THE UNIVERSITY OF WARWICK

Original citation:
Al Fayez, Reem and Joy, Mike (2014) A framework for linking educational medical objects : connecting Web 2.0 and traditional education. In: 15th International Conference on Web Information System Engineering (WISE 2014), Thessaloniki, Greece, 12-14 Oct 2014. Published in: Web Information Systems Engineering – WISE 2014 : 15th International Conference, Thessaloniki, Greece, October 12-14, 2014, Proceedings, Part II, Volume 8787 pp. 156-165.

Permanent WRAP url:
http://wrap.warwick.ac.uk/62765

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions. Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:
"The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-11746-1_11"

A note on versions:
The version presented here may differ from the published version or, version of record, if you wish to cite this item you are advised to consult the publisher's version. Please see the 'permanent WRAP url' above for details on accessing the published version and note that access may require a subscription.

For more information, please contact the WRAP Team at: publications@warwick.ac.uk

warwickpublicationswrap

highlight your research

http://wrap.warwick.ac.uk

A Framework for Linking Educatioal Medical Objects: Connecting Web2.0 and Traditional Education

Reem Qadan Al Fayez and Mike Joy

Computer Science, University of Warwick, Coventry, CV4 7AL, UK
{r.qadan-al-fayez,m.s.joy}@warwick.ac.uk

Abstract. With the emergence of Linked Data principles for achieving web-scale interoperability, and the increasing uptake of open educational content across institutions, Linked Data (LD) is playing an important role in exposing and sharing open educational content on the web. The growing use of the internet has modified quickly our learning habits. Learning in the medical field is unique in its nature. The educational objects are of various types and should be published by trustworthy organizations. Therefore, medical students and educators face difficulties locating educational objects across the web. To address this problem, we propose a data model for describing educational medical objects harvested from the World Wide Web (WWW). The data model is published in Linked Data format in order to build dynamic connections between educational objects harvested from distributed resources. To reduce the burden of navigating through the overwhelming amount of information on the web, we provide a harvesting engine for collecting metadata objects from specified repositories. Then, the harvested educational objects' metadata are represented in our proposed data model named Linked Educational Medical Objects (LEMO). Further enrichment is applied on the metadata records by annotating the textual elements of the records using biomedical ontologies to identify terms inside text and connect them to the Linked Open Data cloud. In this paper, we present the framework proposing the data model LEMO along with its implementation and experiments conducted.

Keywords: Linked Data, Web Data Mode, Medical Ontologiesl

1 Introduction

The volume of what can be considered an educational material is rapidly increasing on the web. This leads to an increased attention to the concerns about managing such online objects throughout the World Wide Web (WWW). The factors behind the growing number of educational content on the web can be narrowed down to two basic factors: firstly, modern techniques applied in education which encourage the move towards concept-based learning instead of content-based learning. Secondly, the incorporation of web2.0 technologies in education such as wikis, blogs, and social networks [1].

As any area of education, educational content on the web plays a significant role in Medical Education. Multimedia objects such as videos, virtual patients and pictures have been heavily used medical e-curricula [2] in the past years. This emphasizes the importance of searching and using such educational objects for both users and educators. Also, researchers in the field of medical education have been studying the potential use of web2.0 technologies in enhancing the learning and teaching experience. Studies such as the ones presented in [3], explains the potential impact which web 2.0 technologies can have on medical education for both students and educators.

Searching and discovering the educational content needed by a user depends on both the search terms he/she uses when searching for online conttent, and on the description publicly provided by the educational object's publisher which is the Metadata of an object. Therefore, Metadata is considered the key component for publishing and managing online content. It has been defined as "data describing the context, content, and structure of records and their management through time" [4]. The definition is wide and accommodates different interpretations of what metadata can be. Hence, there is no perfect metadata standard which fits varying requirements of different organizations and can be called an ideal standard. In e-learning, different standards exist for describing educational materials depending on the organization publishing that content [5]. Every proposed standard or application profile focuses on specific features to emphasize depending on the requirements of the developing organization. Many challenges face the field of metadata such as interoperability of the metadata, which is ranked as the top challenge in digital libraries [6]. Other challenges are also important to consider such as the special nature of the application domain for educational objects to be described such as medical educational materials. More details about such standards will be presented in the following section.

