OPEN PRAXIS

INTERNATIONAL
COUNCIL FOR OPEN AND
DISTANCE EDUCATION

An Architecture based on Linked Data technologies for the Integration and reuse of OER in MOOCs Context

Nelson Piedra, Janneth Alexandra Chicaiza & Jorge López
Universidad Técnica Particular de Loja (Ecuador)
nopiedra@utpl.edu.ec, jachicaiza@utpl.edu.ec & jalopez2@utpl.edu.ec

Edmundo Tovar
Universidad Politécnica de Madrid (Spain)
etovar@fi.upm.es

Abstract

The Linked Data initiative is considered as one of the most effective alternatives for creating global shared information spaces, it has become an interesting approach for discovering and enriching open educational resources data, as well as achieving semantic interoperability and re-use between multiple OER repositories. The notion of Linked Data refers to a set of best practices for publishing, sharing and interconnecting data in RDF format. Educational repositories managers are, in fact, realizing the potential of using Linked Data for describing, discovering, linking and publishing educational data on the Web. This work shows a data architecture based on semantic web technologies that support the inclusion of open educational materials in massive online courses. The authors focus on a type of openness: open of contents as regards alteration i.e. freedom to reuse the material, to combine it with other materials, to adapt, and to share it further under an open license.

Keywords: OER; OCW; MOOC; Linked Data; Integration; Reuse; Open

Introduction

The amount of open educational resources available nowadays is permanently growing, making difficult their use, re-use, organization and their efficient access. There is a current global movement towards open digital reusable educational materials. Open Educational Resources (OER) are currently seen as a practical way for realizing education for all. In particular, developing countries can benefit through OER from developed regions. The term OER is used to mean a small self-contained unit of self-assessable teaching with a measurable learning objective, often in digital format and free to use. To attach an open license to OER is an efficient way to avoid reuse problems.

The arrival of Massive Open Online Courses (MOOCs) and the growth of open and online education -OER, Open Course Ware (OCW)- is increasingly focus on self-learners as the primary target group. The OER movement has tended to define "openness" in terms of access to use and reuse educational materials, and to address the geographical and financial barriers, among students, teachers and self-learners with distinguished educational institutions (Petrides, Nguyen, Jimes & Karaglani, 2008).

The emergent MOOCs (Piedra, Tovar, Dimovska & Chicaiza, 2013) advance is a similar proposes on the idea of "open," frequently promoting unprecedented massive access to the world-class education that has so far been available only for a select group of few students. MOOC initiatives emphasize free access and interactive features rather than static content, the dominant message is of the quantity of access rather than the openness of educational resources for use, re-use, adaptation or repurpose.

In the last years, the amount of Open Educational Resources (OER) on the Web has increased dramatically, especially thanks to initiatives like OpenCourseWare (OCW) and other OER movements. The potential of this vast amount of resources is enormous but in most cases it is very difficult and cumbersome for users (teachers, students and self-learners) to visualize, explore and use these resources, especially for lay-users without experience with search technologies.

The Semantic Web is a collaborative movement led by the World Wide Web Consortium (W3C). Tim Berners Lee, the creator of the Web, coined the term (Berners-Lee, Hendler & Lassila, 2001). The Semantic Web promotes common data formats for publishing content on the World Wide Web, by encouraging the inclusion of semantic content in Web pages. The objective is to convert the current Web, dominated by unstructured and semi-structured documents, into a "Web of Linked Data."

The purpose of this paper is present a framework based on Semantic Web technologies (Bizer, Cyganiak & Heath, 2007) to support the inclusion of open materials in massive online courses. The authors focus on a type of openness: opening of contents as regards alteration i.e. freedom to reuse the material, to combine it with other materials, to adapt it, and to share it further under an open license (Hilton, Wiley, Stein & Johnson, 2010; Hodgkinson-Williams & Gray, 2009). The framework provides a service that allows people to discover and access open educational resources that are extracted from open repositories distributed. Our principal OER providers are OCW institutions. In this context, we opted to apply the principles of Linked Data (Heath & Bizer, 2011; Berners-Lee *et al.*, 2001) to integrate, interoperate and mash up data from distributed and heterogeneous repositories of open educational materials. The purpose is to significantly improve discovery, accessibility, visibility, and promote reuse of open educational content in massive courses (Piedra, Tovar, Colomo-Palacios, López & Chicaiza, 2014).

Open Educational Resources Movement

Main Purpose: ensuring wide access to quality higher education

The main purpose of Open Educational Resources (OER) movement is to provide open and free access to high quality digital learning materials. There is wide participation by universities, global and national organizations, and volunteers (Piedra *et al.*, 2013). The movement has gained an important and widely applicable effects or implications on higher education. In this work, the goal is to seek and combine OERs into a great variety of learning programs custom-made for each user of our architecture.

OERs are defined as

> teaching, learning and research materials in any medium, digital or otherwise, that reside in the public domain or have been released under an open license that permits no-cost access, use, adaptation and redistribution by others with no or limited restrictions. Open licensing is built within the existing framework of intellectual property rights as defined by relevant international conventions and respects the authorship of the work (UNESCO, 2012).

The term was coined by UNESCO (2002) at its 2002 Forum on Open Courseware for Higher Education in Developing Countries, and emphasized at their published Paris OER Declaration (UNESCO, 2012).

They are typically made freely available over the Web. Their principal use is by teachers, students and self-learners. OERs include a wide range of learning objects and free applications, from whole course, open access journals, to lecture material, references and readings, simulations, experiments and demonstrations, as well as syllabi, curricula and teachers' guides. OERs are critically important

for guaranteeing wide access to quality higher education and full participation in the rapidly evolving world higher education system.

