

### 6.2 Documentos y marcos elaborados por el W3C

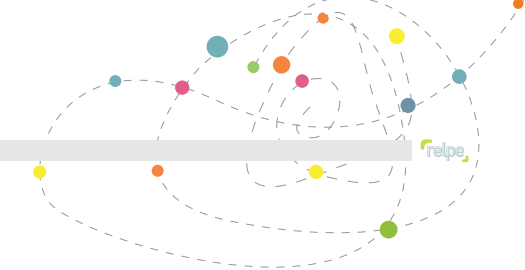
- Arquitectura de la web <http://www.w3.org/TR/webarch/>
- Guía Breve de Linked Data <http://www.w3c.es/Divulgacion/Guias-Breves/LinkedData>
- Library Linked Data Incubator Group Final Report <http://www.w3.org/2005/Incubator/llid/XGR-llid-20111025/>

- Linked Data Current Status [http://www.w3.org/standards/techs/linkeddata#w3c\\_all](http://www.w3.org/standards/techs/linkeddata#w3c_all)
- Linked Data <http://www.w3.org/DesignIssues/LinkedData.html>
- Linked Data <http://www.w3.org/standards/semanticweb/data>
- RDF Current Status [http://www.w3.org/standards/techs/rdf#w3c\\_all](http://www.w3.org/standards/techs/rdf#w3c_all)
- Resource Description Framework <http://www.w3.org/RDF/>
- The SPARQL (1.1) [http://www.w3.org/2009/sparql/wiki/Main\\_Page](http://www.w3.org/2009/sparql/wiki/Main_Page) <http://www.w3.org/TR/sparql11-overview/>

### 6.3 Colecciones de datos en educación

- Education Data: Datos abiertos enlazados publicados por el gobierno del Reino Unido. Corresponde a conjuntos de datos consolidado y estático elaborado en 2009. <http://education.data.gov.uk/>





- Herramientas para convertir datos en plataformas de datos abiertos enlazados: Conjunto de herramientas recopilados por Linked Universities que reúne herramientas para convertir datos desde casi cualquier formato (excel, rss, XML, RDBs, OAI-PMH, citas bibliográficas y más) en plataformas de datos abiertos enlazados. <http://linkeduniversities.org/lu/index.php/tools/index.html>
- **Linked Education Blog:** Relevamiento de catálogos de datos y puntos de consulta SPARQL mantenido por el sitio Linked Education. <https://linkededucation.wordpress.com/data-models/data-sets/>
- **Linked Education Cloud:** Repositorio y catálogo de colecciones de datos abiertos en el ámbito educativo. <http://data.linkededucation.org/linkedup/catalog/>
- **Linked Universities:** Convenio entre universidades europeas orientado a exponer sus datos públicos como datos abiertos enlazados. <http://linkeduniversities.org/>

- **Mejoratuescuola:** Sitio web que permite realizar comparaciones entre datos cuantitativos y cualitativos correspondientes al sistema educativo mexicano. <http://www.mejoratuescuola.org/>
- **Open Discovery Space:** Proyecto orientado a construir una infraestructura multilíngue orientada a recursos de aprendizaje abiertos. Dispone de un catálogo de recursos de aprendizaje expuesto a través de un esquema de datos abiertos enlazado. <http://opendiscoveryspace.eu/> <http://data.opendiscoveryspace.eu/sparql.tpl>
- **Qedu:** Sitio web que permite realizar comparaciones entre datos cuantitativos y cualitativos correspondientes al sistema educativo brasileño. <http://www.qedu.org.br/>

## 6.4 ALGUNAS LECTURAS RECOMENDADAS

- Bizer, Christian, Heath, Tom, Berners-Lee, Tim (2009) Linked Data - The Story So Far. Int. J. Semantic Web Inf. Syst. 5(3): 1-22. Disponible en: <http://tomheath.com/papers/bizer-heath-berners-lee-ijswis-linked-data.pdf>
- Keßler, Carsten, d'Aquin, Mathieu, Dietze, Stefan (2009) Linked Data for Science and Education. Disponible en: [http://www.semantic-web-journal.net/system/files/swj326\\_0.pdf](http://www.semantic-web-journal.net/system/files/swj326_0.pdf)
- Linked data Connecting and exploiting big data (2012). Fujitsu. Disponible en: [http://www.fujitsu.com/uk/Images/Linked-data-connecting-and-exploiting-big-data-\(v1.0\).pdf](http://www.fujitsu.com/uk/Images/Linked-data-connecting-and-exploiting-big-data-(v1.0).pdf)
- LinkedUp Project. Open Education Working Group (2014) Open Education Handbook. Disponible en: <http://education.okfn.org/handbook/>

- Tim Berners-Lee, Wendy Hall, James A. Hendler, Kieron O'Hara, Nigel Shadbolt and Daniel J. Weitzner (2006), "A Framework for Web Science", Foundations and Trends® in Web Science: Vol. 1: No. 1, pp 1-130. <http://dx.doi.org/10.1561/1800000001>
- Tom Heath, Christian Bizer (2011) Evolving the Web Into a Global Data Space. Morgan & Claypool Publishers. Disponible en: <http://linkededatabook.com/editions/1.0/>
- USA. Executive Office of the President (2014). The Big Data and Privacy Review. Disponible en: [http://www.whitehouse.gov/sites/default/files/docs/big\\_data\\_privacy\\_report\\_5.1.14\\_final\\_print.pdf](http://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_5.1.14_final_print.pdf)





## Explotación cooperativa de recursos educativos

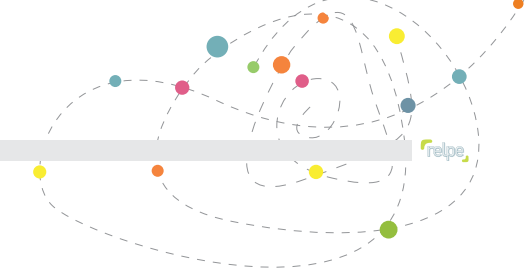
Consultor:

**Diego Ferreyra**



Esta publicación está basada en el informe de consultoría realizado en el contexto de implementación de una propuesta técnica y metodológica para la explotación cooperativa de recursos educativos en el marco de la red RELPE. Dicho proceso fue realizado durante el segundo semestre del año 2010





## CONTENIDO

### Explotación cooperativa de recursos educativos

#### RELPE y el intercambio de contenidos

- Descripción de la Arquitectura de red centralizada para el intercambio de contenidos
- Dificultades con la solución técnica adoptada
- Hacia la explotación cooperativa de contenidos

#### Condiciones que promueven la cooperación en la Web

- utilización de marcos formales de descripción
- uso de estándares abiertos
- esquemas autodescriptivos basados en XML para la estructuración de los datos.
- esquemas de metadatos de bajo nivel para la estructuración de los datos
- desarrollo de productos y servicios basados en la explotación del valor agregado emergente
- La adopción de modelos de cooperación

#### OAI-PMH como modelo para la explotación cooperativa de recursos educativos

- Descripción del protocolo OAI PMH
- Dinámica del protocolo OAI-PMH

#### La red RELPE y el protocolo OAI-PMH

- Escalabilidad





Sustentabilidad

Construcción de capacidades

Modelo polivalente

Cooperación entre actores asimétricos

Generación de valor agregado

#### **RELPE: Implementación del modelo de explotación cooperativa de recursos educativos utilizando OAI-PMH**

Escenario 1: Capacidad instalada

Escenario 2: Articulación tecnológica

Escenario 3: Migración de metadatos

Escenario 4: OAI estático

La prueba piloto

Del intercambio de contenidos a la explotación cooperativa

#### **Guía de redacción de fuentes de datos XML para Servicios OAI basados en fuentes estáticas**

Datos descriptivos del repositorio

Set de metadatos de recursos

#### **Proveedor de datos OAI satélite**

Tutorial de implementación

Características funcionales de la herramienta

Opciones de configuración

Opciones de configuración obligatorias

Opciones de configuración estándar

#### **Requerimientos de base para la implementación**

Instalación

Tabla de datos

#### **Algunos recursos y direcciones útiles**

Documentos y tutoriales

Herramientas y fuentes OAI-PMH





## EXPLOTACIÓN COOPERATIVA DE RECURSOS EDUCATIVOS RELPE Y EL INTERCAMBIO DE CONTENIDOS

La Red Latinoamericana de Portales Educativos <sup>1</sup> (RELPE) está conformada por los portales educativos nacionales de 19 países de Latinoamérica. Uno de los principales objetivos de RELPE desde sus inicios en 2004 fue propiciar la libre circulación de contenidos educativos entre los miembros de la red. Para cumplir tal fin, se desarrolló una solución basada en una arquitectura de red centralizada, con una lógica jerárquica, en la que los nodos dependen de un ente centralizador que concentra y distribuye la información.

Esta solución necesita de criterios comunes de catalogación y empaquetamiento de los contenidos <sup>2</sup>, al igual que el desarrollo de una metodología de articulación que posibilite la interconexión e interoperabilidad de los nodos <sup>3</sup>.

## DESCRIPCIÓN DE LA ARQUITECTURA DE RED CENTRALIZADA PARA EL INTERCAMBIO DE CONTENIDOS

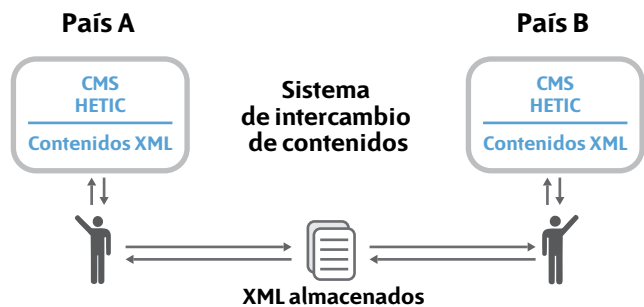
El sistema de intercambio de contenidos requiere que cada país genere y almacene sus contenidos en su CMS <sup>4</sup> y, a su vez, realice, de ser necesario, modificaciones al mismo para poder catalogar sus contenidos según las normas de catalogación de RELPE y para empaquetarlos en el formato especificado. Los pasos de este esquema se pueden resumir de la siguiente manera:

**El país A** genera y almacena sus contenidos en su CMS, los cataloga y empaqueta de acuerdo con las especificaciones de RELPE. Si el CMS no tiene esa capacidad se debe desarrollar un conector llamado HETIC <sup>5</sup> que se integra con el CMS y permite generar el XML RELPE.