In this paper we propose a framework for harvesting and exposing different types of what can be considered medical educational object on the web. According to a comparative study conducted on existing standards and conventions for publishing educational objects, we present a comprehensive data model for storing metadata records harvested from different libraries and websites hosting educational content. The proposed data model applies Linked Data structure in order to easily manage textual content of records and wheere keywords can be annotated and links can be built between the collections of objects harvested. We incorporate annotation of textual content of an object using well recognized medical ontologies in order to annotate that object with keywords and categories it might belong to. This paper represenst a work in progress which requires iterations of experiments and enhancements. In this paper we discuss in details the data modelling framework and present the data model proposed along with results of conducted experiments for harvesting and modelling educational objects using the proposed model.

The rest of the paper is organized as follows. Section 2 presents the related background of existing standards along with a comparative study between these standards. Section 3 introduces an overview framework of the proposed data model. Section 4 describes the development of the data model in details. Further techniques incorporated in designing and testing the data model are explained in section 5. Experiments conducted on harvesting, mapping and annotating educational content are detailed in section 6. Finally, section 7 concludes the paper and discusses further future plans.

2 Background

In this section, we first introduce definitions necessary to understand the paper content. We then give an overview of existing metadata standards applied in e-learning in general and in medical education in particular. Following, results of comparative study conducted between the metadata schemas used in medical education are presented. A standard, as defined by the British standard institute, is a published document that contains a technical specification or other precise criteria designed to be used consistently as a rule, guideline, or definition. Standards are created by bringing together the experience and expertise of all interested parties such as the producers, sellers, buyers, users, and regulators of a particular material, product, process or service. In the context of e-learning, standards are used to govern the design, description, and publishing of educational materials for the purpose of ensuring reusability, and interoperability between different e-learning systems. It is possible that a standard does not satisfy specific systems needs. Hence, developing an Application Profile (AP) is a flexible way to adopt a standard and satisfy the requirements of a system. Any system or organization can implement a modified version of any standard, either by adding or removing characteristics from the standards set of attributes, and still guarantee its compatibility with the original standard, and the interoperability of its content [7].

2.1 Existing Metadata standards

An overview of popular metadata standards and application profiles developed in the medicine field is presented in order to give the needed knowledge about existing data models. Two of the most popular metadata standards used for e-learning are IEEE Learning Objects Metadata (LOM) and Dublin Core Metadata Initiative (DCMI). Several Application Profile where developed in the medical field having one of these standards as their base schema. IEEE LTSC started to develop **Learning Object Metadata (LOM)** standard in 1997. With the aid of different international participants, LOM working group succeeded to have this standard accredited by IEEE [8]. As defined in LOM standard [9], a LOM instance is designed to record the characteristics of the learning object it describes. LOM categorizes these characteristics into 9 groups: general, life cycle, meta-metadata, educational, technical, rights, relation, annotation, and classification. The purpose of such detailed description for a learning object as stated by the working group is to facilitate sharing and exchange of learning objects since the metadata have a high degree of semantic interoperability. **Dublin Core**

Metadata Initiative (DCMI) is an open public organization that is non profitable. It supports metadata design and implementation across broad range of purposes. The initiative work resulted in a simple cross domain metadata statement known as Dublin Core Metadata Element Set (DCMES) which has been standardized as ISO standard [10]. Dublin Core standard is used to describe a wide range of resources, where a resource is defined by DCMI as "anything which might be identified" [11]. The simple Dublin Core metadata element set consists of these fifteen elements: *contributor, coverage, creator, date, description, format, identifier, language, publisher, relation, rights, source, subject, title, and type*. All of these elements are optional and maybe repeated if required for a single resource.