Enhance the reusability of OER

The openness of a content can be measured in terms of the rights a user of the content is granted. One of the primary benefits of an OER is that it can be discovered and adapted to the needs of specific situations. The OER should be designed to be easily adaptable for other users. It should have metadata sufficient for discoverability. OER reusability means that the content is relevant to the specific needs of a user, which is technologically accessible and that it is sufficiently open for use, re-use, re-mix, adapt and re-distribute.

OERs discoverability: Different studies have highlighted the difficulty, finding OERs and how this affects their use. In (White & Manton, 2011) some of the causes that affect the location of OERs are identified: technical issues around search engines and repositories, practical searching skills and the volume of available resources in different subject areas. While discoverability is probably the major barrier to reuse, tutors are still expecting to find useful materials online and are prepared to spend time searching for them (White & Manton, 2011).

Open Licensing: Most of OER repositories are licensed under Creative Commons Licenses. The use of open licenses can help users to discover materials that they know can use, reuse, adapt and redistribute. Reuse is to use a resource for using as another resource, usually for a purpose unintended by the original creator. Thus creators of OER should consider the degree to which they want their OER to be open, and license the resource accordingly. In addition for licensing there are technical aspects that make OER suitable for a new use or purpose, easier to discover, adapt and remix, and consequently affect the level of openness of an OER. This implies the right to adapt, adjusts, modify, or alter the content itself.

OERs Reuse: One of the two fundamental concepts related to OER is "the ability to freely adapt and re-use existing pieces of knowledge" (Abeywardena, 2012, p. 3). (White & Manton, 2011) identified 3 factors that influence the decision to reuse of digital material: improve quality, meet a teaching need and peer suggestion.

Related Works

The need to reduce the workload for educators during the creation of learning material and the need to provide personalized learning paths according to styles and preferences of learners have been addressed in some works.

   The Wiley's concept of learning object, presented in (Koper, 2003), highlights three characteristics related to its reusability: online availability, reproducibility and addressability. In this sense, the OERs extracted from OCWs meet these features and can be reused in different contexts. Koper (2003) explores some underlying issues in the reuse of learning resources and presents these within the context of a teacher and an instructional designer who wish to reuse resources within their own practice. Moreover, Kellar, Stern, Watters and Shepherd (2004) highlight the importance of that a reuse environment includes a component that models the user's profile and thus it supports the dynamic composition of objects into personalized content.

   From the technological point of view, the system based on Adaptative Hypermedia (AH) and the intelligent tutoring systems (ITS) have been used to create courses based on existing educational material. However, as noted in (Brusilovsky & Nijhavan, 2002) the approaches based on these

technologies "are building around a close corpus material. Collecting and preparing this material to use in adaptive systems is an expensive process. Thus these systems can't directly benefit from existing repositories of learning material.

One of the reuse environments of educational material is provided by the ARIADNE Foundation (Klerkx, Vandeputte, Parra, Santos, Van Assche & Duval, 2010). The courseware re-use framework of ARIADNE allows a course author to search for the learning objects in repositories of educational material and include them in their courses (Brusilovsky & Nijhavan, 2002). Despite the large amount of resources indexed by ARIADNE, updated content could not be found and services of recommendation based on particular needs have not been located.

Brusilovsky & Nijhavan (2002) suggest a courseware-reuse approach named KnowledgeTree. It is a framework for adaptive e-learning based on distributed re-usable learning activities, the most recent version that has already been used in several courses at the University of Pittsburgh. The framework allows the presence of multiple portals, activity servers, and user-modeling servers. However, the service of resource discovery has not been addressed in the current version of KnowlegdeTree, it that is one of the core components of our proposal.

Kellar *et al.* (2004) describe an architecture that supports the dynamic composition of Web based lessons based on a database of learning objects tagged according to the IMS Metadata. A prototype has been developed specifically for an e- learning environment (users of health informatics learning modules). The scalability of this proposal can be discussed.

Other proposals found in the reuse of educational material have been defined for closed corpus, i.e. repositories of learning objects previously that have been described and classified by metadata schemas as LOM or IMS.

Instead, the OERs are published on the Web an open and extensible repository where each person can share resources; therefore, an open architecture that captures new resources and classify them according to different criteria is required to provide more relevant resources when developing MOOCs.

Regarding the web technologies and linked data there is an increasing use for topic classification and annotation of digital resources. The Semantic Web approach is about adding formal structures and semantics (metadata and knowledge) to Web content for easy access, management, discovery and integration, to make the resources machine-understandable. Some proposals, as Cano, Varga, Rowe, He and Ciravegna (2013) or Husby and Barbosa (2012) are based on the use of repositories of linked open data to determine the topics that describe social content as micropost or blogs. These and other studies have found similar findings, "DBpedia resources are a good starting point to define keyword meanings due to the fact that a huge part of the knowledge base is related to classes in the DBpedia Ontology" (Cano *et al.*, 2013).

MOOCs and Open Education

The recent emergence of MOOCs has introduced another version of potentially disruptive and likely to threaten existing practices in higher education, connected with the academic open content movement. A massive open online course (MOOC) is a model for delivering learning content online to any person who wants to take a course, with no limit on attendance. New start-up companies such as Coursera and Udacity, and the non-profit edX, have begun to provide free online access to mass-produced courses taught by leading faculty members at the world's most prestigious and prominent universities.

MOOCs represent the next stage in the evolution of open educational resources. First was open access to course content, and then access to free online courses. Accredited institutions are now accepting MOOCs as well as free courses and experiential learning as partial credit toward a degree.

Students do not pay fees to the content provider for basic enrolment in the course, nor do they receive credit from the content-providing institution. Social networking, interactive services, and automated grading or peer assessment are provided by the platform provider, as is a nominal certificate for the completion of assignments.

