

# Uso del Web semántico para la interoperabilidad semántica de recursos educativos en Internet y redes P2P

Sandra Aguirre, Joaquín Salvachúa, Juan Quemada, Alberto Mozo  
Departamento de Ingeniería Telemática. Universidad Politécnica de Madrid  
ETSI de Telecomunicación. Ciudad Universitaria s/n. 28040 – Madrid.  
Telf: +34 91 3367366 Ext.: 3032, 3011  
E-mail: {saguirre, jsalvachua, quemada}@dit.upm.es, amozo@eui.upm.es

## Resumen

*La existencia de una gran cantidad de información en la Web, ha implicado el reto a nivel de investigación por encontrar soluciones que nos permitan realizar búsquedas de una manera más eficiente, a tiempo y con resultados coherentes. El artículo presenta la experiencia del proyecto Elena, en donde se ha implementado una novedosa infraestructura y solución informática usando algunas tecnologías del Web semántico para alcanzar la interoperabilidad semántica en el contexto de los sistemas de información educativos. El uso del Web semántico en el proyecto Elena ha permitido optimizar los procesos de búsqueda, identificar relaciones entre recursos y permitir la personalización de preferencias y rasgos particulares. Como caso práctico se presentará la integración del nodo educativo EducaNext dentro de la red de Elena, empleando el modelo semántico común para la descripción de recursos educativos, el API para consultas y lenguaje QEL.*

**Área Telemática:** Internet.

Tecnologías Web: Web semántica, XML y Web Services.

## 1. Introducción

El crecimiento acelerado de Internet y su uso como medio de apoyo educativo han permitido la creación de un gran número de recursos educativos en formato digital. Sin embargo, gran parte de estos recursos no pueden ser accedidos de una manera eficiente dado que se encuentran almacenados dentro de sistemas de eLearning y sistemas de gestión de contenidos educativos cerrados y propietarios.

La mayoría de estos sistemas emplean sus propios modelos de datos para la descripción de sus recursos educativos, lo que trae consigo la heterogeneidad de la información. Por tanto, las tareas de búsqueda, recuperación e integración de recursos educativos no son fáciles de implementar debido al carácter heterogéneo en la descripción de los recursos.

El primer paso para la definición de una sintaxis y semántica común para la descripción de los recursos educativos ha sido el surgimiento de estándares basados en XML (Extensible Markup Language) que proporcionan una serie predefinida de atributos para la descripción, identificación y búsqueda de recursos educativos a través de metadatos.

Los metadatos para recursos educativos se han convertido en un tema de amplia investigación en el

campo de la interoperabilidad, varios organismos como AICC (The Aviation Industry Computer-Based Training Committee) [2], IEEE [5], IMS (Instructional Management Systems) [10], ADL (Advanced Distributed Learning Initiative) [1], Dublin Core [4], entre otros, trabajan en este campo. Sin embargo, existe mucho contenido útil, disponible bajo diversos entornos, que no está estructurado según los estándares, lo cual provoca serias dificultades para que puedan ser reutilizados.

Las diferencias entre los modelos de datos proporcionados por los sistemas de eLearning y los estándares educativos, plantea la necesidad de buscar alternativas que involucren no sólo el aspecto de sintaxis sino también de semántica (significado) para conciliar estos modelos con el fin de permitir la interoperabilidad entre distintos modelos.

Una solución para este problema es emplear el enfoque del Web Semántico para conseguir la interoperabilidad de recursos educativos entre diferentes sistemas. En concreto usando el estándar RDF (Resource Description Framework) [15].

RDF proporciona un soporte para metadatos flexible y abierto que permitirá la interconexión con otros sistemas similares. RDF junto con XML son dos tecnologías básicas que el Web Semántico utiliza y que ya están funcionando. XML permite estructurar el documento, mientras que RDF le da el significado. RDF es una Recomendación del W3C, basado en XML, que nos proporciona tecnología para escribir metadatos que proporciona interoperatividad entre aplicaciones que intercambian información comprensible por la

máquina para definir esquemas para vocabularios específicos como Dublin Core. Esto permite especificar clases de tipos de recursos y propiedades para dar a conocer descripciones, relaciones entre esas propiedades y clases, y las restricciones en las combinaciones permitidas de clases, propiedades y valores.

Otra tecnología del Web semántico para la construcción de modelos descriptivos de recursos educativos va dirigida hacia la creación de un modelo basado en ontologías, como estructuras mas completas para la descripción de recursos educativos, que permiten una representación formal de un concepto además de la representación semántica y sintáctica del mismo. Se considera que la concepción del Web semántico permitirá un nuevo nivel de los servicios en los sistemas de eLearning. Las ontologías toman un papel clave en la resolución de interoperabilidad semántica entre sistemas de información y su uso.