Luego de ello los editores, de forma manual, eligen los recursos y los suben a la herramienta de intercambio de RELPE.

**El país B** puede ingresar a la herramienta de intercambio y seleccionar manualmente los contenidos que le interesara anexar a su portal.





### DIFICULTADES CON LA SOLUCIÓN TÉCNICA ADOPTADA

Durante los años siguientes a la puesta en marcha de esta solución, RELPE trató de impulsar este sistema de intercambio entre los portales miembros, que tuvo como resultado una baja tasa de aceptación. Esto se debió, principalmente, a que el modelo de intercambio propuesto solo sirve para operar dentro de RELPE y restringe las posibilidades de articulación con otros actores o comunidades.

En el transcurso de la experiencia se detectaron ciertas dificultades tanto de índole tecnológica como de gestión y culturales:

- Los portales debían invertir en desarrollos tecnológicos para poder generar los archivos en el formato para el intercambio.
- El modelo de descripción de recursos adoptado generaba conflictos con las catalogaciones locales, disminuyendo la riqueza propia de la diversidad cultural existente dentro de RELPE. Además, se debe señalar que este modelo solo les servía dentro de la red.
- La adopción creciente de estándares tanto para la catalogación como para el empaquetado, como también la aceptación de los protocolos abiertos de intercambio, llevaron a RELPE a examinar sus estrategias en cuanto a las tecnologías adoptadas, para no quedar aislado de un mundo cada vez más conectado.
- El sistema de intercambio no fue evolucionando con internet.

Con el surgimiento de la Web 2.0 y las redes sociales se tornó cada vez más difícil sostener un esquema de cooperación cerrado.

### HACIA LA EXPLOTACIÓN COOPERATIVA DE CONTENIDOS

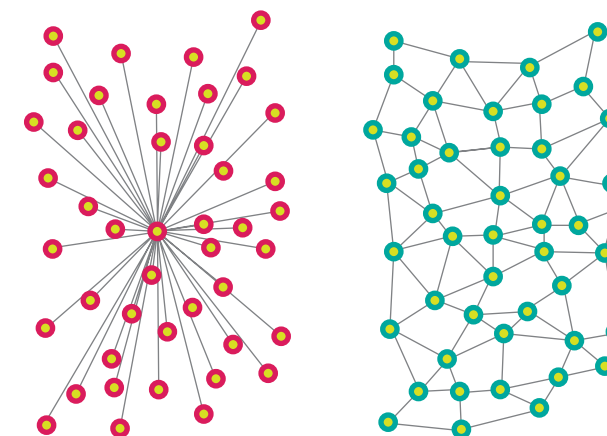
En los procesos de cooperación entre actores autónomos se pueden reconocer los rasgos característicos de los modelos y dinámicas de cooperación, de los cuales podemos destacar, por un lado, la voluntad de articular contextos diversos y modelos heterogéneos, y por otro la voluntad de lograr mejoras y construcciones comunes.

Si bien estos rasgos están presentes en RELPE, las dificultades acarreadas por el modelo de explotación de contenidos complicaban la puesta en práctica de la cooperación.

Teniendo en cuenta las restricciones mencionadas, a partir de 2010 los miembros de RELPE hicieron un replanteo acerca del modelo de intercambio, orientando el cambio hacia uno descentralizado, propiciando de esta manera el desarrollo de una **arquitectura abierta de explotación cooperativa de contenidos**.

Este tipo de modelo propicia el uso compartido de recursos a través del desarrollo de las capacidades locales y de los mecanismos para compartir información. En una arquitectura descentralizada no existe un nodo concentrador sino que cualquier nodo puede

conectarse con cualquier otro y posee una lógica de interacción horizontal; así, la socialización de la información es más eficiente y las redes pueden crecer y conectarse con otras.



Arquitectura centralizada      Arquitectura descentralizada

### CONDICIONES QUE PROMUEVEN LA COOPERACIÓN EN LA WEB

Es posible establecer ciertas condiciones de base que, en general, promueven o facilitan los procesos de cooperación con base tecnológica en el contexto de la Web. Entre las mencionadas condiciones es posible enumerar las siguientes:





## UTILIZACIÓN DE MARCOS FORMALES DE DESCRIPCIÓN

Es posible tanto la adhesión a esquemas de metadatos existentes, como también establecer un esquema de metadatos propio, ya sea creando uno o estableciendo un perfil de metadatos (profile).

Esta última vía resulta interesante y por demás recomendable, ya que permite expresar aspectos locales que muchas veces no están previstos por un determinado esquema, sin alejarse o abandonar la comunidad de prácticas de referencia que mantiene el esquema de metadatos.

## ESQUEMAS DE METADATOS DE BAJO NIVEL PARA LA ESTRUCTURACIÓN DE LOS DATOS

El uso de esquemas de metadatos como Dublin Core no cualificado permite simplificar las condiciones dinámicas de explotación cooperativa maximizando las posibilidades de su reutilización. Considerando los esquemas de metadatos como una lengua de intercambio, podría decirse que, al igual que una lengua franca,

empobrecen sensiblemente la riqueza descriptiva o expresiva a favor de maximizar las posibilidades de la traducción.

El uso de los esquemas de metadatos en tanto estructuras modulares de descripción permite describir los recursos educativos desde distintos puntos de vista, o con distintos grados de especificidad, permitiendo la extensibilidad del modelo de descripción.

## DESARROLLO DE PRODUCTOS Y SERVICIOS BASADOS EN LA EXPLOTACIÓN DEL VALOR AGREGADO EMERGENTE

El modelo de explotación fomenta la articulación de distintos recursos, que agrega valor, más que los recursos por separado. La articulación de los nodos promueve la percepción de que existen beneficios concretos que solo pueden concebirse a partir de la cooperación entre los nodos, y que por separado resultarían imposibles.

Estos productos y servicios constituyen incentivos críticos para sobrellevar los costos y/o adaptaciones que requieren, con distintos grados de impacto según sea el modelo adoptado, los procesos de cooperación.

## LA ADOPCIÓN DE MODELOS DE COOPERACIÓN

En la adopción de modelos de cooperación se deben privilegiar aquellos que permitan mantener condiciones de autonomía local y de interoperabilidad global para los nodos, ya que son un factor significativo para promover y adoptar un modelo de cooperación.

Este tipo de prácticas facilita las dinámicas de cooperación, ya que permite basarse en agendas de discusión compartidas por varias comunidades de prácticas y expresadas en los modelos de descripción formalizados en sus respectivos esquemas de metadatos.

## USO DE ESTÁNDARES ABIERTOS

La adopción de estándares abiertos facilita el desarrollo de marcos de debate, y también permite compartir el conocimiento construido por cada comunidad de práctica en la elaboración de sus esquemas de metadatos. Los estándares abiertos posibilitan

el acceso no solo a los recursos, sino también a los modelos y concepciones que dan base a la gestión de los mismos. Un estándar abierto consiste en una serie de especificaciones públicas, completamente mantenidas por una entidad neutral y sin ánimo de

lucro, cuya propiedad intelectual se ofrece de forma libre de regalías y sin restricciones en cuanto a su uso y reutilización.

## ESQUEMAS AUTODESCRIPTIVOS BASADO EN XML PARA LA ESTRUCTURACIÓN DE LOS DATOS

El uso de XML<sup>6</sup> para exponer los metadatos de los recursos educativos o los recursos mismos facilita la reutilización y portabilidad de los mismos. Esto se debe a que el XML es fácil de procesar por ser texto plano, y también a que los documentos estructurados en XML incluyen referencias a reglas que rigen la conformación del propio documento XML, lo cual facilita su interpretación.

## ESQUEMAS DE METADATOS DE BAJO NIVEL PARA LA ESTRUCTURACIÓN DE LOS DATOS

El uso de esquemas de metadatos como Dublin Core no cualificado<sup>7</sup> permite simplificar las condiciones dinámicas de explotación cooperativa maximizando las posibilidades de su reutilización. Considerando los esquemas de metadatos como una lengua de intercambio, podría decirse que, al igual que una lengua franca,







empobrecen sensiblemente la riqueza descriptiva o expresiva a favor de maximizar las posibilidades de la traducción.

El uso de los esquemas de metadatos en tanto estructuras modulares de descripción permite describir los recursos educativos desde distintos puntos de vista, o con distintos grados de especificidad, permitiendo la extensibilidad del modelo de descripción.

## DESARROLLO DE PRODUCTOS Y SERVICIOS BASADOS EN LA EXPLOTACIÓN DEL VALOR AGREGADO EMERGENTE

El modelo de explotación fomenta la articulación de distintos recursos, que agrega valor, más que los recursos por separado. La articulación de los nodos promueve la percepción de que existen beneficios concretos que solo pueden concebirse a partir de la cooperación entre los nodos, y que por separado resultarían imposibles. Estos productos y servicios constituyen incentivos críticos para sobrellevar los costos y/o adaptaciones que requieren, con distintos grados de impacto según sea el modelo adoptado, los procesos de cooperación.

## LA ADOPCIÓN DE MODELOS DE COOPERACIÓN

En la adopción de modelos de cooperación se deben privilegiar aquellos que permitan mantener condiciones de autonomía local y de interoperabilidad global para los nodos, ya que son un factor significativo para promover y adoptar un modelo de cooperación.

### CONDICIONES DE AUTONOMÍA LOCAL

*Las diferencias constituyen la identidad institucional de cada uno de los nodos participantes, y en gran medida establecen sus condiciones de legitimidad y autoridad con respecto a su propia comunidad.*

La cooperación como dinámica, la interrelación entre actores en base a acuerdos formales o a circunstancias por las cuales pudieran compartir objetivos u agendas técnicas, metodológicas o políticas, no invalida el hecho de que, tal como se ha visto antes, mantienen entre sí diferencias. Son las diferencias que justamente constituyen la identidad institucional de cada uno de los nodos

participantes, y que en gran medida establecen sus condiciones de legitimidad y autoridad con respecto a su propia comunidad.

