Using Linked Data to Annotate And Search Research Articles For Supporting Research Community

Dr. R.RajaRence [1], Dr.T.Nalini[2]
[1]Manager, Department of Computer Education, Apollo computer education, Chennai-73,TN, India,
[2]Professor Department of Computer Science&engineering Bharath University,Chennai-73,TN,India,

Abstract *: Multimedia educational resources play an important role in reasearch, particularly for distance learning environments. With the rapid growth of the multimedia web, large numbers of research articles video resources are increasingly being created by several different organizations. It is crucial to explore, share, reuse, and link these educational resources for better e-learning experiences. Most of the video resources are currently annotated in an isolated way, which means that they lack semantic connections. Thus, providing the facilities for annotating these video resources is highly demanded. These facilities create the semantic connections among video resources and allow their metadata to be understood globally. Adopting Linked Data technology, this paper introduces a video annotation and browser platform with two online tools: Annomation and SugarTube. Annomation enables users to semantically annotate video resources using vocabularies defined in the Linked Data cloud. SugarTube allows users to browse semantically linked educational video resources with enhanced web information from different online resources. In the prototype development, the platform uses existing video resources for research articles. The result of the initial development demonstrates the benefits of applying Linked Data technology in the aspects of reusability, scalability, and extensibility.*

Keywords : *Video resources, explore, share, reuse, linked data, semantic search, web services.*

I. INTRODUCTION

This paper adopts Semantic Web technology, more precisely, the Linked Data approach to address the above challenges. The following lists the major contributions of our approach.

1. A video annotation ontology is designed by following Linked Data principles and reusing existing ontologies. It provides the foundation for annotating videos based on both time instance and duration in the video streams. This allows more precise description details to be added to the video.

2. A semantic video annotation tool (Annomation) is implemented for annotating and publishing research articles video resources based on the video annotation ontology. Annomation allows annotators to use domain specific vocabularies from the Linked Open Data cloud to describe the video resources. These annotations link the video resources to other web resources.

3. A semantic-based video searching browser (Sugar-Tube) is provided for searching videos. It generates links to further videos and research articles video resources from the Linked Open Data cloud and the web.

II. PROBLEM STATEMENT:

Video resources should be described precisely. It is difficult to use only one general description to accurately tell the whole story of a video because one section of the video stream may have plenty of information (e.g., on historical figures and hidden events in the conversations) but some of them might not related to the main points of the video when it was created. Therefore, the normal paragraph-based description process is not good enough for annotating videos precisely. A more accurate description mechanism, based on the timeline of the video stream, is required.

The descriptions of the educational resources should be accurate and machine-understandable, to support related search functionality. Although a unified and controlled terminology can provide accurate and machine- understandable vocabularies, it is impossible to build such a unified terminology to satisfy different description requirements for different domains in practice.

Linking video resources to useful knowledge data from the web. More and more knowledge and scientific data is published on the web by different research and educational organizations (e.g., Linked Open

Data), and so it is useful to break the teaching resource boundaries between closed institutions and the Internet environment to provide richer learning materials to both educators and learners.

III. BACKGROUND AND RELATED WORK

3.1 Requirements for enhancing research articles Video Resources to support research community

Videos are important educational resources that enable students to gain knowledge more efficiently and intuitively than text-based educational resources. Video resources play an important role in distance learning courses (e.g., history courses). For example, a five-minute long video of a speech may contain plenty of information such as event background,
location, time and related people. However, traditional research articles video resources usually lack labelled vocabularies and structured metadata. These drawbacks limit the usability, efficiency, and reusability of the research article video resources.

3.2 Semantic Web, Linked Data, and Web Services

The Semantic Web the Semantic Web [5] is an evolving development of the World Wide Web, in which the meanings of information on the web is defined; therefore, it is possible for machines to process it. The basic idea of Semantic Web is to use ontological concepts and vocabularies to accurately describe contents in a machine readable way. These concepts and vocabularies can then be shared and retrieved on the web.

3.3 Linked Data

Linked Data is the recent revolutionary development of the Semantic Web. Linked Data create typed links between different data from different resources. From the technical point of view, Linked Data means to publish data on the web in such a way that they are readable by machines and their meanings are explicitly expressed. These data are then linked to external data sets, and in turn are linked from external data sets

3.4 Multimodal Fusion for Video Search Reranking

Videos are traditionally searched by syntactic matching mechanisms. Recently, with more videos being annotated or tagged in the Linked Data manner, researchers have begun to search videos in a more Semantic-Web oriented fashion. The two major approaches are the
semantic indexing process and the natural language analysis process.