Metadata schemas or Application Profiles (APs) were developed by organizations responsible of libraries and repositories for publishing medical educational materials. A brief overview of four popular AP is presented and comparative study is conducted on these four APs in order to conclude set of common elements across these schemas. **HealthCare LOM**[1], which was developed within the MedBiquitous Learning Object Working Group, extends the IEEE Learning Object Metadata (LOM). In this schema, one category named HealthCare Education was added to the original 9 categories composing LOM, in which fields to describe health care related metadata were added such as continuing education credits, patient and professional resources, and others [12]. **mEducator**[2] was developed as part of mEducator project, which is a European funded best practice network launched in order to host medical educational materials from European higher academic institutions. The objective of mEducator schema was to enable ease of sharing, discovery, and reuse of medical educational content across EU higher academic institutions [13] Another AP developed for managing medical educational content is **The Health Education Assets Library (HEAL)**[3]. HEAL project is a result of collaboration between numerous faculties, medical schools, the Association of American Medical Colleges, and the National Library of Medicine. HEAL metadata schema extends Educause IMS which is based on IEEE LOM specification and builds upon it specific fields needed in health science education such as clinical history field and disease process field [14]. Finally, **National Library of Medicine (NLM)**[4] implemented by the world's largest biomedical library. NLM schema is based on Dublin Core Metadata Initiative and incorporates additional elements identified as requirements by NLM for publishing its content [15].

2.2 Existing Metadata Schemas in Medical Educational Libraries: a Comparative Study

A comparative study is conducted in order to identify common elements used in the different AP schemas described in the background section. The results of the

[1] http://ns.medbiq.org/lom/extend/v1
[2] www.meducator.net/mdc/schema.rdf
[3] http://www.healcentral.org
[4] http://www.nlm.nih.gov/tsd/cataloging/metafilenew.html

comparative study is outlined in table 1. The tables lists the common elements between the four schemas and how they are represented in each schema. An element in any metadata schema might be represented as a Parent Node (PN) which states that it is a composite element of other leaf nodes, or it might be Leaf Node (LF) which states that this element is a child node of another PN, or it might be a Single Node (SN) which states that the element is represented as a node directly descendent from the root and does not have children.

Table 1. Results of comparative analysis conducted on four medical metadata schemas

	HealthCareLOM		mEducator		HEAL		NLM	
1	General	PN	General	PN	General	PN	General	-
2	Identifier	LN	Identifier	LN	Resource URN	LN	Identifier	PN
3	Title	LN	Title	LN	Title	LN	Title	SN
4	Description	LN	Description	LN	Description	PN	Description	PN
5	Lifecycle	PN	Lifecycle	PN	Creation Date	LN	Date	PN
6	Rights	PN	Rights	PN	IPR	LN	Rights	SN
7	Resource Type	LN	Resource Type	LN	Resource Type	LN	Resource Type	SN
8	Keywords	LN	Keywords	LN	Keywords	LN	Keywords	LN
9	Classification	PN	Classification	PN	Classification	PN	Subject	PN
10	Educational	PN	Educational	PN	Educational	PN	Educational	-
11	Relation	PN	Relation	-	Companion	PN	Relation	SN
12	Technical	PN	Technical	PN	Technical	LN	Technical	-

From the analysis of the four metadata schemas, 12 elements have been found common between these four schemas as shown in table 1. Identifier, Title, and Description are main elements in all of the four schemas and in most cases they are LN or PN. Another common attribute is the Date, which might be used to capture several types of dates such as important revision dates or re-purposing of the resource using the Lifecycle. Intellectual Property Rights is another common element where in some schemas the element is used to state copyrights and other restrictions such as cost if required. All of the four schemas describe the type of the resource using Type property. Also, Keywords and Classification elements are from the common elements between the entire collection of metadata schemas analysed. The last three attributes in table 1 Educational, Relation, and Technical are not implemented in all four schemas. The elements used in each schema represent the need of the organization developing the AP. From the analysis above, we can say that the goal of the application or the organization for which an AP is developed plays vital role in the elements chosen and their structure in the schema.