Although MOOCs may be considered open in the sense of "free to try," they are not offered under an open license. Any use of the content or services for academic credit-bearing purpose is restricted and requires payment to the MOOC provider. MOOCs have become attractive because their technology brings modularity to several components of higher education content such as lectures and recordable demonstrations to reach mass audiences in a flexible and accessible on-demand format.

Unlike the OCW and OER model, MOOCs promote training scenarios at large-scale of participation and open access via the web. MOOCs are a progression of the kind of open education ideals suggested by open educational resources. Though the design of and participation in a MOOC may be similar to college or university courses, MOOCs typically do not offer credits awarded to paying students at schools. However, assessment of learning may be done for certification.

Linked Data Vision

To date, most OER data are collected in heterogeneous and distributed repositories, such as OER Commons, (http://www.oercommons.org) OCW initiatives (http://www.ocwconsortium.org and http://ocw.universia.org), Merlot (http://www.merlot.org/merlot/index) and other OER repositories, where data are annotated using different metadata mechanisms [e.g. IEEE LOM (http://ltsc.ieee.org/wg12/), ADL SCORM (Advanced Distributed Learning (ADL) Sharable Content Object Reference Model (SCORM) http://www.adlnet.gov/capabilities/scorm), custom metadata schemas], and retrieved by ad-hoc mechanisms, individual Web APIs/Services or other mechanisms (e.g. OAI-PMH, Open Archives Initiative—Protocol for Metadata Harvesting http://www.openarchives.org/pmh/); however, these technologies are limited because the data cannot be de-referenced.

This work explores on how reuse, integrate and interoperate isolated OER repositories using Semantic Web Technologies. Semantic Web technologies and, more precisely, Linked Data are changing the way information is stored, published and exploited. The term "Linked Data" refers to a set of best practices for publishing and connecting structured data on the Web (Heath & Bizer, 2011; Berners-Lee *et al*., 2001). Linked data is mainly about publishing structured data in RDF using URIs rather than focusing on the ontological level or inference. OER provided with Linked Data (Linked Open Educational Resources Data) supports the process of discovery, reuse, integration and interoperability of open educational materials.

The W3C's Semantic Web provides a common framework namely Resource Description Framework (RDF) for describing resources on the Web. With RDF, automated software can store, exchange, and use machine-readable information distributed throughout the Web, in turn enabling users to deal with the information with greater efficiency and certainty; also, RDF data can be shared and reused through application, enterprise, and community boundaries.

RDF is based upon the idea of making statements about resources (in particular web resources) in the form of subject-predicate-object expressions. These expressions are known as triples in RDF terminology. The subject denotes the resource, and the predicate denotes features or aspects of the resource and expresses a relationship between the subject and the object. Uniform Resource Identifiers (URIs) are used to identify these resources. RDF Schema (RDFS) is to represent the web resource and SPARQL (Standard Protocol for RDF Query language) is to extract information from RDF graphs for machine understandable representation.

The Linked Data Design Issues, outlined by Tim Berners-Lee back in (Berners-Lee, 2006), provide guidelines on how to use standardized Web technologies to set data-level links between data from different sources (Heath & Bizer, 2011). Linked data is an opportunity to mitigate complexity in OER reuse. These Linked Data Design Issues, in OER context, are:

1. Use URIs as names for things, which can be unambiguously identified (e.g. OERs, courses, MOOCs, OER creators, OCW providers, knowledge areas,)
2. Use HTTP URIs so that people can look up those names. With the aid of URIs, the corresponding OER data and relevant interlinked data can be dereferenced.
3. When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL) to describe linked OER data, which are machine-readable and repurposed to serve the proposed architecture to enhance integration with reused and interoperated OER data.
4. Include links to other URIs, so that they can discover more entities. Linked Data—particularly data available using open licenses—has an important role to play on information systems and could be a key feature for Open Education based on OER data on the Web of Data.

In Piedra *et al*. (2014), authors apply the Linked Data Design Issues to explore, visualize and use information that is semantically related to open educational resources that are accessible via the OCW Consortium. Linked data have the potential of create bridges between OCW data silos. The authors demonstrate that OCW resource metadata can be enriched using datasets hosted by the Linked Open Data cloud. Additionally, the Linked OER and OCW Data environment enabled us to discovery and reuse open educational materials.

Architecture of OER-reuse

The vision of Semantic Web is the idea of having data on the Web described and linked in a way that it can be used by machines not just for display purposes, but for automation, interoperability, integration and reuse of data through various applications and contexts. It provides a promising platform for Open Educational Initiatives. In this section, we describe our proposal for the integration and reuse of OER: an architecture based on Linked Data technologies**.**

The main objective of this work is to propose a linked OER data architecture (Figure 1), able to adapt, reuse and re-mix OERs in the MOOC context. The architecture is composed by five services which have been designed to carry out this task in a collaboratively way. This linked data architecture enabling us to ask questions and solve open educational problems across a heterogeneous and distributed information landscape extending beyond the traditional boundaries of each OER contributor.