3.5.Semantic Video Search Using Natural Language Queries

The indexing process assumes that the video annotations are made from a fixed set of vocabularies that change infrequently. Although this process can be efficient, the fixed set of vocabulary may introduce a gap between user's knowledge and indexed annotations, especially in the research environment, in which videos are often annotated by different groups of teachers or students, who may apply different annotation terms to the same video in the context of different courses and key points.

3.6. Semantic Web for Content Based Video Retrieval

The natural language analysis process focuses more on adding semantic tags to the user's search inputs. However, most of these approaches require machine learning mechanisms to assist dynamically adding tags. Hence, they restrict their applications to small and closed domains of discourse.

Advantages:

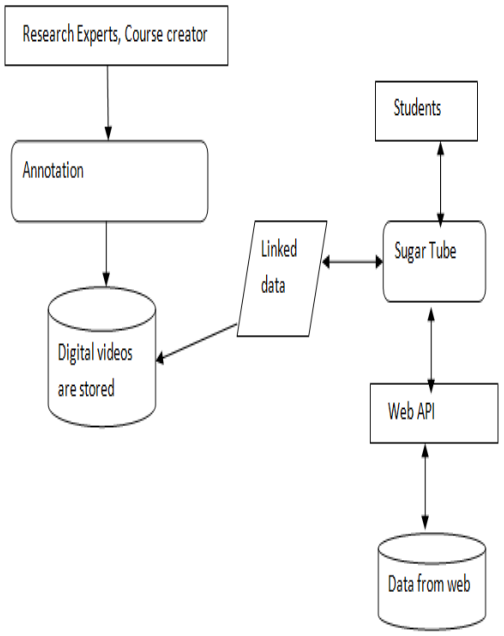
Four most important advantages of using Linked Data to create video annotations for the research domain.

- Each video annotation is unique and explicitly identified
- Information is linked to the big knowledge net
- Videos are linked to each other
- More useful knowledge is gathered from the web

Benefit of society:

Distance learning education.
Research development.

Process flow diagram



IV. USING LINKED DATA TO ANNOTATE VIDEOS

Traditional video annotations using free-text keywords or predefined vocabularies are insufficient for a collaborative and multilingual environment. They do not properly handle the annotation issues, such as accuracy, disambiguation, completeness, and multilinguality. For example, free-text keywords annotation easily fails on accuracy issues as they may contain spelling errors or be ambiguous. Furthermore, they are insufficient for a collaborative and multilingual environment. Our approach uses Linked Data to tackle the above issues in video annotations. It brings the following benefits. Each vocabulary is controlled and accurately defined in the Linked Data Cloud. It owns a unique URI to distinguish it from other vocabularies, so there are no conflicts between different vocabularies and meanings.

Different vocabularies, which describe the same thing, are linked using the owl: same As property as an equation definition. Meanwhile, a number of semantic annotations are used to build the relationships between different vocabularies, such as rdfs: subClassOf and rdfs: seeAlso. Once a vocabulary is applied to an annotation, the related vocabularies are associated with the annotation. Therefore, the collaborative and multilingual issues are well addressed. The Linked Open Data Cloud, which has the most complete data sets to describe the current world, helps to find a good number of related educational resources. Five Linked Data Services are currently used as a foundation to annotate videos, and they are embedded in the Annomation functions to facilitate the annotation process. More services can be easily added into the system by adding a tab option to show the query results of the new service when required. The five Linked Data Services are: Dewey Decimal: The top level Dewey Decimal Classification (covering the first three digits of a Dewey number) has been published in RDF form by the Online Computer Library Center,4 and the resulting taxonomy is presented to the annotator as a browsable tree.

Library of Congress classifications: The Library of Congress has published its entire classification system in RDF,5 but this is much too large to present directly to the user. Instead, Annomation provides an interface to the Library of Congress keyword search service, which returns suitable RDF files for the user to choose from.

GeoNames [23]: The GeoNames API is used to identify named locations using a keyword search, or to perform reverse lookup to find named locations in a vicinity. The results provide the position and the category information with URI indentifiers.

OU Bluelist: An OU service is used to get Open University course taxonomies.

Zemanta [24]: A service which provides analysis of natural language text to identify various concepts and named entities returning their URIs to Linked Data such as DBpedia and Freebase.