3 Overview of System Framework for Developing LEMO Data Model

Several hypothesis were derived after conducting the comparative study. The main aim for proposing this data model is having a metadata in which we can describe different types of educational objects, in addition to the ability of building dynamic connections between the educational objects. The hypothesis focused on the set of elements necessary to have in the proposed data model to achieve its goal which are:

1. General content descriptors of the objects content
2. People responsible for the development of the object
3. The category in which the object might be classified into
4. The objects type
5. The objects access rights

The framework shown in figure 1 illustrates the main aims of proposing the data model and the conceptual model design on this model.

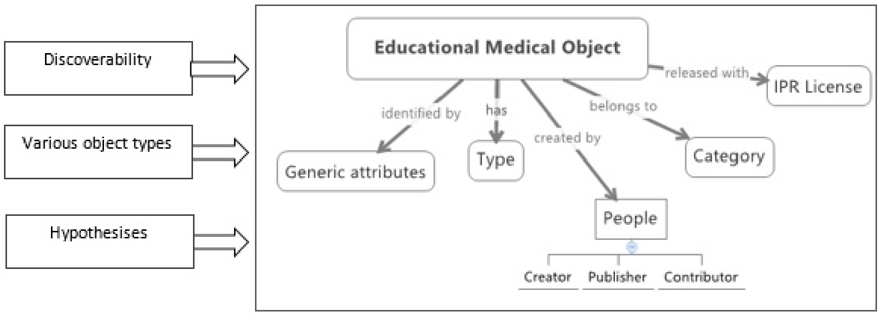

Fig. 1. An Overview of Our Framework for the proposed Data Model Design

The proposed data model focuses on discoverability since the main goal of the project is to expose and link harvested medical educational objects from various sources having different metadata standards or schemas.

4 LEMO Proposed Data Model

Our proposed data model is based on DCMI standards. DCMI was chosen to be the base schema for what we named Linked Educational Medical Objects (LEMO) Application Profile because of its wide usage in different repositories. DCMI is a simple schema with flat structure of elements which is beneficial for introducing further refined elements to the original set of elements.The LEMO AP adopts most of DCMI elements in its schema though changing the structure of some elements based on the conceptual model, shown in the framework presented in the previous section. The changes on DCMI standards are only refinements on its elements without adding new elements. The refinements are applied to specific elements when possible to add extra data in order to build dynamic connections between the educational objects based on the content of their metadata records. Figure 2 represent the schema of the proposed model with DCMI labels to specify the original elements derived from DCMI standard.

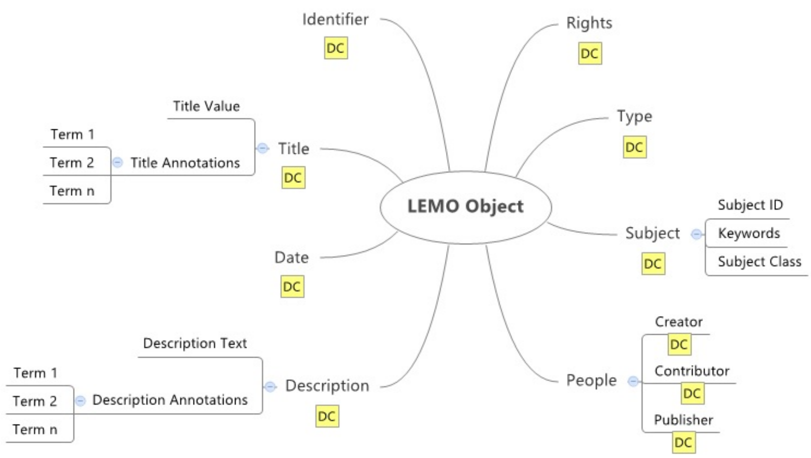

Fig. 2. Proposed LEMO Data Model with DCMI elements labeled

LEMO AP is implemented as Linked Data using RDF/XML which helps in exposing and connecting the educational objects when enriched with existing ontologies and Linked Open Data sets. Technical implementation of LEMO Data Model is detailed in the next section.

5 Harvesting, Mapping and Enriching Educational Objects

The term Linked Data is a new method for publishing data on the web using URIs and RDF. The data published in Linked Data is machine readable and can be linked to external data sets in a simple way [16]. In this data model, every educational medical object's metadata is represented using URIs and RDF which make the whole data set a Linked Open Data set of medical educational objects. The goal is to harvest and map educational objects from various websites and libraries into one Linked Open Data set.