Esas diferencias son las que aportan componentes significativos en términos locales para los actores que conforman la base y audiencia de cada portal o nodo cooperante. En tal sentido, resulta recomendable adoptar esquemas de cooperación que permitan mantener las condiciones de autonomía local de cada uno de los nodos participantes, o al menos que minimicen el impacto de las restricciones o requerimientos de los procesos de cooperación en las dinámicas y servicios en el contexto local de gestión de cada nodo.

Es indispensable no perder capacidades que reflejen las condiciones locales, ya que la condición para la adopción de la innovación está en directa relación con la capacidad de resultar significativa para los actores más inmediatos.

La autonomía puede definirse desde distintas facetas:

- **Autonomía de diseño:** libertad de cada actor sobre elecciones de implementación tecnológica, gerenciamiento de datos, funcionalidades, etc.
- **Autonomía de comunicación:** capacidad para decidir cuándo y cómo responder a requisitos de terceros.
- **Autonomía de ejecución:** en la que cada nodo puede realizar sus tareas sin interferir en la ejecución de tareas no locales.
- **Autonomía de participación:** capacidad para decidir cuánto de sus funciones, operaciones e información compartir.



## CONDICIONES DE INTEROPERABILIDAD GLOBAL

En un contexto de explotación abierta y cooperativa de contenidos existen ciertas condiciones que deben darse para lograr la cooperación y la colaboración. Básicamente deben garantizarse ciertas condiciones de autonomía local y condiciones de autonomía global.

Los intercambios a través de internet generan un escenario macro en el que las condiciones de interoperabilidad global son las que garantizan, más allá de los ciclos y condiciones de gestión local, los modelos y herramientas de articulación con otros actores y ambientes.

Interoperabilidad es la capacidad de dos entidades para trabajar conjuntamente con el propósito de realizar una tarea. Significa ampliar la visibilidad, participación y acceso a los sistemas de información.

La mayor ventaja es que no se precisa acuerdo previo entre las entidades que gestionan los sistemas de información para cooperar.

En efecto, con la interoperabilidad no se necesitan acuerdos pre-

establecidos sobre la gestión y el tratamiento de los recursos que se comparten. Esto facilita la existencia de pluralidad de métodos y estándares en los procesos de cada nodo, los cuales pueden conservar sus cualidades regionales únicas en vez de resignarlas para poder normalizar el intercambio.

A continuación se resaltan las ventajas en relación con la diversidad que cada nodo puede disfrutar en un ámbito de interoperabilidad.

Estas ventajas se consideran un punto de partida para la construcción en común:

- **Diversidad de modelos de gestión:** los nodos cooperantes implementan políticas y prácticas locales de gestión específicas, dinámicas y cambiantes con respecto a los recursos educativos.

Este tipo de cambios puede afectar al modo en que se producen o incorporan los recursos educativos, y también a las políticas de descripción, disponibilidad y acceso de los recursos.

- **Diversidades en las políticas de descripción de recursos:** más allá de los acuerdos establecidos para la adopción del esquema RELPE de descripción de recursos educativos, conviven y perviven diversas prácticas y políticas de descripción e identi-

ficación de recursos. Esta diversidad puede ser expresada tanto en la adopción de un determinado vocabulario controlado o un esquema descriptivo local específico, como en diferencias acerca de los métodos para la identificación unívoca y persistente de los recursos educativos.

- **Diversidad de recursos educativos:** En ocasiones, al tratarse de diferentes realidades, comunidades y necesidades, no siempre es posible establecer un acuerdo único acerca del estatus de recurso educativo de un determinado contenido u objeto digital. Se presenta entonces el desafío de un modelo que sea capaz de contener diversas concepciones acerca de la especificidad y definición de la noción de recurso educativo.

- **Diversidad tecnológica:** Cada uno de los nodos que participa una dinámica de cooperación con base tecnológica utiliza diferentes aplicaciones, dispositivos y plataformas tecnológicas para administrar sus recursos. En este sentido, un modelo de cooperación deberá poder articular esta diversidad, más allá de las tecnologías de base utilizadas por los nodos cooperantes.

Hasta aquí se ha intentado ofrecer una breve descripción de cier-

tos rasgos que definen el horizonte de posibilidades y requerimientos a la hora de establecer un modelo descentralizado para la explotación cooperativa de recursos educativos, en el contexto de una red de portales educativos como lo es RELPE.

En virtud de estos factores y consideraciones, durante el año 2010 se evaluó la conveniencia de adoptar un modelo de explotación cooperativa que pudiera soportar estándares de metadatos formales y abiertos para la descripción de recursos educativos, y cuyo nivel de especificidad, granularidad y complejidad no fuera alto. Asimismo, se estableció la necesidad, con respecto al modelo de explotación cooperativa, de respetar y garantizar condiciones de autonomía local en un marco de interoperabilidad tecnológica global.

*En el ámbito de las TIC, los estándares abiertos garantizan la interoperabilidad entre agentes emisores y receptores de la información, con independencia del software utilizado por ellos, avalando la no dependencia de un único fabricante, fomentando la competitividad en el mercado [y con ello el ahorro de costos] e incentivando la neutralidad tecnológica y la innovación.*



Establecidos estos parámetros se resolvió proponer la adopción del protocolo OAI-PMH como modelo técnico para la explotación cooperativa de recursos educativos en el marco de la red RELPE.

El uso del protocolo OAI-PMH<sup>8</sup> es una propuesta superadora con respecto al modelo anterior, ya que está basado en estándares abiertos.

En el ámbito de las TIC, los estándares abiertos garantizan la interoperabilidad entre agentes emisores y receptores de la información, con independencia del software utilizado por ellos, avalando la no dependencia de un único fabricante, fomentando la competitividad en el mercado (y con ello el ahorro de costos) e incentivando la neutralidad tecnológica y la innovación.

## OAI-PMH COMO MODELO PARA LA EXPLOTACIÓN COOPERATIVA DE RECURSOS EDUCATIVOS

La iniciativa de archivos abiertos se creó con la finalidad de desarrollar y promover normas de interoperabilidad para facilitar la difusión eficaz de los contenidos digitales. Actualmente promueven y forman parte de esta iniciativa universidades e instituciones de todo el mundo.

El Protocolo de OAI-PMH surge en el contexto de la búsqueda de una serie de instituciones y comunidades profesionales por construir alternativas técnicas y metodológicas que faciliten y dinamicen el acceso hacia los recursos digitales, en principio producidos por ellos mismos.

El origen de estas iniciativas, que pueden ser ubicadas temporalmente en los finales de la década de los 90 y principios de nuestro siglo, fue cristalizándose en la llamada Convención de Santa Fe, una reunión realizada en Nuevo México (Estados Unidos) en 1999, en la que participaron algunas instituciones y archivos científicos preocupados y abocados a la tarea de mejorar las alternativas para com-

partir e intercambiar información a través de un modelo federado.

En esta reunión se establecieron una serie de principios y directrices orientadas al desarrollo de un modelo que permitiera realizar búsquedas y consultas federadas, es decir, buscar a la vez en una serie de repositorios, proveedores de datos o archivos.

En este contexto la reunión centró sus objetivos en tres aspectos:

- Establecer un protocolo para las reglas de consulta y las posibles respuestas.
- Un esquema de metadatos para la descripción de los recursos.
- Un sistema para la identificación de los recursos.

En la actualidad, el protocolo OAI-PMH se ha convertido en una solución ampliamente reconocida para compartir, difundir y exponer recursos en internet, y constituye una estrategia utilizada tanto por iniciativas académicas como gubernamentales, no-gubernamentales y comerciales.

En términos de modelo, el marco OAI-PMH puede ser definido por tres aspectos articulares:

- **Descripción de recursos:** propone una catalogación poco compleja y por ello de fácil adopción, a través de la utilización del esquema de metadatos Dublin Core sin cualificar como línea base. Eventualmente, según sean los casos de implementación, es posible utilizar otros esquemas de metadatos más específicos, como por ejemplo LOM<sup>9</sup>.
- **Identificación unívoca de los recursos:** el marco recomienda una sintaxis de identificadores únicos para cada recurso en el contexto de cada repositorio. Los mencionados identificadores únicos pueden ser URL<sup>10</sup>, URI<sup>11</sup>, o cualquier cadena de caracteres que resulten únicas en el contexto de cada repositorio.
- **Protocolo regulador de los intercambios:** la especificación del mismo se encuentra orientada al intercambio de metadatos a través de un set de peticiones posibles y parámetros requeridos. Implica la especificación de una serie de reglas y cláusulas que se deben cumplir en el marco de un proceso de comunicación entre los repositorios.





## DESCRIPCIÓN DEL PROTOCOLO OAI-PMH

El protocolo OAI-PMH establece una arquitectura de comunicación cliente-servidor. Los procesos de intercambio se realizan a partir de un servicio ofrecido por un servidor o proveedor OAI y una aplicación informática cliente que solicita los datos.

Un proveedor OAI (proveedor de datos) es una aplicación informática conectada a un repositorio de recursos capaz de ofrecer servicios de comunicación de datos de acuerdo al protocolo OAI-PMH para terceros sobre los recursos mencionados. Esto implica entender las peticiones formales y acordes al protocolo y devolver resultados correspondientes a las peticiones.

Un cliente OAI (proveedor de servicios) es una aplicación informática capaz de realizar peticiones a un servidor o proveedor de datos OAI-PMH y entender y procesar los datos que el servidor le envía. Es decir: actúa como un agregador de los recursos ofrecidos por el proveedor OAI, utilizando los metadatos recolectados con el objetivo de incorporarles valor agregado y presentarlos a sus usuarios finales. OAI-PMH establece sus transacciones sobre el protocolo http<sup>12</sup>. Las peticiones se realizan dentro del esquema cliente servidor como parámetros GET o POST a través de URL. OAI-PMH soporta seis tipos de peticiones, llamadas verbos, y las respuestas están codificadas en XML.