Our approach is based on identifying distinctive features with the help of MOOC preferences and resources needs data. As with all recommender systems, the main goal is to help users to find information or resources and match information that is important about needs with information that is important about resources. Figure 1 summarizes the architecture in a general model of OER recommendation for MOOC Designers. Accordingly, the process can be broken down to the following steps:

1. OER collecting.
2. OER Data Publication and Semantic Annotation of OER.
3. MOOC profiles provider.
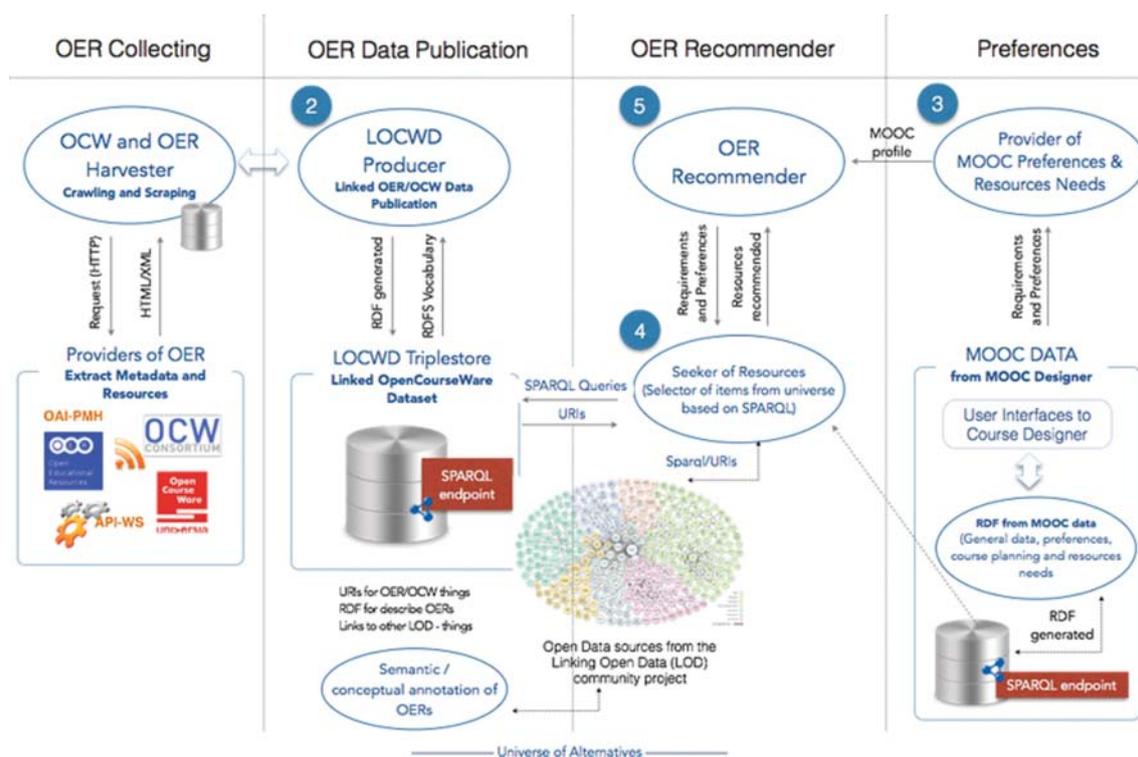4. Seeker of resources.
5. OER recommender.

**Figure 1: Architecture for Seeker of OER from Linked OpenCourseWare and
OER Data (LOCWD) Triplestore**

Component 1, OER collecting

Goal: Identify and select OER repositories, then extract metadata and educational resources with Open Licenses

Description: The initial step is to identify and select the OER repositories that are available in the Web. There is a large amount of unstructured data of an OCW resource available on the Web, but only in a human-readable representation, HTML.

Most OCW web sites do not have APIs for data consumption. So, the only other alternative for automatically reconstitute the underlying data from an OCW web site is to use web-scraping techniques (Piedra *et al*., 2014).

Examples of extracted OER properties include the name of the resource, its creation date, abstract, keywords, information about creator, language, open license information, format, MIME type, expected study duration, expected level of difficulty, and so on. On the other hand, content metadata correspond to the properties of the knowledge and skills designed, such as learning objectives, learning pathways, and examinations.

Component 2, OER—Data Publication and Semantic Annotation of OER

Goal: Development and delivery of open educational resources data as Linked Data.

Description: Linked Data design principles are increasingly employed to publish and consume heterogeneous datasets in a distributed way. Data is still locked up in applications. The technical problem with today's most common information architecture is that content, metadata and schema information are not separated well from application logics and presentation layer. Data cannot be re-used as easily as it should be.

Publication of OER data:

Using Linked Data design issues, developers can query Linked Data from multiple sources at once and merge it without the need for a single common schema that all data shares. Linked data technologies can also help to integrate the work of disperse institutions producing diverse linked data.

The following is an outline for producing Linked Data in OER context:

1. Identify and select heterogeneous repositories.
2. Model vocabularies for OER domain.
3. Data extraction from OER repositories.
4. Generate OER data as Linked Data.
5. Publish linked data.
6. Consume and display linked data.

Vocabularies and ontologies provide the mechanism to organize the Web information in structured way. The web contents can be understood by the computer as well as by human beings. Piedra *et al*. (2014) described LOCWD RDFS vocabulary using W3C's RDF technology, for open educational resources with the aim to describe the specific types and classes of resources in OCW domain. This vocabulary was called Linked OpenCourseWare Data (LOCWD). A machine-friendly version is also available in http://purl.org/locwd/schema on RDF/XML format. LOCWD reuses a set of RDF(S) vocabularies. Each vocabulary includes a set of terms and classes that are common to a particular knowledge domain. The aim of these vocabularies is to connect the described OCW domain with Datasets in the LOD cloud.

LOCWD is a RDF(S) vocabulary devoted to linking OERs, open licenses, OCW repositories, and other academic information using the Web. Different kinds of applications can use or ignore different parts of LOCWD. With LOCWD, the OER/OCW initiatives can retain some control over their information of materials and courses in a non-proprietary format.

Semantic annotation:

In this section we also focus on semantic or conceptual annotation of OERs, which consists in attaching "semantic labels" to a resource or parts of a course using semantic features provided by a formal knowledge organization called ontology or semantic vocabulary. The use of ontologies in the annotation process of OERs has a particular benefit: the queries for information retrieval and annotated resources share the same vocabulary.

We present an approach to semi-automatic semantic or conceptual annotation of OER that aims at bridging the gap between the ontology formal language and user's natural language terms or keywords. The approach provides candidate concepts for a user's term.