Sample of a Record Presented in LEMO AP

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
```

```
        xmlns:dc="http://purl.org/dc/elements/1.1/"
        xmlns:oai="http://www.openarchives.org/OAI/2.0/"
        xmlns:lemo="http://www.warwick.ac.uk/ias/lemo#"
        xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
        xmlns:cc="http://creativecommons.org/ns#"
        xmlns:foaf="http://xmlns.com/foaf/0.1#">
 <rdf:Description rdf:about="http://www.warwick.ac.uk/ias/lemo#1234">
        <dc:identifier>a</dc:identifier>
        <dc:creator>
            <foaf:name> Author Name </foaf:name>
        </dc:creator>
        <dc:title>
            <rdf:Description>
                <rdf:value>title text</rdf:value>
                <lemo:textAnnotation rdf:parseType="Resource">
                    <lemo:term rdf:parseType="Resource">
                        <lemo:ID>term id</lemo:ID>
                        <lemo:from>1</lemo:from>
                        <lemo:to>2</lemo:to>
                        <lemo:text>word</lemo:text>
                        <lemo:class>class</lemo:class>
                    </lemo:term>
                </lemo:textAnnotation>
            </rdf:Description>
        </dc:title>
</rdf:Description>
</rdf:RDF>
```


Harvesting and Mapping EMOs In order to implement and test the AP proposed, we harvest Educational Medical Objects from repositories. OAI-PMH harvesting protocol is one of the protocols which can be used to harvest of educational materials from such repositories if the repository supports. This protocol supports harvesting content in different formats based on the library or repository support of this protocol. The harvested records are mapped into LEMO AP using XSL Transformation (XSLT) where XSL stands for Extensible Stylesheet Language and is a style sheet for XML documents. Harvesting other types of content such as blogs and videos is implemented in LEMO. RSS feeds reader is implemented to harvest RSS feeds from YouTube Channels or Blogs.

Enriching the content of LEMO AP After mapping the records harvested into LEMO AP, further enrichments are applied to textual elements of the records. The enrichment are basically implemented on two fields *title*, and *description* elements by annotating the raw text using medical ontologies and adding the annotated terms to the refined elements of the AP in *dc:title* and *dc:description* elements. Annotating textual content is done using SNOMED ontology. It is one of the most widely used ontologies for categorizing medical libraries content.

6 Experiments and Results

In this section, we present the experiments conducted to test the validity of the proposed framework used for developing LEMO Application Profile. The experiments were conducted on three different repositories for medical educational content. The repositories chosen varies in their content types in order to validate the data model against the goals determined in the framework.

The experiments were conducted on three different repositories hosting different types of medical educational materials. The repositories participating in the experiments were chosen based on the results of a survey conducted with Warwick Medical School educators. The results showed that most of the educators use popular search engines to find educational resources. They also mentioned using PubMed Central Library and medical journals when searching for educational content to help them in their teaching process. Based on that we have chosen *PubMed Central library*[5] and two other Journals hosting blogs and videos. The video collection is hosted by *The New England Journal of Medicine* (NEJM) on their YouTube channel[6], and the blogs are provided by *The American Journal of Medicine* (AMJMED)[7].

At the first stage of the experiments we harvested medical educational objects from the three repositories. The educational objects were harvested from PubMed Central using OAI-PMH harvesting protocol. The library provides OAI-PMH service which allows us to access and download metadata records of the PubMed Central library content. The service supports different data formats, and DCMI schema is supported by PubMed Central Library. We implemented OAI-PMH service in our application for harvesting recent records published in PubMed Central library and retrieved more than 10,000 records representing educational materials from different subjects. The data is retrieved in Dublin Core format and stored as XML files. Using RSS feeds for reading videos and blogs, we added an RSS Reader feature to the application developed in order to read RSS feeds of a given RSS feeds URL. Using this feature we were able to collect 208 videos posted by NEJM channel on YouTube, and a sample of 10 blogs collected using the same RSS Reader from AMJMED blogs.

The second stage of the experiments was concerned with mapping the harvested files into the proposed data model. All the content harvested was stored as XML files representing educational objects in different metadata formats. The first step was to build XSL files for transforming the content harvested to the proposed data model format. The data model is presented as RDF/XML file in order to enable the use of Linked Data for exposing and connecting the educational objects together. The benefits of using XSL are: converting the XML files into RDF/XML files, transforming the existing data format of the XML files to the desired LEMO AP proposed in section 4, and enabling the addition of new refined elements needed for data enrichment. Figure 3 is a snippet of a metadata record for an item retrieved from NEJM video collection in its XML format retrieved by the RSS Reader. All metadata records are mapped to LEMO AP using XSL Transformation services developed in the application. Figure 4 shows snippet of the same metadata record in RDF/XML format after the mapping process.

[5] http://www.ncbi.nlm.nih.gov/pmc/
[6] https://www.youtube.com/user/NEJMvideo
[7] http://amjmed.org/