Los seis verbos o acciones que se admiten para las peticiones son:

- 1. Identify:** petición utilizada para obtener información sobre el repositorio: nombre, versión del protocolo, datos del administrador, etc.
- 2. ListMetadataFormats:** petición utilizada para obtener la lista de formatos de metadatos que utiliza el servidor.
- 3. ListSets:** recupera la estructura de colecciones de un repositorio.
- 4. ListIdentifiers:** petición que devuelve solo los encabezados de los registros. Permite delimitar el rango de fechas a recuperar.
- 5. ListRecords:** petición que devuelve los datos completos de los registros. Permite delimitar el rango de fechas a recuperar.
- 6. GetRecord:** petición que recupera un registro concreto. Utiliza dos argumentos: identificador del registro pedido y especificación del formato de metadatos en que se debe entregar.

Analizando la acotada variedad de acciones que permite el protocolo, es posible descubrir el grueso de las definiciones conceptuales que dan marco al protocolo. De esta manera es posible ver conceptos como identificadores, esquemas de metadatos, registros, colecciones y repositorios.

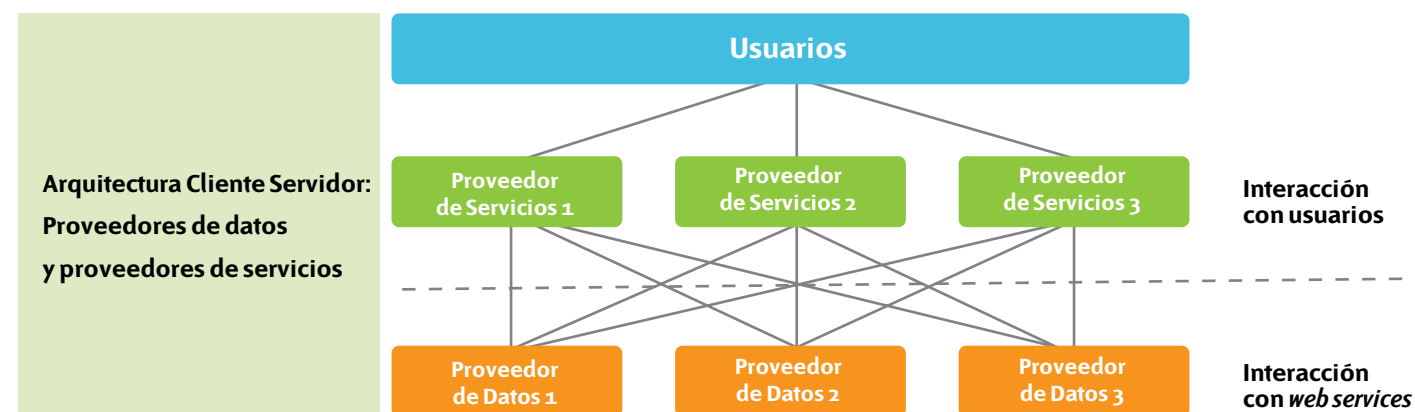
A continuación se describe cada uno de los conceptos y las posibles acciones asociadas en el marco del protocolo.

### REPOSITORIO

Un repositorio es un proveedor de servicios de metadatos a través del protocolo OAI-PMH. Es decir que es un servicio implementado en un servidor que expone metadatos que describen recursos, según sean las peticiones que recibe. La petición asociada a este concepto es **Identify**, que permite conocer información acerca del propio repositorio.

### COLECCIONES

En el contexto de cada repositorio es posible definir colecciones o agrupamientos de recursos. Estas definiciones son arbitrarias y pueden establecerse con criterios de todo tipo: pueden ser temáticos, según audiencias, según el formato de los recursos, según su





orden de ingreso al repositorio o cualquier otro esquema según el cual se desee organizar o clasificar los recursos. También es posible definir una única colección que comprenda todo el repositorio.

Se utilizan para acotar (filtrar) partes específicas de un repositorio, para poder solicitar los registros de un repositorio, pero solo los que pertenecen a una determinada colección o agrupamiento. La petición asociada a este concepto es ListSets, que permite consultar y conocer las colecciones que conforman el repositorio.

### IDENTIFICADORES

Un identificador único refiere de forma unívoca un recurso en el contexto de un repositorio; se utiliza para la extracción y consulta de metadatos que describen el recurso, para conocer la última fecha en que fue modificado el recurso o las colecciones o agrupaciones a las que pertenece el recurso. Este identificador puede ser una URL, URI o cualquier cadena de caracteres que resulte única en el contexto de cada repositorio. Los identificadores son utilizados en varios tipos de peticiones, pero una petición directamente vinculada a este concepto es **ListIdentifier**, que permite consultar la lista de encabezados de registros de un repositorio según un determinado set, o a los registros que hubieran sido creados o

modificados partir de una determinada fecha.

### REGISTROS

Cada recurso en el contexto del repositorio es descrito como un ítem compuesto por un identificador único y los metadatos que describen el recurso. Cada uno de los conjuntos de metadatos que describen el recurso es un registro. De esta manera, uno o más esquemas de metadatos pueden describir un mismo recurso, por ejemplo utilizando Dublin Core, LOM y el esquema RELPE, pero el recurso seguiría teniendo un identificador único.

Cada registro se compone de los metadatos que describen el registro propiamente dicho y de un componente llamado cabecera (header), que contiene el identificador único del recurso en el contexto del repositorio, una indicación temporal acerca de la última fecha de modificación o creación de recurso (estampilla de tiempo del recurso) y la indicación de las colecciones o agrupaciones (sets) a las que pertenece el recurso. Esto permite a las aplicaciones que consultan un repositorio conocer, antes de procesar los metadatos de los recursos, la cantidad de recursos disponibles por colección o agrupamiento o la cantidad de recursos que fueron modificados o creados desde una determinada fecha.

Las peticiones asociadas a este concepto son dos, **GetRecord** y **ListRecords**. GetRecord permite ver los metadatos de un registro específico. En cuanto a ListRecords: este verbo permite recuperar los datos completos disponibles para cada registro.

### ESQUEMAS DE METADATOS

Los metadatos son una estrategia utilizada para denominar, describir y organizar entidades a través de una instancia diferenciada (meta, del griego: después, más allá de) de las entidades mismas de descripción, o sea, los recursos.

Los esquemas de metadatos, como podrían ser el Dublin Core, el LOM, el esquema RELPE, etc., son formalizaciones explícitas y estructuradas de elementos de descripción de los datos que una determinada institución o comunidad de prácticas considera conveniente utilizar para comunicarse con otras entidades (instituciones, otras comunidades, a veces hacia la propia comunidad).

Estas formalizaciones no siempre se corresponden necesariamente con las modelizaciones utilizadas cotidianamente en la gestión real de los datos. Bajo esta perspectiva, los esquemas de metadatos (schemas) son abordados como una modalidad de exposición y explotación de un modelo de metadatos subyacente,

que no necesariamente coincide con el esquema de metadatos.

Es una estrategia que permite a una comunidad exponer sus metadatos manteniendo condiciones de autonomía local. En el caso del OAI-PMH el esquema de metadatos básico recomendado es Dublin Core sin cualificadores, pero es posible utilizar otros. La petición asociada a este concepto es **ListMetadataFormats**, que posibilita la consulta de la lista de esquemas de metadatos utilizados por el repositorio.

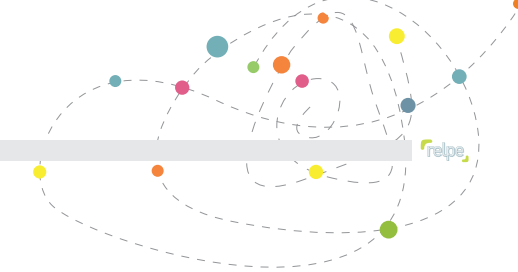
## DINÁMICA DEL PROTOCOLO OAI-PMH

Una vez descriptos los conceptos y componentes básicos previstos en el protocolo, se realizará una reseña acerca de la dinámica de funcionamiento prevista por el protocolo.

### EL COSECHADO DE METADATOS

Como se dijo antes, el protocolo está orientado a una arquitectura cliente-servidor, es decir que un servidor OAI-PMH ofrece servicios que pueden ser consultados por una aplicación informática (el cliente) siempre y cuando cumpla con las reglas del protocolo,





o sea que pueda dirigir las peticiones previstas en el protocolo a través de URL y procesar y entender las respuestas estructuradas en XML generadas por el servicio.

El protocolo como tal no está orientado a usuarios finales, sino que debe mediar una aplicación que realice las peticiones y procese las respuestas.

A partir de la consulta que realiza la aplicación informática (el cliente), literalmente “cosecha” los metadatos que describen los recursos ofrecidos por el servidor. La aplicación que consulta los metadatos puede establecer algunos límites con respecto a la consulta, como, por ejemplo, puede consultar todos los recursos que hubieran sido modificados o creados luego de determinada fecha, o todos los recursos que pertenezcan a determinada colección (set) o agrupamiento de recursos. Estas capacidades del protocolo son en gran medida las responsables de la versatilidad y escalabilidad de este modelo de solución aplicado a la explotación cooperativa de recursos. A través de estas consultas resulta sencillo sincronizar datos acerca de los recursos, pese a que los nodos participantes de una red cooperativa trabajan de manera asincrónica.