4.1 Annomation Implemetation

The Annomation interface is divided into four sections: a Flash video player (top left); a list of current annotations (top right); controls for the video player, and for entering new annotations (across the centre) and a set of panels to help the user to find new Linked Data URLs (bottom). The bottom panels provide quick access to previously used tags, to the Dewey and Library of Congress classifications schemes, Open University course taxonomies, a service for suggesting URLs based on the Zemanta service, and a visual map tool that uses GeoNames to find named geographical entities.

The system is implemented in the Clojure language [25], using Sesame as an RDF quad store, and RDF2Go6 as an abstraction over the store. Annomation provides programmatic APIs in the form of a SPARQL end point for querying and RESTful interfaces for adding and removing annotations, and exploring existing ones. The client side uses Javascript with the Yahoo YUI library7and the JQuery plugin RDF Query,8 and FlowPlayer9 for video playback. OpenId [26] is used for user identification and authentication.

4.2. SUGARTUBE: SEMANTICS USED TO GET ANNOTATED VIDEO RECORDING

SugarTube is developed to facilitate the usage of the OU's educational video resources that are annotated by Annomation. It adopts the Semantic Web approach to search videos and explore their related online resources in a mashup navigation interface. In SugarTube, the annotations are semantically matched to other annotated educational resources from the web. The SugarTube application includes three layers. Users interact with the application layer when specifying the concepts, documents (e.g., lecture notes), or website contents in order to get educational video resources. Based on different types of concept data, user requests are then sent to the semantic data mining and reasoning layer for generating different queries to the service layer. The service layer includes both Linked Data Services and nonsemantic based services.

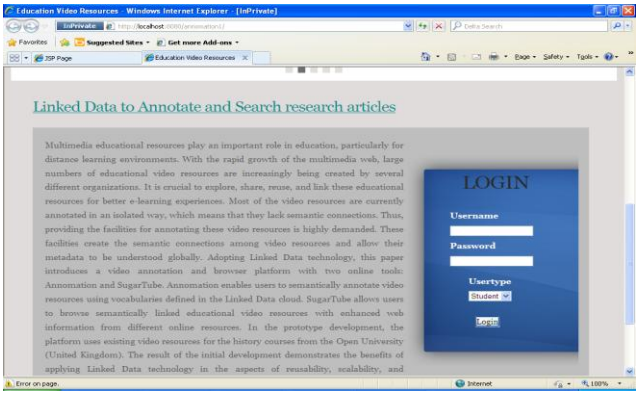
4.3.Service Layer

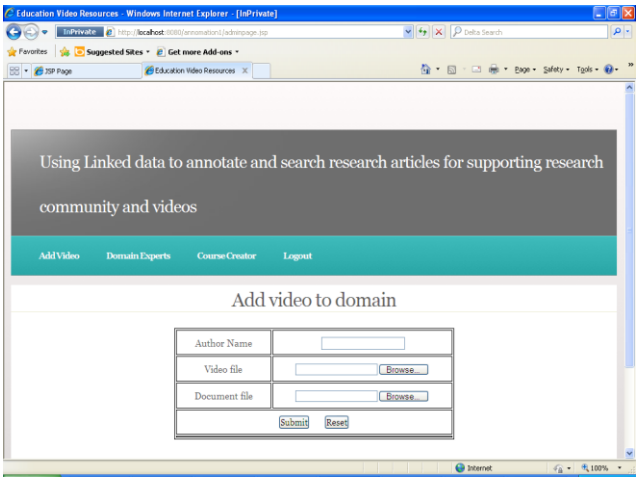
In addition to the Linked Data Services that are applied in the Annomation process, some other Linked Data Services and nonsemantic services are used in SugarTube. The Linked Data Services are:
. WorldHistory11 that provides API access to retrieve the information about people, events, places, and genealogy in history;. The OU Linked Data12 that is currently under development and aims to extract and interlink previously available educational resources in various disconnected institutional repositories of the Open University and publish them into the Linked Open Data cloud; and . Sindice [27], [28] that is a semantic search engine, which crawls and collates the Semantic Web (including microformats), and provides services such as keyword-based searching for linked data and accessing cached fragments of the Semantic Web. In order to reuse the data sets and services listed above, we use two ontologies: the Simple Knowledge Organization System (SKOS) [29] and the WGS84-based RDF vocabulary for geographical data [30].