```xml
<item>
    <title>Leg Venographic Images Obtained before and after Thrombolysis</title>
    <link>http://www.youtube.com/watch?v=a5iuJQ0XFxA&amp;feature=youtube_gdata</link>
    <atom:updated xmlns:atom="http://www.w3.org/2005/Atom">2014-05-06T08:44:24.000Z</atom:updated>
    <gd:comments>
    <media:group xmlns:media="http://search.yahoo.com/mrss/">
        <media:category label="Science &amp; Technology" scheme="http://gdata.youtube.com/schemas/2007/ca
        <media:content url="http://www.youtube.com/v/a5iuJQ0XFxA?version=3&amp;f=videos&amp;app=youtube_g
        <media:description type="plain">In the first video, a cine loop of contrast-material injection be
        <media:keywords />
        <media:title type="plain">Leg Venographic Images Obtained before and after Thrombolysis</media:ti
        <yt:duration xmlns:yt="http://gdata.youtube.com/schemas/2007" seconds="28" />
    </media:group>
    <gd:rating xmlns:gd="http://schemas.google.com/g/2005" average="5.0" max="5" min="1" numRaters="2" re
    <yt:statistics xmlns:yt="http://gdata.youtube.com/schemas/2007" favoriteCount="0" viewCount="105" />
    <description>In the first video, a cine loop of contrast-material injection before thrombolysis shows
    <category domain="http://schemas.google.com/g/2005#kind">http://gdata.youtube.com/schemas/2007#video<
    <category domain="http://gdata.youtube.com/schemas/2007/categories.cat">Tech</category>
    <pubDate>Mon, 05 May 2014 14:17:15 GMT</pubDate>
    <guid isPermaLink="false">http://gdata.youtube.com/feeds/api/videos/a5iuJQ0XFxA</guid>
    <dc:creator>NEJMvideo</dc:creator>
    <dc:date>2014-05-05T14:17:15Z</dc:date>
</item>
```


Fig. 3. XML Snippet for RSS feeds object from NEJM YouTube Channel

As shown in figure 4, the metadata record is represented in RDF format where the record metadata are represented as triples (subject, predicate, object) i.e. the subject is the video resource identified by *rdf:about*, and has predicate such as *dc:title* and the object is another resource representing the title as seen in the picture. The flexible nature of RDF/XML format make it easier for us to dynamically add elements to the records content for enriching it.