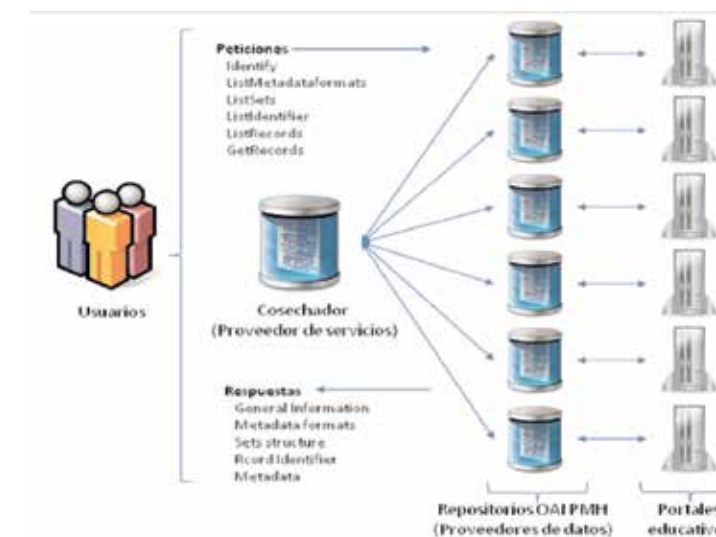
Si una aplicación, que denominaremos Centro educativo 1, “cosecha” todos los metadatos provistos por un servidor que dispone de 10.000 recursos educativos, procesará y copiará los metadatos que referencian 10.000 recursos educativos. Un mes después el servidor creció, y ahora dispone de 10.050 recursos educativos y ha modificado y actualizado 5 recursos ya existentes. Cuando la aplicación Centro educativo 1 vuelva a intentar “cosechar” todos los metadatos de los recursos ofrecidos por el servidor, solo obtendrá los 55 registros correspondientes a la sumatoria de los registros agregados y los modificados. Si al mismo tiempo una segunda aplicación, a la que llamaremos Centro educativo 2, intenta “cosechar” el mismo repositorio, a esta aplicación el servidor le brindará los metadatos correspondientes a los 10.050 registros. De esta manera, pese a los distintos ciclos de gestión y los distintos modos y necesidades, los diferentes actores o nodos que “cosechan” el servidor logran sincronizarse y disponer todos de los mismos datos. Asimismo, las consultas pueden delimitarse también según colecciones (set) de metadatos. Estos filtros se combinan, a su vez, con los filtros según fechas de creación o última modificación antes descriptos.

La capacidad para sincronizar el estado de descripción de los recursos entre distintos nodos a través de un procedimiento sencillo

es uno de los factores que facilitan la escalabilidad del modelo federado. Es un modelo asincrónico de cooperación y articulación de servicios, ya que no depende de la articulación en tiempo real de todos los repositorios o de los nodos cooperantes.

Otro factor relevante a considerar con respecto al modelo de cooperación es el que se refiere a la naturaleza del intercambio. El protocolo contempla la distribución y exposición de metadatos, pero no de los recursos en sí o de los objetos digitales asociados a los metadatos descriptivos. Está orientado a compartir metadatos y no necesariamente los objetos referenciados; esta restricción concentra la fortaleza del modelo en el descubrimiento de recursos más que en la granularidad o riqueza descriptiva.

La descripción extensa, la apropiación y valoración de cada recurso, estaría basada en gran medida en la capacidad para la integración de datos y servicios y la generación de valor agregado que pueda aportar el “cosechador” de metadatos. Asimismo, esta arquitectura les permite a los servicios que exponen los metadatos mejorar la visibilidad web de sus recursos, sin perder tráfico web o sin redundar los recursos educativos o los objetos digitales en miles de servidores, sino solamente los metadatos que los describen.





## LA RED RELPE Y EL PROTOCOLO OAI-PMH

Hasta aquí se han reseñado algunos factores y condiciones consideradas favorables para las dinámicas de explotación cooperativa de recursos en el contexto de la Web. En una segunda instancia se han señalado las características más salientes del protocolo OAI-PMH como tal. A continuación se describen algunas características que podrían favorecer sus condiciones de adopción en el contexto de la red RELPE.

### ESCALABILIDAD

Es un modelo de cooperación sin limitaciones intrínsecas o conocidas con respecto a la cantidad de nodos cooperantes. Cada uno de los miembros RELPE puede exponer sus recursos a través del protocolo OAI-PMH, y es posible crear un repositorio central con todos los metadatos de los recursos educativos de la región y, utilizando el mismo modelo, otro repositorio con todos los metadatos de los recursos educativos disponibles en la región y en otras redes que estuvieran orientados, por ejemplo, a alfabetización informacional. El modelo puede crecer en cantidad de nodos y variar en la especificidad o cualidad de sus miembros.

### SUSTENTABILIDAD

Es un modelo basado en estándares abiertos, utiliza tecnologías basadas en normas ISO o recomendaciones del Consorcio W3C y la URI como sistema de direcciones para la identificación de recursos en el contexto de internet. Estas bases favorecen soluciones neutrales en términos tecnológicos, facilitando la autonomía tecnológica y evitando escenarios de dependencia de un único proveedor.

El marco OAI-PMH se encuentra inscripto en la agenda de investigación y desarrollo compartida por cientos de comunidades e instituciones, lo cual facilita el mantenimiento del modelo y su constante evolución, y suma a la comunidad RELPE al diálogo global acerca de cómo mejorar la difusión y el intercambio de los recursos y las producciones intelectuales.

En lo que a sustentabilidad se refiere, existen soluciones tecnológicas y aplicaciones informáticas para prácticamente todas las plataformas y lenguajes de programación disponibles.

### CONSTRUCCIÓN DE CAPACIDADES

La adopción del protocolo OAI-PMH no es un proceso de innovación basado en la incorporación de nuevas tecnologías o mejores aplicaciones: es una innovación centrada en la adopción de metodologías de trabajo específicas. De esta manera, es una innovación basada en el desarrollo de capacidades locales, en la documentación y difusión de buenas prácticas. En este sentido, constituye una vía de innovación accesible a cualquier actor en cualquier nivel.

### MODELO POLIVALENTE

La adopción del protocolo OAI-PMH permite exponer los metadatos de los recursos educativos no solo hacia RELPE sino también hacia actores locales de cada uno de los portales, como las redes escolares o las instituciones de formación profesional docente, por ejemplo. Asimismo, es posible participar de otras redes de cooperación que actúan a nivel internacional. Otro factor a considerar estaría dado por la mejora en las condiciones de visibilidad y disseminación de las producciones propias: la exposición de datos descriptivos estructurados, como los son los servicios web basados en metadatos, aporta mejoras en las condiciones de visibilidad de las producciones locales en los motores de búsqueda (ej.: Google).

### COOPERACIÓN ENTRE ACTORES ASIMÉTRICOS

Los modelos de cooperación posibles a través del uso del protocolo OAI-PMH permiten diferentes esquemas de articulación y modulación entre los cooperantes. De esta manera, posibilitan escenarios en los que, por ejemplo, un portal nuevo reutiliza los metadatos de un portal ya consolidado, o quizás de la sumatoria de un conjunto de portales o de toda la red RELPE. Todo esto es posible en un marco de absoluta transparencia acerca del origen y destino de los recursos educativos, ya que los mismos siempre serán referenciados hacia su contexto de origen.

### GENERACIÓN DE VALOR AGREGADO

Utilizando el protocolo es posible concebir un escenario en el que RELPE (u otro actor interesado) implementa un repositorio central a partir del cosechado de los portales miembros de la red. Esto permitiría realizar búsquedas a través de todos los portales de RELPE a través de una única interfaz de búsqueda. También permitiría desarrollar modelos de análisis que permitan caracterizar la demografía de los recursos educativos producidos en la región, y eventualmente desarrollar métricas acerca de sus usos.





## RELPE: IMPLEMENTACIÓN DEL MODELO DE EXPLOTACIÓN COOPERATIVA DE RECURSOS EDUCATIVOS UTILIZANDO OAI-PMH

Tal como se indicó al comienzo, sobre la base de las condiciones y actualidad de la red RELPE, las características de la WWW como contexto tecnológico y las posibles potencialidades y ventajas evaluadas en el protocolo OAI-PMH, se estableció una agenda de trabajo orientada a desarrollar una experiencia piloto para la adopción de un modelo de explotación federada de recursos educativos basado en el protocolo de OAI-PMH, que fue presentado en la reunión de República Dominicana.

Dicha agenda estuvo orientada a desarrollar experiencias capaces de dotar a los portales miembros de RELPE de las capacidades técnicas y metodológicas que les permitan exponer los metadatos asociados a sus recursos educativos a través del mencionado protocolo (OAI-PMH).

Asimismo, se implementó un repositorio central en los servidores de RELPE a través del cual es posible consultar los metadatos de los recursos referidos y acceder a los mismos en los portales participantes.

El plan inicial de trabajo, en lo referente a prototipo, contempló la inclusión y exposición de metadatos por parte de tres portales a través de métodos y vías de implementación distintos para cada caso. Esta metodología estaba orientada a establecer la viabilidad del modelo, aun partiendo de distintas arquitecturas, dinámicas de gestión o plataformas tecnológicas.

La implementación del repositorio central de metadatos de recursos educativos RELPE permite brindar una interfaz única, consistente y homogénea para la consulta de los recursos educativos brindados por los portales miembro de la red RELPE.

La experiencia piloto contó con la participación de cuatro países miembros de RELPE: España, Argentina, Brasil y El Salvador.

Como se describió antes, con el fin de contemplar distintos escenarios de partida se establecieron esquemas de solución o procedimientos de implementación para cuatro escenarios a considerar en el marco de la red de portales RELPE. Estos escenarios se describen a continuación.

### ESCENARIO 1: CAPACIDAD INSTALADA

Se trata de portales cuyas herramientas de gestión de contenidos ya disponen de la capacidad para exponer servicios web de acuerdo al protocolo OAI-PMH, o, en su defecto, están en condiciones de desarrollarla e implementarla.

Para estos escenarios, los portales deberán indicar a la red RELPE la dirección web a través de la cual exponen el servicio de metadatos compatible con el protocolo OAI-PMH.



### ESCENARIO 2: ARTICULACIÓN TECNOLÓGICA

Se trata de portales que tengan la posibilidad de intervenir en sus herramientas de gestión de contenidos y establecer una articulación con alguna de las herramientas, módulos o aplicaciones ya disponibles dentro del marco OAI.

Para estos escenarios, los portales deberán implementar los módulos o aplicaciones y luego indicar a la red RELPE la dirección web a través de la cual exponen el servicio de metadatos compatible con el protocolo OAI-PMH.







### ESCENARIO 3: MIGRACIÓN DE METADATOS

Se trata de portales que pueden ofrecer un archivo autónomo codificado en XML según el esquema RELPE, para que sea migrado hacia una herramienta capaz de ofrecer servicios de acuerdo al protocolo OAI-PMH. Para este caso se adoptará una solución tecnológica ya existente para operar como OAI provider, y se desarrollará una utilidad para la importación e ingesta de metadatos desde una fuente XML según el esquema RELPE.