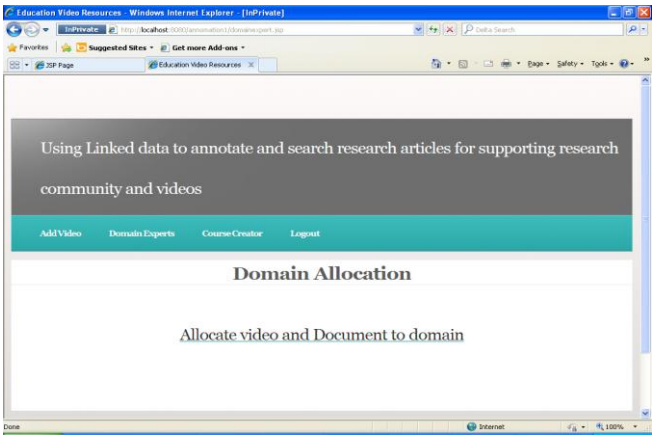
The nonsemantic services are:
BBC Web API13 that offers machine readable program information;
Map Services that provides interactive Ajax mapping services from Google14and Yahoo;15
OpenLearn [31] that gives free access to a subset of the Open University's course resources, along with metadata; and
Experiment & Result
YouTube data services16 that finds videos via a keyword-based search.

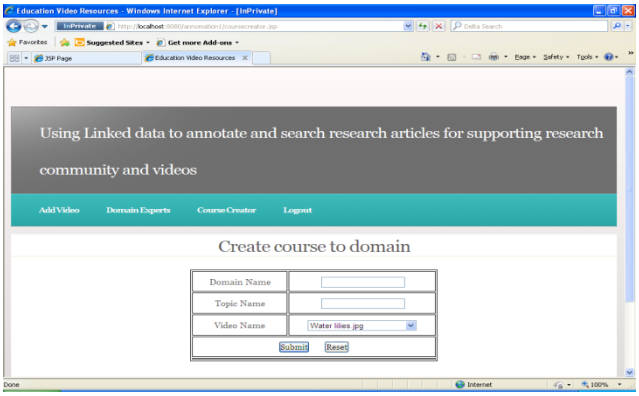
4.4 Application Layer

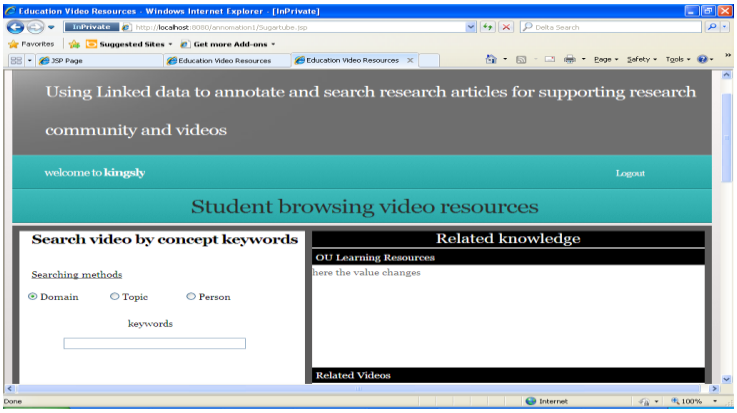
The SugarTube functionalities are divided into two groups, namely basic concept search and advanced search. The basic concept search divides the concepts into "Person," "Event," "Place," and "Others." For different types of concepts, different service queries are generated. For instance, searching by the name of a person queries the searchByPerson WorldHistory service, while searching by the name of a place queries the searchByName GeoName service. The advanced search supports searching videos by automatically analyzing documents, highlighting web contents, and pointing to locations on a map. Behind this, the Zemanta17 service is used. For example, when a user copies and pastes the learning content from lecture notes into the text field, all related knowledge

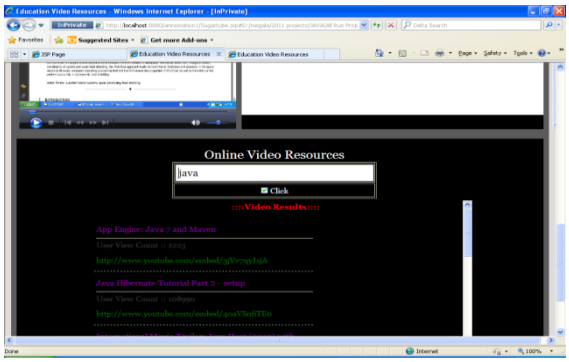

Linked Data to Annotate and Search research articles
LOGIN
Username
Password
Usertype
Student
Login