```xml
<rdf:Description rdf:about="http://gdata.youtube.com/feeds/api/videos/a5iuJQ0XFxA">
    <dc:identifier>http://www.youtube.com/watch?v=a5iuJQ0XFxA&amp;feature=youtube_gdata</dc:identifier>
    <dc:title rdf:resource="http://gdata.youtube.com/feeds/api/videos/a5iuJQ0XFxA:title"/>
    <dc:description rdf:resource="http://gdata.youtube.com/feeds/api/videos/a5iuJQ0XFxA:desc"/>
    <dc:date>2014-05-05T14:17:15Z</dc:date>
    <dc:creator>NEJMvideo</dc:creator>
    <dc:type rdfs:type="Video">28 seconds</dc:type>
</rdf:Description>
<rdf:Description rdf:about="http://gdata.youtube.com/feeds/api/videos/a5iuJQ0XFxA:title">
    <rdf:value>Leg Venographic Images Obtained before and after Thrombolysis</rdf:value>
</rdf:Description>
<rdf:Description rdf:about="http://gdata.youtube.com/feeds/api/videos/a5iuJQ0XFxA:desc">
    <rdf:value>In the first video, a cine loop of contrast-material injection before thrombolysis shows
</rdf:Description>
```


Fig. 4. Snippet for the Same Record in LEMO AP RDF/XML style

The third stage of the experiments was applied for annotating textual elements in LEMP AP records which are the *Title* and *Description* fields. For the annotation service we used a popular and widely used ontology in the medical field called SNOMED ontology. The annotation service was implemented in our application using BioOntology annotation API[8] in addition to the services

[8] http://bioportal.bioontology.org/annotator

developed for harvesting and mapping the educational objects. The results of annotating the textual fields of a specific record were dynamically added to the records RDF/XML representation. The *Title* resource of any object after mapping contains only the textual value of a title without further enrichments. Figure 5 represents a sample of the *Title* resource after annotating its textual value. Predicates and their objects were dynamically added to the *Title* resource with details about the annotated term and its annotated class in SNOMED ontology which contains that term.


```xml
<rdf:Description rdf:about="oai:pubmedcentral.nih.gov:2606822:title">
    <rdf:value>Dipeptidyl Peptidase-4 Inhibition by Vildagliptin and the Effect on
        Insulin Secretion and Action in Response to Meal Ingestion in Type 2 Diabetes </rdf:value>
    <lemo:lemoTitleAnnotation rdf:resource="oai:pubmedcentral.nih.gov:2606822:title:term1"/>
    <lemo:lemoTitleAnnotation rdf:resource="oai:pubmedcentral.nih.gov:2606822:title:term2"/>
    <lemo:lemoTitleAnnotation rdf:resource="oai:pubmedcentral.nih.gov:2606822:title:term3"/>
    <lemo:lemoTitleAnnotation rdf:resource="oai:pubmedcentral.nih.gov:2606822:title:term4"/>
    <lemo:lemoTitleAnnotation rdf:resource="oai:pubmedcentral.nih.gov:2606822:title:term5"/>