Para estos escenarios, se proveerá de una herramienta disponible para el desarrollo de servicios de acuerdo al protocolo OAI-PMH y se asistirá a los portales en el desarrollo de una dinámica de migración de metadatos. Asimismo, los portales disponen de un tutorial para la implementación de esta herramienta. Luego los portales deberán indicar a la red RELPE la dirección web a través de la cual exponen el servicio de metadatos compatible con el protocolo OAI-PMH.



### ESCENARIO 4: OAI ESTÁTICO

Se trata de portales que optan por exportar los metadatos referidos a sus recursos hacia un archivo autónomo codificado en XML de acuerdo con las pautas del protocolo OAI-PMH para repositorios estáticos. La vía de solución que se propone permite a los portales exponer sus metadatos con una intervención tecnológica mínima.

Para estos casos, los portales disponen de un tutorial que explica los procedimientos para exponer los metadatos de sus recursos educativos según las pautas del protocolo para repositorios estáticos. Luego los portales deberán indicar a la red RELPE la dirección web a través de la cual exponen el servicio de metadatos compatible con el protocolo OAI-PMH.

En todos los casos se requiere que los portales puedan establecer los siguientes datos obligatorios para cada set de metadatos:

- Una URL única por set de metadatos.
- Indicación de la colección o agrupamiento de pertenencia dentro del contexto del portal.
- Una fecha de última modificación por set de metadatos.



### LA PRUEBA PILOTO

Una vez establecidos los escenarios y los portales participantes de la experiencia piloto, cada uno de los portales adoptó alguno de los escenarios y llevó adelante las tareas necesarias para exponer sus recursos según lo acordado. Esto permitió verificar los escenarios propuestos con datos reales.

- Se evaluó el escenario 1 accediendo al servicio OAI provisto por el proyecto Agrega (España).
- Se evaluó el escenario 2 accediendo al módulo OAI implementado en el contexto del actual plan de trabajo por parte del BIOE (Brasil).
- Se evaluó el escenario 3 implementando un módulo externo al CMS utilizado por Educ.ar. Dicho módulo utilizó como fuente de datos un Excel provisto por Educ.ar. El mismo fue migrado hacia el módulo y actualmente brinda servicios OAI.
- Se evaluó el escenario 4 a través de la generación de una





fuelle OAI estática por parte del portal Mi Portal (El Salvador). Se elaboró un tutorial para facilitar la generación de la fuente estática (ver anexo).

Finalmente, se implementó el repositorio central en los servidores de RELPE y se cosecharon los metadatos desde las cuatro fuentes que se corresponden con los escenarios antes descriptos.

Actualmente el repositorio central se encuentra implementado en <http://oai.relpe.org>

A su vez, el cosechado de los portales participantes permitió concentrar cerca de 11.000 recursos, pudiendo este número incrementarse sucesiva e indefinidamente a través de los mismos procedimientos utilizados para el desarrollo del piloto.

## DEL INTERCAMBIO DE CONTENIDOS A LA EXPLOTACIÓN COOPERATIVA

La conclusión exitosa del piloto le brinda a RELPE las herramientas necesarias para realizar un cambio en cuanto a la circulación de los recursos tanto dentro como fuera de la red, pasando de un modelo de intercambio de contenidos cerrado a una arquitectura abierta de explotación cooperativa de contenidos.

Una de las principales ventajas de la solución adoptada es su pluralidad, que resuelve tanto la forma de compartir recursos en RELPE con iniciativas locales, otras comunidades de pares y comunidades de práctica, como también brinda los medios para la descripción de los recursos.

Al ser un esquema tecnológicamente abierto, permite que los actores dispongan libremente de los pisos tecnológicos a utilizar y los modelos de gestión posibles, además de ofrecer una articulación con la agenda de investigación y desarrollo compartida con otros actores e instituciones, discutiendo sobre estos temas y generando mejores prácticas.

A continuación se anexan los tutoriales orientados a asistir a los equipos técnicos en los procesos de implementación del esquema basado en repositorios estáticos y el esquema basado en migración de metadatos.

## GUÍA DE REDACCIÓN DE FUENTES DE DATOS XML PARA SERVICIOS OAI BASADOS EN FUENTES ESTÁTICAS

Este tutorial servirá para obtener la fuente estática XML necesaria para el escenario 4: OAI Estático. Siguiendo estos pasos se obtendrá un archivo que contenga el listado total de recursos que el portal comparte. El archivo se alojará en los servidores del portal y estará disponible para ser cosechado por el proveedor de servicios OAI de RELPE.

### DATOS DESCRIPTIVOS DEL REPOSITORIO

Cada fuente de datos estática deberá disponer de un set de etiquetas que describan la propia fuente de datos. A continuación se describen las etiquetas a incluir:

- oai:repositoryName: Nombre del repositorio.
- oai:adminEmail: correo electrónico del administrador del repositorio.
- oai:earliestDatestamp: fecha más temprana de creación de set de metadatos.

Los demás elementos, para los fines de esta experiencia, serán fijos. A continuación se presenta un ejemplo de set de metadatos descriptivos de un repositorio:





```
<Identify>
  <oai:repositoryName>Educar</oai:repositoryName>
  <oai:baseURL>http://www.educ.ar/URL_Fuente_XML.xml</oai:baseURL>
  <oai:protocolVersion>2.0</oai:protocolVersion>
  <oai:adminEmail>admin@educ.ar</oai:adminEmail>
  <oai:earliestDatestamp>2010-01-01</oai:earliestDatestamp>
  <oai:deletedRecord>no</oai:deletedRecord>
  <oai:granularity>YYYY-MM-DD</oai:granularity>
</Identify>
```

## SET DE METADATOS DE RECURSOS

El set de metadatos de cada recurso se compone de dos tipos de datos, datos meta-descriptivos de los metadatos (header) y los metadatos propiamente dichos (metadata). A continuación se describen los elementos que los componen.

### ETIQUETAS DE DATOS:

Datos del encabezamiento del set de metadatos:

```
xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
http://www.openarchives.org/OAI/2.0/oai_dc.xsd"
```

- oai:identifier: identificador único del set de metadatos del recurso en contexto del servicio OAI.
- oai:datestamp: fecha de última modificación del set de metadatos del recurso.

Datos propios del set de metadatos:

Este apartado se compone del set de metadatos 1.1 de Dublin Core <sup>13</sup>.

Se incluye a su vez la referencia hacia los espacios de nombres (namespace) utilizados:





A continuación se presenta un ejemplo de set de metadatos codificado en XML:

```
<oai:record>
<oai:header>
<oai:identifier>oai:mined.sv/0809201001</oai:identifier>
<oai:datestamp>2010-09-08</oai:datestamp>
</oai:header>
<oai:metadata>
<oai_dc:dc xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
<dc:identifier>
http://www.miportal.edu.sv/sitios/14869/bb/Poema_de_M%EDo_Cid.doc
</dc:identifier>
<dc:title>Poema del Mio Cid</dc:title>
<dc:creator>Mio Cid</dc:creator>
<dc:subject>Lenguaje</dc:subject>
<dc:subject>Poesia</dc:subject>
<dc:subject>Bachillerato General</dc:subject>
<dc:subject>Educación Secundaria</dc:subject>
<dc:description>
Poema que canta las aventuras y hazañas del Cid Campeador, Ruy Díaz de Vivar en su reconquista de España.
Este libro se utiliza en la segunda unidad del Primer Año de Bachillerato.
</dc:description>
<dc:type>actividades para el aula</dc:type>
<dc:type>Presentación</dc:type>
<dc:type>Sugerencia metodológica</dc:type>
<dc:type>Actividad ejercitación</dc:type>
<dc:date>2010-09-08</dc:date>
<dc:source>Bilbioteca Tecpan</dc:source>
<dc:language>es</dc:language>
<dc:publisher>mined.gob.sv</dc:publisher>
<dc:format>text/doc</dc:format>
<dc:rights>mined.gob.sv</dc:rights>
</oai_dc:dc>
```

## PROVEEDOR DE DATOS OAI SATÉLITE

Este tutorial sirve para implementar el proveedor de datos OAI propuesto en el Escenario 2: Articulación tecnológica. y el Escenario 3: Migración de datos. Es útil para aquellos países que no cuentan con una solución integrada dentro de su CMS y poseen la capacidad tecnológica para instalar y gestionar el componente.

## TUTORIAL DE IMPLEMENTACIÓN

A continuación se presenta un instructivo para la implementación de un proveedor de datos de acuerdo con el protocolo OAI-PMH (versión 2.0) en un esquema satelital, es decir, como una aplicación dinámica independiente y autónoma de las herramientas o sistemas utilizados para administrar los contenidos o metadatos de los recursos educativos contenidos en el portal.

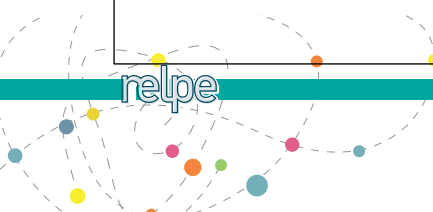
Luego de un relevamiento de las herramientas y aplicaciones disponibles y adecuadas para el mencionado modelo de servicios, se ha optado por utilizar la herramienta desarrollada por Heinrich

Stamerjohanns (Institute for Science Networking, Carl von Ossietzky Universität Oldenburg) denominada PHP OAI Data Provider<sup>14</sup>. La herramienta fue seleccionada ponderando criterios relativos a la sencillez de implementación y la posible disponibilidad de los requerimientos de la herramienta en los portales miembros de la red. Igualmente, se realizaron pequeñas modificaciones al código fuente original con el fin de actualizar ciertas funciones desactualizadas y lograr mayor portabilidad. El código modificado está disponible en el portal de RELPE.

## CARACTERÍSTICAS FUNCIONALES DE LA HERRAMIENTA

La herramienta permite implementar servicios basados en las cláusulas previstas por el protocolo OAI-PMH versión 2.0. Asimismo, permite definir y subdividir los metadatos ofrecidos según sets (colecciones de metadatos).