The primary external source of semantic annotation of OER is DBpedia, the most important part of the Linked Open Data. The Linking Open Data cloud currently provides access to hundreds of datasets in various areas such as Media, Geography, Publications, Government, and Life Sciences. As a consequence of Linking Open Data community project, datasets in a wide range of domains are now semantically described and connected to each other.

Some of the key resources for OER annotation in DBpedia include "dbpedia:Category" and the SKOS semantic relations between SKOS concepts that is inherent in the meaning of the linked concepts that provide information about concepts (see Figure 2). In this work, SKOS properties (W3C, 2009) are used to infer hierarchical links, which can then be used to annotate resources and access direct or indirect hierarchical links between OER and concepts.
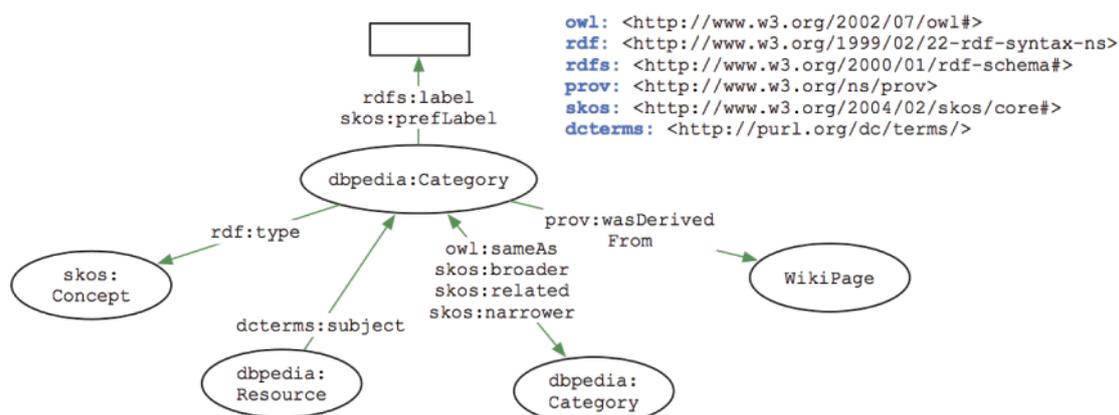
Figure 2: Key resources used for OER annotation and enrichment from DBPedia

Several relations link resources to their concepts, such as "owl:sameAs," "skos:broader" (a hierarchical link between two concepts what indicates that one is in some way more general -broader- than the other -narrower-), "skos:narrower," "skos:related" (an associative link between two concepts indicates that the two are inherently "related," but that one is not in any way more general than the other) is used to assert an associative link between two SKOS concepts., and "dcterms:subject" to describe the topic of the resource (typically, the subject will be represented using keywords, key phrases, or classification codes. In DBpedia is used as a controlled vocabulary). The properties skos:broader and skos:narrower are used to assert a direct hierarchical link between two resources using SKOS concepts.

In the architecture proposed, it is necessary to make use of both direct and indirect hierarchical links between concepts, to improve search recall through query expansion. For this purpose, the properties skos:broaderTransitive and skos:narrowerTransitive are provided (W3C, 2009).

Component 3, MOOC profiles provider

Goal: Extract Data from a course profile that serves as a view of filter onto the whole universe.

Description: The idea is evolving into a more interoperable and integrated system to share, connecting and discovering data and metadata of MOOC profiles. Users don't know precisely what they can find on OER site, or what to search for. Self-learners are trying to discover relationships or trends between MOOC profile and OER data.

Component 4, Seeker of resources

Goal: Merge the functionalities of recommendation seeker and Course Profile.

Description: The architecture uses seekers based on SPARQL to express preferences and resources needs by rating OERs. The system focuses on SPARQL query-based algorithms for matching OER based on MOOC preferences and weighting the interest of MOOC designer with similar taste to produce a recommendation for the resources seeker. Table 3 shows the properties of a topic for a Java Course.

This module is designed to access two data sources: The first one reline on LOCWD data, which provides RDF data extracted from OCW and OER websites. The second one use the LOD Project, particularly from DBPEDIA, which provides RDF data extracted from the info-boxes of Wikipedia pages in a structured way.

Component 5, OER Recommender

Goal: For our purpose, a preference is an individual mental state concerning a subset of items from the universe of alternatives. Users can use the architecture because a single taxonomic order or a single folksonomy is not suitable or sufficient for explorer OER resources.

Description: The architecture proposed attempt to recommend OERs that are similar to educational resources planned by the MOOC designer and others records of social activity, such as OCW Syllabus and system usage history.

Performance Evaluation

Experimental Context: Selection and Collection of OER repositories

Scope: 15 associate consortia, as well as 212 higher education institutions and 57 organizational members compose OCWC; all courses are available for adoption and adaptation by faculty and students around the world.

When publishing an OER Repository, the providers are sharing digital educational resources with an attached open license that allowing others to reuse, adapt and share their work. Ideally, this resource when combined with others OERs provides great value.

On this paper, we focus on finding OER published by open licenses and useful to MOOC or Open Course Ware production. Authors selected and extracted information from 80 heterogeneous OCW repositories from OCWC and OCW-Universia members (Piedra *et al*., 2014), sifting through a total of 7,239 OCW courses and 90.000 OERs approx. Data scraping was used to extract data from OCW platforms that was later structured and stored in a database. Scraping eliminated the need for having to do the retrieval manually.

In this context, authors opted to apply the principles of Linked Data to integrate, interoperate and mash up data from distributed and heterogeneous repositories of open educational materials. The purpose is to significantly improve discovery, accessibility, visibility, and to promote reuse of open educational content in massive course.