Interesantes iniciativas en el área de la interoperabilidad semántica de repositorios educativos se han desarrollado. Proyectos como Elena [3], EduSource [6] y OKI (Open Knowledge Initiative) [13] plantean interesantes arquitecturas de integración de repositorios educativos heterogéneos que se encuentran localizados en forma distribuida.

Se dará un enfoque al proyecto Elena, una infraestructura de mediación e interoperabilidad entre repositorios educativos heterogéneos, soportada por las tecnologías del Web semántico, mediadores educativos y redes Peer to Peer, cuyo fin principal es crear un único espacio de integración de servicios y recursos educativos.

## 2. Interoperabilidad semántica en Elena

Para alcanzar la interoperabilidad semántica en el contexto de los sistemas de información educativos, el proyecto Elena [3] ha implementado una novedosa infraestructura y solución informática usando varias tecnologías del Web semántico. El objetivo principal de la red de Elena es la creación de un espacio inteligente de aprendizaje que permite integrar nodos educativos en una red semántica y proporcionar servicio de búsqueda de recursos educativos de una manera más eficiente a través de técnicas de personalización.

Los componentes principales de esta infraestructura son:

1. Un API para establecimiento de sesión y realización de consultas sincrónicas y asincrónicas llamado SQI [16] (Simple Query Interface) que define los servicios que un repositorio puede tener disponibles para recibir y responder consultas de otros repositorios (Figura 1).

SQI es parte de la arquitectura para la interoperabilidad de repositorios educativos LORI, la cual define los servicios necesarios para permitir la interoperabilidad entre repositorios educativos. Estos servicios incluyen servicios básicos como por ejemplo servicios de autenticación, gestión de la sesión y servicios de aplicación como gestión de las consultas.

SQI es neutral en términos de lenguaje de consultas y modelo semántico usado.

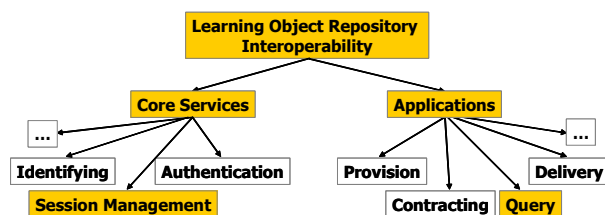


Figura 1. API SQI: Simple Query Interface

Además de ELENA, la red de Excelencia PROLEARN esta coordinando algunas de las experimentos prácticos con estas especificaciones.

2. Un modelo semántico común para la descripción de recursos educativos, con el fin de facilitar las búsquedas.
3. Uso de un lenguaje común basado en Datalog para el formato de las consultas que se enviarán a cada uno de los nodos educativos, llamado QEL (Query Exchange Language) [12]. El formato de los resultados generados esta basado en sentencias RDF.
4. Nodos Educativos que contienen recursos educativos. Algunos de los nodos conectados en la red de Elena son: Amazon, Clix, EducaNext, Lason, Seminarshop.com, Iteach you, ULI.

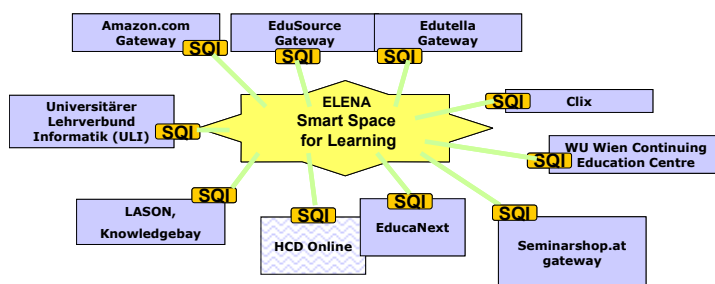


Figura 2. Nodos que conforman la red de Elena.

Como caso práctico se presentará la integración del nodo educativo EducaNext [10] dentro de la red de Elena, empleando el modelo semántico común para la descripción de recursos educativos, el API para consultas y lenguaje QEL.

## 2. Integrando Educanext dentro de la red de Elena

EducaNext, es un portal multilingüe de colaboración académica, donde instituciones de educación superior, de investigación y profesionales pueden colaborar e intercambiar recursos educativos y de conocimiento. Al registrar un recurso educativo en EducaNext, el usuario debe indicar el tipo de oferta (pública o con restricciones de acceso) y el tipo de licencia (libre distribución o licenciado). Cuando un usuario elige ofrecer un recurso educativo públicamente, a su vez elige que el recurso pueda ser consultado y reutilizado a través de otras plataformas que hagan parte de la red de Elena.