Una vez implementada, la herramienta expone los servicios a través del archivo oai2.php, ubicado en el directorio base del código.





## OPCIONES DE CONFIGURACIÓN

La herramienta dispone de un archivo de configuración que permite establecer los parámetros necesarios para su operación.

Entre los mencionados parámetros hay algunos cuya modificación resulta indispensable, otros cuya configuración permite adecuar el funcionamiento de la herramienta a las necesidades puntuales de cada implementación, y otros parámetros que permiten extender u orientar las funcionalidades de manera específica y detallada. En el presente instructivo serán cubiertos los dos primeros grupos de parámetros.

## OPCIONES DE CONFIGURACIÓN OBLIGATORIAS

A continuación se enumeran las opciones de configuración obligatorias.

### UBICACIÓN DE LA LIBRERÍA PEAR

PEAR es un gestor de clases para PHP que permite utilizar de manera sencilla y documentada funciones extendidas o patrones funcionales o transaccionales en PHP. Las clases PEAR necesarias para la operatoria de la herramienta ya se encuentran incluidas en la presente distribución de la herramienta, y este parámetro ya se encuentra configurado para su correcta operación. Sin embargo, si su servidor dispone ya de una instalación local de PEAR, puede modificar este parámetro para no redundar código y mantenimiento.

## CONEXIÓN CON LA BASE DE DATOS

Estos parámetros permiten establecer una conexión con la base de datos MySQL. Deben ser modificados de acuerdo con los permisos, usuarios y bases de datos de su entorno de trabajo.

```
$DB_HOST = 'nombre_del_servidor';
$DB_USER = 'nombre_de_usuario';
$DB_PASSWD = 'clave_de_usuario';
$DB_NAME = 'nombre_de_la_base_de_datos';<oai:record>
```

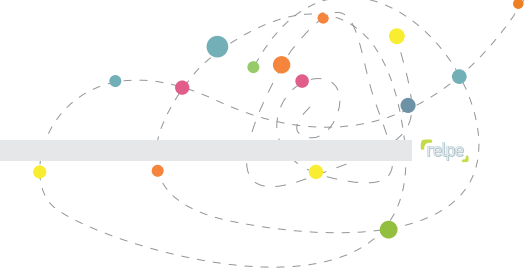
## CODIFICACIÓN UTILIZADA

Este parámetro permite establecer la codificación de caracteres utilizada en los datos almacenados. La presente distribución está configurada para almacenar datos codificados en iso8859-1. Eventualmente puede modificar este parámetro y establecer la codificación en UTF-8 (utf-8).

## OPCIONES DE CONFIGURACIÓN ESTÁNDAR

A continuación se describen las opciones de configuración que permiten adecuar las capacidades de la herramienta a ciertas particularidades funcionales, o según determinaciones de las estructuras de datos de las fuentes de origen.





## OPCIONES DE CONFIGURACIÓN ESTÁNDAR

### Nombre del repositorio

Nombre descriptivo del repositorio. Debe ser único.

```
$repositoryName = 'dominio_portal.edu.pais';
```

### Fecha inicial

Fecha inicial desde la cual se disponen y exponen recursos a través del protocolo. Debe ser una única fecha.

```
$earliestDatestamp = '2001-05-23';
```

### Correo de contacto

Correo de contacto del administrador o administradores del servicio. Pueden ser más de uno.

```
$adminEmail=array('mailto:admin@dominio_portal.edu.pais',  
'mailto:otromail@dominio_portal.edu.pais');
```

### Identificador del repositorio

Nombre e identificador global para el repositorio. Debe ser único y compuesto por el nombre del dominio del proveedor OAI.

```
$repositoryIdentifier = 'dominio_portal.edu.pais';
```

## Publicar detalle del identificador del repositorio

Esta opción permite definir si se publicarán el detalle y ejemplo de un identificador de recursos como respuesta a la cláusula identify prevista en el protocolo OAI-PMH.

```
$show_identifier = true;
```

Si se consigna el valor true incluirá los siguientes datos de identificación:

```
<oai-identifier xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai-  
identifier http://www.openarchives.org/OAI/2.0/oai-identifier.xsd">
```

```
<scheme>oai</scheme>
```

```
<repositoryIdentifier>educ.ar</repositoryIdentifier>
```

```
<delimiter>:</delimiter>
```

```
<sampleIdentifier>oai:educ.ar:anIdentifier</sampleIdentifier>
```

```
</oai-identifier>
```

Número máximo de registros a servir

```
$MAXRECORDS = 50;
```

## Número máximo de registros a servir

Indicación del número máximo de registros a servir en la petición ListRecords (Lista de registros). Luego se entregará un indicador para retomar el suministro de registros (ResumptionToken)

```
$MAXRECORDS = 50;
```





### Número máximo de identificadores a servir

Indicación del número máximo de registros a servir en la petición ListIdentifiers (Lista de identificadores). Luego se entregará un indicador para retomar el suministro de identificadores (ResumptionToken).

```
$MAXIDS = 200;
```

### Colecciones (sets)

Declaración de las colecciones o sets en las que se encuentran distribuidos los recursos.

Pueden declararse múltiples colecciones. Las mismas deben coincidir con las declaradas en el campo oai\_set incluido en la tabla. Puede utilizar un identificador de colección (setSpec), un nombre descriptivo (setName) y un texto descriptivo (setDescription). Solo es obligatorio utilizar al menos un identificador de colección.

```
$SETS = array (
    array('setSpec'=>'inicial','setName'=>'nivel inicial', 'setDescription'=>'recursos
educativos de nivel inicial'),
    array('setSpec'=>'basica','setName'=>'egb', 'setDescription'=>'recursos educativos de
educación general básica'),
    array('setSpec'=>'polimodal','setName'=>'nivel polimodal',
'setDescription'=>'recursos educativos de nivel polimodal')
```

### Delimitador de ocurrencias de campo o etiqueta

Este parámetro permite establecer el carácter de separación utilizado entre varias ocurrencias de un mismo campo, por ejemplo un recurso con dos autores. Por defecto, se utiliza un punto y coma (;) como carácter de separación.

```
$$SQL['split'] = ';';
```

### Tabla única de registros

Este parámetro establece el nombre de la tabla única en la que se encuentran los sets de metadatos. Por defecto y de acuerdo con el script de instalación incluido en el presente código, se utiliza la tabla oai\_records. Eventualmente, puede utilizar otras o modificar su nombre.

```
$$SQL['table'] = 'oai_records';
```

### Columna de clave única

Este parámetro establece el nombre de la columna que define la clave única y secuencial de la tabla en la que se encuentran los sets de metadatos. Por defecto, y de acuerdo con el script de instalación incluido en el presente código, el nombre de la columna es serial. Eventualmente, puede utilizar otra o modificar su nombre.

```
// the name of the column where you store your sequence
// (or autoincrement values).
$$SQL['id_column'] = 'serial';
```





### Columna de identificador único

Este parámetro permite establecer el nombre de la columna del identificador único del set de metadatos en el contexto de su portal. Por defecto, y de acuerdo con el script de instalación incluido en el presente código, el nombre de la columna es url. Eventualmente, puede utilizar otro o modificar su nombre.

```
// the name of the column where you store the unique identifiers
// pointing to your item.
// this is your internal identifier for the item
$SQL['identifier'] = 'url';
```

## REQUERIMIENTOS DE BASE PARA LA INSTALACIÓN

La herramienta requiere para su implementación de un ambiente que cuente con las siguientes aplicaciones instaladas:

- Servidor HTTP (por ejemplo, Apache) con soporte para lenguaje de programación interpretado PHP (versión 4 o superior).
- Servidor de base de datos: La herramienta dispone de una clase de abstracción que permite utilizar prácticamente cualquier servidor de base de datos (MySQL, Oracle, Postgres, MS server, etc.). El presente instructivo se encuentra basado en una implementación que utiliza MySQL como servidor de bases de datos.

1. Descargue y descomprima la versión del código disponible en: [http://oai.relpe.org/phpoai2\\_relpe\\_ver20101004.zip](http://oai.relpe.org/phpoai2_relpe_ver20101004.zip)
2. Copie los archivos en un directorio ubicado dentro del directorio de documentos de su servidor HTTP, por ejemplo /var/www/portal/oai/.
3. Establezca permisos de escritura para el servidor HTTP sobre el directorio tokens (incluido en el código), por ejemplo:

```
chown wwwrun:nogroup tokens;
```

```
chmod ug+rx tokens
```

Este permiso es necesario para que la herramienta pueda gestionar los datos de suministro y reanudación de suministros de datos (resumptionToken).

4. Edite según sea necesario el archivo de configuración oai2/oaidp-config.php incluido en el presente código. Según haya elegido utilizar la







estructura de datos provista con el código o utilizar una propia, puede crear la tabla única de datos utilizando el script de creación de tabla provisto en el código (en el directorio doc/) o modificar el archivo de configuración de acuerdo con sus propias estructuras de datos.

5. Puede revisar la correcta instalación de la herramienta invocando la URL [http://www.direccion\\_de\\_su\\_servidor/oai/oai2.php?verb=Identify](http://www.direccion_de_su_servidor/oai/oai2.php?verb=Identify)

Para incorporar los metadatos de su portal al servicio puede usar utilidades de migración desde sus estructuras de datos (bases de datos, Excel, etc.) hacia la tabla única provista, respetando la estructura de columnas descrita a continuación o adecuar el archivo de configuración hacia la estructura de columnas provista por su estructura de datos local.

### TABLA DE DATOS

La herramienta almacena todos los metadatos afectados al servicio en una única tabla. Este parámetro permite establecer el nombre de esa tabla. En el código de la presente distribución el nombre es oai\_records. El script SQL para la creación de la tabla se encuentra en el directorio doc de la presente distribución.