Integration and reuse of OER in MOOCs Context

In the context of this paper, authors opted to apply the design issues of Linked Data to integrate, interoperate and mash up OER data from MOOC Designer requirements. See table 1 by examples of MOOC profile properties.

SPARQL queries are used to semantically annotate OER materials. SPARQL is a query language designed to gather data from multiple sources for anything that asks a question. The next code shows the SPARQL Query that permits annotating the resource Java Programing, an entity of type Concept (see table 2).

SPARQL Query:

```
SELECT distinct ?area ?labelarea ?relatedTopic ?labelrelatedTopic WHERE {
    {<http://dbpedia.org/resource/Category:Java_programming_language> <http://www.w3.org/2004/02/skos/core#broader> ?area.
    ?area rdfs:label ?labelarea .
    FILTER( lang(?labelarea) = "en" )
    ?relatedTopic <http://purl.org/dc/terms/subject> ?area .
    ?relatedTopic rdfs:label ?labelrelatedTopic .
    FILTER( lang(?labelrelatedTopic) = "en" )
    }
```

Table 1: Some properties for describe a MOOC profile for Java Course in tripletes

Course as RDF resource	Property	Value
:JavaCourse	rdf:type	:MOOCProfile
:JavaCourse	:title	Java Fundamentals
:JavaCourse	:description	This course will cover the main concepts about Java. Students will learn the fundamentals of Java. The focus is on developing high quality, working software that solves real problems.
:JavaCourse	:language	English
:JavaCourse	:alternative_language	Spanish
:JavaCourse	:level	Basic
:JavaCourse	:requirements	The course is designed for students with some programming experience
:javaCourse	:learningOutcome	Learners will be guided through the fundamentals of object-oriented programming on the Java platform.
:JavaCourse	:relatedConcept	"Java"
:JavaCourse	:relatedConcept	"Programming"

Table 2: Some properties and values about the entity Java Programing

Property	Value
owl:sameAs	http://dbpedia.org/resource/Category:Java_programming_language
rdf:type	skos:Concept
rdfs:label	Java programming language
is skos:broader of	category:Java_APIs category:Java_specification_requests category:Eclipse_(software) category:Software_programmed_in_Java category:Articles_with_example_Java_code category:Java_libraries category:Java_programmers

```
UNION {
    ?area <http://www.w3.org/2004/02/skos/core#broader> <http://dbpedia.org/resource/Category:Java_programming_language> .
    ?area rdfs:label ?labelarea .
    FILTER( lang(?labelarea) = "en" )
    ?relatedTopic <http://purl.org/dc/terms/subject> ?area .
    ?relatedTopic rdfs:label ?labelrelatedTopic .
    FILTER( lang(?labelrelatedTopic) = "en" )
    }
}
```

Table 3: A specific topic for Java Course: Introduction to Java programming

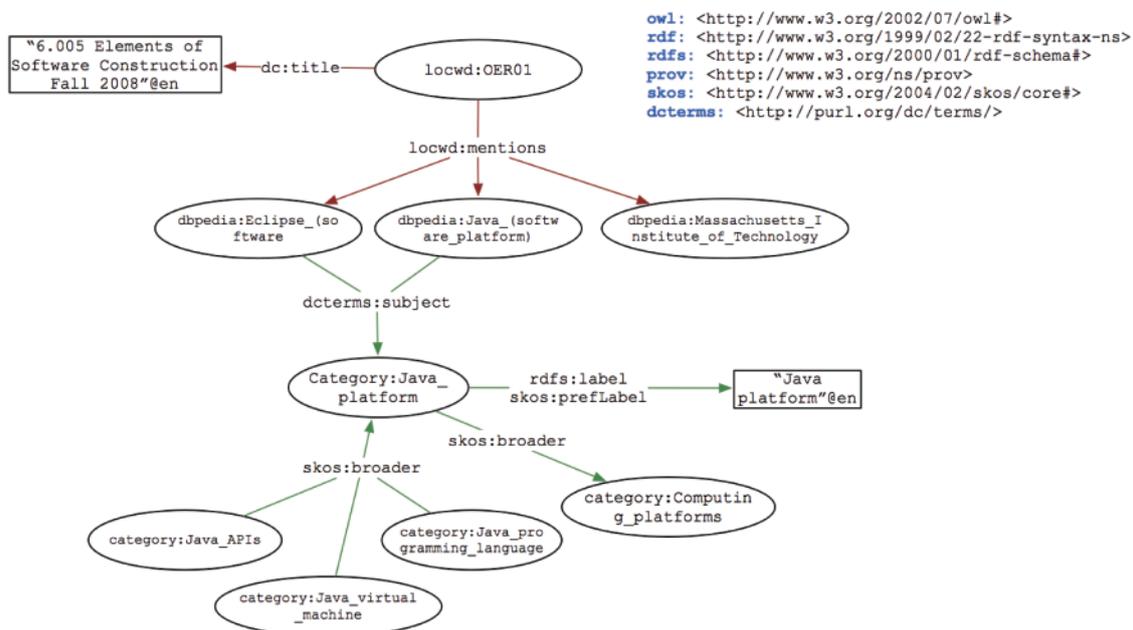| Property	Value
:topic	Introduction to Java programming
:description	Studying the necessary elements that allow you to build simple Java programs
:time	One week
:subTopic	Programming Introduction Java program structure Program flow Arithmetic Operators Primitive data types

Figure 3: An example of annotation of OER from DBPedia. The annotation process not only provides access to a large amount of structured data sources but also enables machines and software agents to automatically analyze this semantic knowledge

An example of annotation of OERs can be seen in Figure 3. This graph structure of Linked Data consists of resources described and connected with other concepts through semantic relations.

Table 4 shows the resources retrieved from DBPedia related to "programming and Java." The query returned 899 related subjects.