La arquitectura usada en EducaNext-UPM esta basada en un wrapper que permite la comunicación entre SQI y la base de datos de EducaNext. A través de la interfaz SQI, el wrapper recibe las consultas en formato QEL enviadas a través de la red de Elena, las transforma a lenguaje SQL y reenvia la consultas hacia la base de datos relacional. Los resultados son retornados en formato RDF (Figura 3).

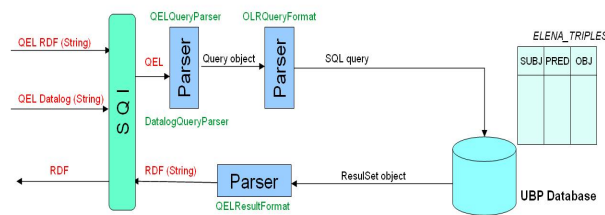


Figura 3: Arquitectura de integración EducaNext-UPM

Edutella [7] ha creado una librería (edutella-sqi.jar), la cual implementa todos los métodos de SQI y usa uno de los wrappers de Edutella para el envío de consultas a repositorios. EducaNext-UPM usa esta librería para conectar el sistema hacia la red de Elena (Figura 4).

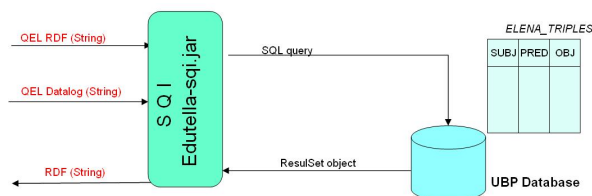


Figura 4: Arquitectura de integración EducaNext-UPM (2)

El proceso de interoperabilidad semántica se permite almacenando la descripción del recurso educativo en forma de metadatos almacenados como tripletas en una base de datos relacional. El

proceso de almacenamiento se realiza de la siguiente manera:

Primero se construye un archivo RDF que contiene la descripción del recurso educativo en forma de metadatos, basándose en un esquema o plantilla en XML/RDF que define los atributos obligatorios y opcionales necesarios para describir los recursos. El esquema está basado en Dublin Core, LOM, OpenQ, and VCard.

```
namespaces
...
<elena_ont:LearningMaterial rdf:about="lr-upm-saguirre-1082994806547"
dc:identifier="lr-upm-saguirre-1082994806547" dc:title="Programación en Java"
dc:language="en" dc:description="Referencia práctica para la programación en Java">

  <dc:subject
rdf:resource="http://www.educanext.org/rdf/taxonomies/dutchBasicClassification.rdf#42.40"/
  >
  <dc:terms:requires>Java</dc:terms:requires>
  <dc:created>
  <dc:terms:W3CDTF>
  <rdf:value>2004-04-26</rdf:value>
  </dc:terms:W3CDTF>
  </dc:created>
  <dc:creator>
  <lom:Entity>
  <vcard:FN>Sandra Aguirre</vcard:FN>
  </lom:Entity>
  </dc:creator>
  <lom-life:editor>
  <lom-life:Contribution>
  <rdf:value>
  <rdf:Seq>
  <rdf:li>
  <lom:Entity>
  <vcard:FN>Juan Quemada</vcard:FN>
  </lom:Entity>
  </rdf:li>
  </rdf:Seq>
  </rdf:value>
  </lom-life:Contribution>
  </lom-life:editor>
  <lom-edu:learning_resource_type>Texto</lom-edu:learning_resource_type>
  <dc:issued>
  <dc:terms:W3CDTF>
  <rdf:value>2004-04-26</rdf:value>
  </dc:terms:W3CDTF>
  </dc:issued>
  <dc:format>
  <dc:terms:IMT>
  <rdf:value>text</rdf:value>
  </dc:terms:IMT>
  </dc:format>
  <lom-tech:version>0.1</lom-tech:version>
```

Figura 5: Ejemplo de un fichero RDF generado en EducaNext-UPM

Luego el archivo RDF es parseado con Jena (A Semantic Web Framework for Java) [11] para generar las tripletas almacenadas en la base de datos relacional.

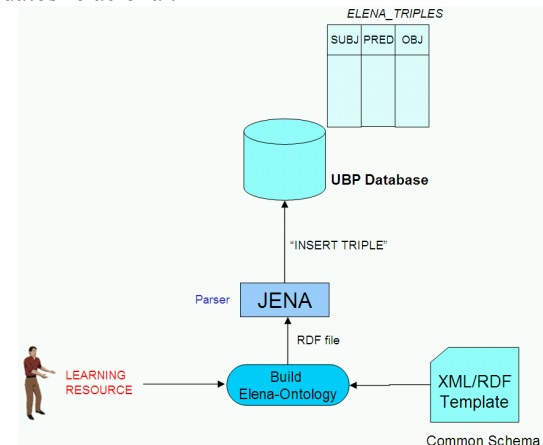


Figura 5. Almacenamiento de un recurso educativo aplicando el esquema común propuesto en Elena.