La tabla resultante dispone de los siguientes campos:

- serial: clave única y autoincremental de la tabla;
- provider: nombre del proveedor OAI (por ejemplo: educ.ar);
- URL: URL del recurso descrito;
- enterdate datetime: registro de fecha (con horas, minutos y segundos) de la fecha de registro del set de metadatos;
- oai\_identifier: identificador OAI del set de metadatos. debe ser único en el contexto del servicio;

• oai\_set: mención del set o colección al que pertenece el recurso descrito por el set (conjunto de campos) de metadatos;

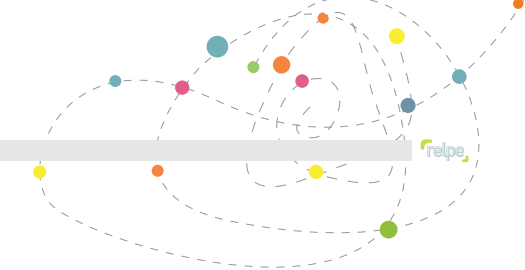
• datestamp datetime: estampilla de tiempo que consigna la fecha de última modificación del set de metadatos;

• deleted: indicador del estado del set de metadatos (eliminados sí o no).

Luego la tabla contiene los campos correspondientes al set de metadatos Dublin Core sin cualificar <sup>16</sup>:

- dc\_title
- dc\_creator
- dc\_subject
- dc\_description
- dc\_contributor
- dc\_publisher
- dc\_date
- dc\_type
- dc\_format
- dc\_identifier
- dc\_source
- dc\_language
- dc\_relation
- dc\_coverage
- dc\_rights





## ALGUNOS RECURSOS Y DIRECCIONES ÚTILES

### DOCUMENTOS Y TUTORIALES

---

- **OAI-PMH Static Repository Gateway**

Tutorial sobre cómo hacer un repositorio OAI estático.

<http://srepod.sourceforge.net>

---

- **Open Archives Initiative Protocol for Metadata Harvesting Protocol Version 2.0 of 2002-06-14**

Este documento describe el protocolo OAI PMH, características, componentes y modo de implementación.

<http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm>

---

- **Buenas prácticas para proveedores OAI (Open Archive Initiative)**

Sitio de la Digital Library Federation <sup>27</sup> y la National Science Digital Library <sup>18</sup> que recopila recomendaciones y ejemplos de buenas prácticas para implementaciones de proveedores de datos OAI y servicios orientados a compartir metadatos.

[http://webservices.itcs.umich.edu/mediawiki/oaibp/index.php/Main\\_Page](http://webservices.itcs.umich.edu/mediawiki/oaibp/index.php/Main_Page)

---

- **OAI para principiantes - OA-Forum tutorial**

Tutorial claro y completo (incluye glosario), basado en los programas impartidos junto a los talleres segundo (2002) y tercero (2003) del Open Archives Forum. <sup>29</sup> Elaborado por Leona Carpenter y traducido por Domingo Arroyo Fernández.

<http://travesia.mcu.es/portalnb/jspui/html/10421/1823/intro.htm>

---





• **Open archives initiative. Protocol for Metadata Harvesting (OAI-PMH): descripción, funciones y aplicación de un protocolo**

Artículo de José Manuel Barrueco.

<http://www.uv.es/~barrueco/cardedeu.doc>

• **DRIVER: infraestructura para repositorios digitales de la Unión Europea**

Documentos, lineamientos, recomendaciones técnicas y herramientas para desarrollar repositorios digitales en el contexto de la iniciativa DRIVER <sup>20</sup>

<http://www.driver-repository.eu/Downloads>

• **DRIVER 2.0 Directrices para proveedores de contenido** – Exposición de recursos textuales con el protocolo OAI-PMH [Noviembre 2008]

[http://www.driver-support.eu/documents/DRIVER\\_2\\_0\\_Guidelines\\_Spanish.pdf](http://www.driver-support.eu/documents/DRIVER_2_0_Guidelines_Spanish.pdf)

**HERRAMIENTAS Y FUENTES OAI-PMH**

• **OAI Repository Explorer**

Sitio que dispone de una interfaz que permite testear proveedores de servicio OAI.

<http://re.cs.uct.ac.za>

• **Herramientas OAI-PMH**

Lista de herramientas OAI mantenida por la Digital Library Federation

<http://webservices.itcs.umich.edu/mediawiki/oaibp/index.php/TurnKeyPackages>

• **UIUC OAI Metadata Harvesting Project**

Herramienta desarrollada en Visual Basic para OAI

<http://sourceforge.net/projects/uilib-oai>

• **Liblivecd**

Ubuntu LiveCD que incluye DSpace, PKP OAI Harvester, Koha Library Management Software y dbwiz

<http://sourceforge.net/projects/liblivecd>

• **XqOAI**

Herramienta para data provider de OAI desarrollada con Xquery

<http://code.google.com/p/xqoai>

• **Open Archives Initiative: aplicaciones**

Directorio de aplicaciones y herramientas para brindar servicios basado en el protocolo de cosechado de metadato OAI-PMH. El directorio es mantenido por Open Archives.

<http://www.openarchives.org/pmh/tools/tools.php>





• **OJAX: Ajax-powered metasearch service**

Práctica herramienta Web desarrollada en JAVA que permite implementar de manera sencilla una interfaz de búsquedas federadas en base al protocolo OAI-PMH sobre cualquier (uno o varios) repositorios OAI.

<http://ojax.sourceforge.net>

• **OAI Cat**

Proyecto Open Source basado en Java Servlet web application para proveer servicios según OAI-PMH v2.0

<http://www.oclc.org/research/software/oai/cat.shtm>

• **PHP OAI**

Aplicación web escrita en PHP que permite implementar servicios de provisión de datos OAI-PMH.

<http://physnet.uni-oldenburg.de/oai>

• **Open Archives Harvester**

Aplicación web desarrollada por el Public Knowledge Project <sup>21</sup> que permite cosechar masivamente registros de repositorios OAI a través del protocolo OAI-PMH. Soporta esquemas Dublin Core MODS <sup>22</sup>, MARCXML <sup>23</sup> y cuenta con un módulo que permite extenderlo a otros esquemas. Cuenta con una interfaz de búsqueda y exploración y un módulo de administración. Esta desarrollada en PHP y MySQL bajo licencia GPL.

<http://pkp.sfu.ca/?q=harvester>

• **Islandora**

Herramienta para gestionar repositorio a partir de la integración de Drupal y Fedora. Desarrollada por la Robertson Library de la University of Prince Edward Island <sup>24</sup>.

<http://islandora.ca/about>

• **mod\_oai: OAI como un módulo del apache**

mod\_oai es un módulo desarrollado por Old Dominion University <sup>25</sup> que permite dar servicios OAI con solo disponer de un servidor http Apache

<http://www.modoi.org>

• **OAIbiblio**

herramienta PHP para data provider OAI

<http://www.ibiblio.org/oaibiblio>

• **Zentity: los repositorios según Microsoft**

Zentity es una herramienta gratuita de Microsoft para gestionar repositorios. Dispone de servicios OAI

<http://research.microsoft.com/en-us/projects/zentity>

• **Proveedores de datos OAI registrados**

Lista de los proveedores de datos OAI registrados en el sitio de la iniciativa Open Archives

<http://www.openarchives.org/Register/BrowseSites>

• **Repository maps**

Servicio que combina el directorio de repositorios OpenDOAR <sup>26</sup> y Google maps para geolocalizar fuentes de repositorios OAI-PMH

<http://maps.repository66.org>

• **Scientific Commons.org**

Proyecto sostenido por el Institute for Media and Communication Management de la University of St.Gallen <sup>27</sup> dedicado a reunir en una misma interfaz de búsqueda múltiples repositorios OAI-PMH.

<http://en.scientificcommons.org>





**1** Relpe <http://www.relpe.org>

**2** Documento Técnico 1 - Normas para la catalogación de contenidos educativos. <http://www.relpe.org/documentos-tecnicos/documento-tecnico-1>

**3** Documento Técnico 2 - Intercambio de contenidos en el contexto RELPE. Descripción de la solución tecnológica. <http://www.relpe.org/documentos-tecnicos/documento-tecnico-2>

**4** CMS Content Manager System

**5** Herramienta para la Exportación, Traducción e Importación de Contenidos

**6** eXtensible Markup Language <http://www.w3.org/XML>

**7** Dublin Core Metadata Element Set, Version 1.1 <http://dublincore.org/documents/dces>

**8** Open Archives Initiative Protocol for Metadata Harvesting <http://www.openarchives.org>

**9** LOM Learning Object Metadata [http://ltsc.ieee.org/wg12/files/LOM\\_1484\\_12\\_1\\_v1\\_Final\\_Draft.pdf](http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf)

**10** Uniform Resource Locators <http://tools.ietf.org/html/rfc1738>

**11** Uniform Resource Identifier <http://tools.ietf.org/html/rfc3986>

**12** HyperText Transfer Protocol <http://www.w3.org/Protocols>

**13** Según se describe en: <http://dublincore.org/documents/dces> y en los documentos técnicos de RELPE

**14** PHP OAI Data Provider <http://physnet.uni-oldenburg.de/oai>

**15** Para más detalles: <http://www.openarchives.org/OAI/2.0/guidelines-oai-identifier.htm>

**16** Para más detalles consultar la documentación relativa a la iniciativa Dublin Core [http://www.sedic.es/usando\\_dublin\\_core.pdf](http://www.sedic.es/usando_dublin_core.pdf) o el esquema utilizado <http://dublincore.org/schemas/xmls/simpledc20021212.xsd>

**17** Digital Library Federation <http://www.diglib.org>

**18** National Science Digital Library <http://nsdl.org>

**19** Open Archives Forum <http://www.oaforum.org>

**20** Digital Repository Infrastructure Vision for European Research <http://www.driver-community.eu>

**21** Public Knowledge Project <http://pkp.sfu.ca>

**22** Metadata Object Description Schema <http://www.loc.gov/standards/mods>

**23** MARC 21 XML Schema <http://www.loc.gov/standards/marcxml>

**24** Robertson Library <http://library.upei.ca>

**25** Old Dominion University <http://www.odu.edu>

**26** OpenDoar <http://www.opendoar.org>

**27** Institute for Media and Communication Management <http://www.mcm.unisg.ch>







relpe

red latinoamericana  
portales educativos

[www.relpe.org](http://www.relpe.org)