Table 5 shows some of the resources returned. The major complexity with the existing query is the lack of information on the particular domain, how to locate the accurate data in repository.

Linked Data technologies can also help to integrate the work of disperse institutions producing diverse linked data. Linked Open Data (LOD) is well known for providing an extensive amount of detailed and structured information (see LOD Cloud maintained by Cyganiak & Jentzsch, 2011).

As we discussed, the architecture not only uses LOCWD dataset, but also uses information from Linked Open Data project. This allows exploiting the LOD community benefits. Table 6, summarizes OER found in LOCWD data source.

Table 4: Topics and Subject related to Programing and Java

URI of Category	Related subjects
http://dbpedia.org/resource/Category:Java_platform	231
http://dbpedia.org/resource/Category:Object-oriented_programming_languages	162
http://dbpedia.org/resource/Category:Java_libraries	86
http://dbpedia.org/resource/Category:Concurrent_programming_languages	83
http://dbpedia.org/resource/Category:Java_specification_requests	77
http://dbpedia.org/resource/Category:Articles_with_example_Java_code	55
http://dbpedia.org/resource/Category:Class-based_programming_languages	48
http://dbpedia.org/resource/Category:Java_APIs	37
http://dbpedia.org/resource/Category:Eclipse_(software)	36
http://dbpedia.org/resource/Category:Sun_Microsystems	33
http://dbpedia.org/resource/Category:Java_programming_language_family	28
http://dbpedia.org/resource/Category:Software_programmed_in_Java	21
http://dbpedia.org/resource/Category:Java_programmers	2

The framework provides transparent access to RDF data sources for OER stored in OCW repositories. In the architecture, the recommendation seeker is based on SPARQL. With Sparql it is possible to filter OER using multiple category or taxonomy terms at the same time, and combine text searches, category term filtering, and other search criteria (table 7). Then, it may ask for an OER recommendation based on MOOC data profile (Module 3).

Conclusions

The use of linked data approach on OCW repositories provides the framework for their evolution into a more interoperable and integrated system for sharing, connecting and discovering data and metadata of OER and OCW initiatives. The framework provides an approach that allows to MOOC designers to discover and access open educational resources that are extracted from open repositories distributed; our principal OER providers are OCW institutions.

This study advocates the use of Linked Data technologies as an enabler for the development of the next generation of Open Educational Resources, allowing the separation of semantics from syntax, the improvement of discoverability and access, and the use of common vocabularies. In addition, the proposed architecture provides to data consumers an opportunity to merge data distributed across different libraries.

Acknowledgment

Table 5: Some of the resources returned

Concept URI	Concept label	Related Topic URI	Related Topic URI
dbpedia:Category:Java_platform	Java platform	dbpedia:Classpath_(Java)	Classpath (Java)
dbpedia:Category:Java_platform	Java platform	dbpedia:Jreality	Jreality
dbpedia:Category:Java_platform	Java platform	dbpedia:Virtual_Database_Manager	Virtual Database Manager
dbpedia:Category:Java_platform	Java platform	dbpedia:JavE	JavE
dbpedia:Category:Java_platform	Java platform	dbpedia:Java_security	Java security
dbpedia:Category:Java_programmers	Java programmers	dbpedia:Markus_Persson	Markus Persson
dbpedia:Category:Java_programming_language_family	Java programming language family	dbpedia:BeanShell	BeanShell
dbpedia:Category:Java_programming_language_family	Java programming language family	dbpedia:Scala_(programming_language)	Scala (programming language)
dbpedia:Category:Java_programming_language_family	Java programming language family	dbpedia:Processing_(programming_language)	Processing (programming language)
dbpedia:Category:Java_specification_requests	Java specification requests	dbpedia:Java_Persistence_API	Java Persistence API
dbpedia:Category:Java_specification_requests	Java specification requests	dbpedia:Java_virtual_machine	Java virtual machine
dbpedia:Category:Java_specification_requests	Java specification requests	dbpedia:Java_3D	Java 3D
dbpedia:Category:Java_specification_requests	Java specification requests	dbpedia:Java_Database_Connectivity	Java Database Connectivity
dbpedia:Category:Java_specification_requests	Java specification requests	dbpedia:Java_Servlet	Java Servlet

**Table 6: Summary of Resources found in LOCWD data source for Topic
"Introduction to Java programming"**

Number of recommended resources	SubTopic	Kind of resources recommended
8	Arithmetic Operators	Learning guides, Lecture, Exercises, Tutorial
7	Java program structure	Learning guides, Lecture, Exercises, Labs
3	Program flow	Lecture, Exercises, Tutorial
7	Programming introduction	Lectures, Learning resource, Tutorial, Labs
7	Primitive data types	Learning guides, Lectures, Exercises, Tutorial, Labs, Tutorial

Table 7: OER recommended for subtopic: Programming introduction

OER Title	Kind of resource	URL of OER recommended	Some metadata
6.170 Laboratory in Software Engineering. Java Style Guide	Learning material	http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-092-java-preparation-for-6-170-january-iap-2006/study-materials/java_style.pdf	:Language English :Provider MIT :File java_style.pdf
1: Types, Variables, Operators	Lecture	http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-092-introduction-to-programming-in-java-january-iap-2010/lecture-notes/MIT6_092IAP10_lec01.pdf	:Language English :Provider MIT :File MIT6_092IAP10_lec01.pdf
Problems: Getting Started	Lab	http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-092-java-preparation-for-6-170-january-iap-2006/labs/problems1_4.pdf	Language English :Provider MIT :File problems1_4.pdf
Aprenda java como si estuviera en primero	Tutorial	http://ocw.uc3m.es/ingenieria-informatica/programacion/manuales/java2-U-Navarra.pdf	:Language: Spanish :Provider: UC3M :File java2-U-Navarra.pdf
Otros ejercicios propuestos del capítulo de operadores	Lecture	http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-092-introduction-to-programming-in-java-january-iap-2010/lecture-notes/MIT6_092IAP10_lec02.pdf	Language English :Provider MIT :File MIT6_092IAP10_lec02.pdf
Capítulo 1: Introducción a la Programación	Learning material	http://ocw.utpl.edu.ec/sistemas-informaticos-y-computacion/fundamentos-de-la-programacion/capitulo1-introduccionprogramacion.pdf	:Language: Spanish :Provider: UTPL :File capitulo1-introduccionprogramacion.pdf
Tema 2. Fundamentos de Java	Lecture	http://ocw.uc3m.es/ingenieria-informatica/programacion/transparencias/tema2.pdf	:Language: Spanish :Provider: UC3M :File tema2.pdf