<rdf:Description rdf:about="oai:pubmedcentral.nih.gov:2606822:title:term1">
    <lemo:lemoClassLabel>Hydrolase</lemo:lemoClassLabel>
    <lemo:lemoTermClass>http://purl.bioontology.org/ontology/SNOMEDCT/74628008</lemo:lemoTermClass>
    <lemo:lemoTermText>PEPTIDASE</lemo:lemoTermText>
    <lemo:lemoTo>20</lemo:lemoTo>
    <lemo:lemoFrom>12</lemo:lemoFrom>
    <lemo:lemoTermID>http://purl.bioontology.org/ontology/SNOMEDCT/130202003</lemo:lemoTermID>
</rdf:Description>
<rdf:Description rdf:about="oai:pubmedcentral.nih.gov:2606822:title:term3">
    <lemo:lemoClassLabel>General metabolic function</lemo:lemoClassLabel>
    <lemo:lemoTermClass>http://purl.bioontology.org/ontology/SNOMEDCT/47722004</lemo:lemoTermClass>
    <lemo:lemoTermText>INHIBITION</lemo:lemoTermText>
    <lemo:lemoTo>33</lemo:lemoTo>
    <lemo:lemoFrom>24</lemo:lemoFrom>
    <lemo:lemoTermID>http://purl.bioontology.org/ontology/SNOMEDCT/61511001</lemo:lemoTermID>
</rdf:Description>
<rdf:Description rdf:about="oai:pubmedcentral.nih.gov:2606822:title:term5">
    <lemo:lemoClassLabel>Dipeptidyl peptidase IV inhibitor</lemo:lemoClassLabel>
    <lemo:lemoTermClass>http://purl.bioontology.org/ontology/SNOMEDCT/422403005</lemo:lemoTermClass>
    <lemo:lemoTermText>VILDAGLIPTIN</lemo:lemoTermText>
    <lemo:lemoTo>46</lemo:lemoTo>
    <lemo:lemoFrom>35</lemo:lemoFrom>
    <lemo:lemoTermID>http://purl.bioontology.org/ontology/SNOMEDCT/428807006</lemo:lemoTermID>
</rdf:Description>
```


Fig. 5. A sampe of an Object's Title Resource after annotation

7 Conclusions and Future Work

Searching for educational content on the web is a challenge. In this paper, we proposed a data model that is part of a work in progress project for harvesting and exposing Medical educational objects for facilitating the process of search and finding of educational objects. We have presented the framework for developing and implementing the proposed data model called LEMO AP and conducted experiments for applying this AP on different repositories. The repositories used in the experiments contain different types of content such as articles, video, and blogs. The proposed data model was able to accommodate different metadata schemas harvested from these repositories, and annotations were successfully added to the records harvested. Based on the proposed data model, we are aiming at building a metadata repository for harvesting and querying content from repositories and Web2.0 technologies such as RSS Feeds. In the near future, we will study the possibility of inferring categories for the objects retrieved using the annotation terms discovered from SNOMED ontology. Further work will focus on dynamic linking of objects and evaluating the accuracy of these linkages discovered.

References

1. Caswell, T., Henson, S., Jensen, M., Wiley, D.: Open content and open educational resources: Enabling universal education, The International Review of Research in Open and Distance Learning 9(1), (2008)
2. Fleiszer, D. M., Nancy H. P., Sean P. S.: New directions in medical e-curricula and the use of digital repositories. Academic Medicine 79(3), 229-235 (2004)
3. Popoiu, M. C., Grosseck, G., Holotescu, C.: What do we know about the use of social media in medical education?. Procedia-Social and Behavioral Sciences 46, 2262-2266 (2012)
4. Franks, P., and Kunde, N.: Why METADATA Matters. Information Management Journal 40(5), 55-61 (2006)
5. Devedic, V.: Semantic web and education. Vol. 12. Springer, (2006)
6. Park, J. R., Tosaka, Y.: Metadata creation practices in digital repositories and collections: Schemata, selection criteria, and interoperability. Information Technology and Libraries 29(3), 104-116 (2013)
7. Reba, I., de La Passardire, B., Labat, J. M.: A Formalism to Describe Open Standards in order to Generate Application Profiles. In Proceedings of the International Conference on Computer in Education, pp. 491-498, (2008)
8. Learning Technology Standards Committee of the IEEE: Final 1484.12.1-2002 LOM Draft Standard. http://ltsc.ieee.org/wg12/20020612-Final-LOM-Draft.html, (2005)
9. Learning Technology Standards Committee of the IEEE: Draft Standard for Learning. http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf, (2002)
10. Information and documentation: The Dublin Core metadata element set, http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumbe=r52142, (2009)
11. Powell, A., Nilsson, M., Naeve, A., Johnston, P., Baker, T.: DCMI abstract model. DCMI Recommendation, http://dublincore.org/documents/abstract-model, (2007)
12. Healthcare LOM, ANSI/MEDBIQ LO. 10.1-2008, Healthcare Learning Object Metadata, http://txlor.org/docs/medbiquitoushealthcarelomspecification-200607.pdf, (2011)
13. Bamidis, P. D., Kaldoudi, E., Pattichis, C.: mEducator: a best practice network for repurposing and sharing medical educational multi-type content. Leveraging Knowledge for Innovation in Collaborative Networks, 769-776, (2009)
14. Candler, C. S., Uijtdehaage, S. H. J., Dennis, S. E.: Introducing HEAL: The health education assets library, Academic medicine, 78(3) 249-253 (2003)

15. National Library of Medicine, Fact Sheet: Medical Subject Headings(MeSH), http://www.nlm.nih.gov/pubs/factsheets/mesh.html (2013)
16. Bizer, C., Heath, T., Berners-Lee, T.: Linked Data - the story so far. International Journal on Semantic Web and Information Systems 5(3), 1-22 (2009)