Using Linked data to annotate and search research articles for supporting research community and videos
Add Video
Domain Experts
Course Creator
Logout
Add video to domain
Author Name
Video file
Browse...
Document file
Browse...
Submit
Reset


Using Linked data to annotate and search research articles for supporting research community and videos
Add Video
Domain Experts
Course Creator
Logout
Domain Allocation
Allocate video and Document to domain


Using Linked data to annotate and search research articles for supporting research community and videos
Add Video
Domain Experts
Course Creator
Logout
Create course to domain
Domain Name
Topic Name
Video Name
Water lilies.jpg
Submit
Reset





is deployed for gathering the geoinformation about a place so that the user may click on the map to search related videos. The searching results do not only contain the OU educational video resources with their annotations but also include relevant learning resources about the videos and related videos from other services.

4.5 SugarTube Evaluation Results

Whether students think the SugarTube can help their studies is the most interesting part of our evaluation task for SEG. The SugarTube evaluation task contains:

[1] Using the basic search functions to find as many as possible (no more related new video can be found) videos that related to "Cold war" topics and were stored in OU video repository. We examined the quantity of the found videos.

[2] Using the place search function and enter "Cape Canaveral" to search for related videos. Go through all video resources from different video search providers to identify at least five videos that relate to the "Moon landing" or "Apollo 11" topics.

[3] Using the person search function and enter any person's name that you believe is related to "Cold War" class and identify at least two videos from different resources and two URIs that describe either the person you searched for or "Cold war" related topics.

[4] Using the map search function to search for videos and information related to "Berlin" in "Cold War" topic (at least five related resources need to be identified).

[5] Taking a particular text content from a lecture note that is used for the "Cold war" class to search the related and useful resources to prepare the class based on the highlights in the lecture note.concepts are listed, which enables the user to select further video searching activities. The Google map service

V. CONCLUSION

This paper illustrated the Annomation and the SugarTube and no student thought platform that uses Linked Data technologies to semantically.annotate and search researchl video resources from the Open University video repository and link the videos to all other research resources on the web. In the semantic annotation process, 1) an annotation t of ontology is defined to support Linked Data annotations. 2) dynamic annotation URI suggestions are fully supported by integrating Linked Data Services into the Annomation interface; and 3) collaborative functionalities are implemented to enhance the teamwork capability.. In the semantic search process, the search methods are based on the data retrieved through Linked Data Services and URIs, which links different resources together to enrich the original video search results. SugarTube shows that e-learning resources distributed across different research organizations can be linked together to provide more value-added information.

REFERENCES

[1] Hong Qing Yu, Carlos Pedrinaci, Stefan Dietze, and John Domingue
[2] E. Allen and J. Seaman, "Class Differences Online Education in the United States," http://sloanconsortium.org/publications, 2010.
[3] J.W. Brackett, "Satellite-Based Distance Learning Using Digital Video and the Internet," IEEE Multimedia, vol. 5, no. 3, pp. 72-76, July-Sept. 1998.
[4] T.-D. Wu, Y.-Y. Yeh, and Y.-M. Chou, "Video Learning Object Extraction and Standardized Metadata," Proc. Int'l Conf. Computer Science and Software Eng., vol. 6, pp. 332-335, 2008.
[5] M. Hausenblas and M. Karnstedt, "Understanding Linked Open Data as a Web-Scale Database," Proc. Second Int'l Conf. Advances in Databases Knowledge and Data Applications (DBKDA), pp. 56-61, Apr. 2010.
[6] T. Berners-Lee, J. Hendler, and O. Lassila, "The Semantic Web," Scientific Am. Magazine, 2001.
[7] The Description Logic Handbook: Theory, Implementation, and Applications, F. Baader, D. Calvanese, D.L. McGuinness, D. Nardi Univ. of Denmark, master's thesis, 2010. and P.F. Patel-Schneider, eds. Cambridge Univ., 2003.
[8] T. Berners-Lee, "Linked Data," http://www.w3.org/Design Issues/LinkedData.html, 2006.
[9] C. Bizer, T. Heath, and T. Berners-Lee, "Linked Data— The Story So Far," Int'l J. Semantic Web and Information Systems, vol. pp. 1- 22, 2009.
[10] E. Prud'hommeaux and A. Seaborne, "SPARQL Query Language for RDF," technical report, World Wide Web Consortium, Jan. University.2008.