### 3. Modelo semántico para la descripción de recursos educativos

Se entiende como recurso educativo a los servicios y materiales educativos que son empleados para facilitar y dar soporte en la realización de un objetivo educativo específico. Un servicio educativo hace referencia a un evento educativo cuyo objetivo principal es educar o capacitar personas con la ayuda de un agente humano (profesor, ponente, etc.) en un tiempo predefinido. Cursos, seminarios, tutoriales, etc., son ejemplos de servicios educativos.

El término material educativo es usado para denotar todas las unidades de contenidos que puedan ser almacenados o transferidos en formato digital o no digital. Son consideradas estáticas en el sentido que pueden ser producidas para ser (re)usadas para por los educadores como soporte a sus actividades educativas. Ejemplos de materiales educativos son libros, material de lectura, cursos o clases grabados sobre cintas de video, CD-ROM multimedia.

El foco central de investigación y desarrollo del proyecto Elena se encuentra en la semántica que se va a emplear para describir los recursos educativos con el fin de proveer un marco flexible para la integración de los recursos [8]. Actualmente se emplea un esquema en RDF [9] basado en LOM, Dublin Core, Vcard y OpenQ.

Un ejemplo de la descripción de un documento sobre Java (material educativo) en formato RDF usando el esquema común propuesto por Elena es el siguiente:

Este esquema contiene los atributos necesarios para describir un recurso educativo. Contiene atributos comunes tanto para los materiales como los servicios educativos, por ejemplo: Identificador, Título, Lenguaje, Prerrequisitos, Tema, entre otros. Asimismo contiene atributos específicos tanto para materiales como para recursos educativos. Ejemplos de atributos para materiales educativos tenemos: Fecha de publicación, formato, versión. Ejemplos de atributos para servicios educativos tenemos: Número máximo y mínimo de participantes, horario inicial y final del servicio.

### Agradecimientos

Este trabajo ha sido soportado por el proyecto Elena y es parcialmente patrocinado por la Comisión Europea (IST-2001-37264).

### Referencias

- [1] ADL, Advanced Distributed Learning Initiative, US Department of Defense. Disponible en: <http://www.adlnet.org/>
- [2] AICC, Aviation Industry Computer-Based Training Committee. Disponible en: <http://www.aicc.org/>
- [3] B. Simon, Z. Miklós, M. Sintek, & J. Salvachua. (2003). Smart space for learning: A mediation infrastructure for learning services. In Proceedings of the 12th World Wide Web Conference. Budapest.
- [4] Dublin core metadata initiative. Retrieved October 4, 2003. Disponible en: <http://dublincore.org/>
- [5] E. Law, K. Maillat, J. Quemada, B. Simon: EducaNext: A Service for Knowledge Sharing, in: Proceedings of Ariadne Conference, Leuven, Belgium, November, 2003.
- [6] EduSource Canada: Canadian Network of Learning Object Repositories. Disponible en: <http://www.edusource.ca/>
- [7] Edutella. Disponible en: <http://edutella.jxta.org/>
- [8] G. Rivera, B. Simon, J. Quemada, J. Salvachua: Improving LOM-Based Interoperability of Learning Repositories, in: Proceedings of the OTM 2004 Workshop on Ontologies, Semantics and E-learning, October, 2004.
- [9] IEEE Learning Technology Standards. Disponible en: Committee. <http://ltsc.ieee.org/>
- [10] IMS Global Learning Consortium.. Disponible en: <http://www.imsproject.org/>
- [11] Jena. A Semantic Web Framework for Java. Disponible en: <http://jena.sourceforge.net/>
- [12] Nilsson, M., & Siberski, W. (2003). RDF Query Exchange Language (QEL) - Concepts, Semantics and RDF Syntax. Disponible en: <http://edutella.jxta.org/spec/qel.html>.
- [13] OKI (Open Knowledge Initiative). Disponible en: <http://www.okiproject.org/>.
- [14] PROLEARN project. Disponible en: <http://www.prolearn-project.org/>.
- [15] RDF, Resource Description Framework. W3C recommendation. Disponible en: <http://www.w3.org/RDF/>
- [16] Simple Query Interface. Disponible en: [www.prolearn-project.org/lori](http://www.prolearn-project.org/lori)