References

Abeywardena, I. S. (2012) *The Re-use and Adaptation of Open Educational: Resources (OER). An Exploration of Technologies Available*. Commonwealth of Learning. Retrieved from http://www.col.org/PublicationDocuments/ExplorationOfTechnologiesAvailable_OER.pdf

Berners-Lee, T. (2006) Linked Data: Design Issues. *W3C*. Retrieved from http://www.w3.org/DesignIssues/LinkedData.html

Berners-Lee, T., Hendler, J. & Lassila, O. (2001). The Semantic Web. *Scientific American*, *284*(5), 34–42.

Bizer, C., Cyganiak, R. & Heath, T. (2007). *How to Publish Linked Data on the Web*.

Brusilovsky, P. & Nijhavan, H. (2002). A Framework for Adaptive E-Learning Based on Distributed Re-usable Learning Activities. *Proceedings of World Conference on E-Learning, E-Learn 2002.*

Cano, A., Varga, A., Rowe, M., He, Y. & Ciravegna, F. (2013). Harnessing Linked Knowledge Source for Topic Classification in Social Media. *Proceedings of the ACM Hypertext and Social Media Conference*. Paris, France.

Cyganiak, R. & Jentzsch, A. (2011). *The Linking Open Data cloud diagram*. Retrieved from http://lod-cloud.net/

Heath, T. & Bizer, C. (2011). Linked Data: Evolving the Web into a Global Data Space. *Synthesis Lectures on the Semantic Web: Theory and Technology*, *1*(1), 1–136. http://dx.doi.org/10.2200/S00334ED1V01Y201102WBE001

Hilton, J., Wiley, D., Stein, J. & Johnson, A. (2010). The four R's of openness and ALMS Analysis: Frameworks for Open Educational Resources. *Open Learning: The Journal of Open and Distance Learning*, *25*(1), 37–44. Retrieved from http://contentdm.lib.byu.edu/cdm/ref/collection/IR/id/774

Hodgkinson-Williams, C. & Gray, E. (2009). Degrees of Openness: The emergence of Open Educational Resources at the University of Cape Town. *International Journal of Education and Development using Information and Communication Technology* (IJEDICT), *5*(5), 101–116. Retrieved from http://ijedict.dec.uwi.edu/viewarticle.php?id=864

Husby, S. D. & Barbosa, D. (2012) Topic Classification of Blog Posts Using Distant Supervision. *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics* (pp. 28–36).

Kellar, M., Stern, H., Watters, C. & Shepherd, M. (2004). An Information Architecture to Support Dynamic Composition of Interactive Lessons and Reuse of Learning Objects. *Proceedings of the 37th Hawaii International Conference on System Sciences.*

Klerkx, J., Vandeputte, B., Parra, G., Santos, J. L., Van Assche, F. & Duval, E. (2010). How to Share and Reuse Learning Resources: The ARIADNE Experience. *Sustaining TEL: From Innovation to Learning and Practice Lecture Notes in Computer Science,* 6383, (pp. 183–196).

Koper, R. (2003). Combining reusable learning resources and services to pedagogical purposeful units of learning. In A. Littlejohn (Ed.), *Reusing Online Resources: A Sustainable Approach to eLearning* (pp. 46–59). London: Kogan Page.

Petrides, L., Nguyen, L., Jimes, C. & Karaglani, A. (2008). Open educational resources: inquiring into author use and reuse. *International Journal of Technology Enhanced Learning, 1*, 98–117.

Piedra, N., Tovar, E., Colomo-Palacios, R, López, J. & Chicaiza, J. (2014). Consuming and producing linked open data: the case of Opencourseware. *Program: electronic library and information systems*, *48*(1), 16–40. http://dx.doi.org/10.1108/PROG-07-2012-0045

Piedra, N. Tovar, E. Dimovska, A. & Chicaiza, J. (2013). OCW-S: enablers for building sustainable Open Education. *Proceeding of IEEE Global Engineering Education Conference (EDUCON)* (pp. 1262–1271). Berlin, March 2013.

UNESCO. (2002). *Forum on the Impact of Open Courseware for Higher Education in Developing Countries. Final report*. Retrieved from http://portal.unesco.org/ci/en/files/2492/10330567404 OCW_forum_report_final_ draft.doc/OCW_forum_report_final_draft.doc

UNESCO. (2012). 2012 *Paris OER Declaration*. 2012 World Open Educational Resources (OER) Congress. Paris, June 20–22. Retrieved from http://www.unesco.org/new/fileadmin/ MULTIMEDIA/HQ/CI/CI/pdf/Events/Paris%20OER%20Declaration_01.pdf

White, D. & Manton, M. (2011). *Open Educational Resources: The value of reuse in higher education. JISC-funded OER Impact Study*. University of Oxford.

W3C. (2009, August 18). *SKOS Simple Knowledge Organization System. W3C Recommendation.* Retrieved from http://www.w3.org/TR/skos-reference/