

# Linked Data Based Video Annotation and Browsing for Distance Learning

Dave Lambert and Hong Qing Yu

Knowledge Media Institute  
The Open University  
Milton Keynes, United Kingdom

**Abstract.** We present a pair of prototype tools that enable users to mark up video with annotations and later explore related materials using Semantic Web and Linked Data approaches. The first tool helps academics preparing Open University course materials to mark up videos with information about the subject matter and audio-visual content. The second tool enables users, such as students or academics, to find video and other materials relevant to their study.

## 1 Introduction

The Open University (OU) is a leading distance learning institute, and the largest provider of tertiary education in the United Kingdom. It makes innovative use of multimedia resources in its teaching materials, and is famous for its television programmes, both those designed for its students, and those targeted at the general population. The OU course programme transcripts alone occupy 92 metres of shelving in the university's library, and the associated video material represents a substantial investment and resource, and the university is considering how to better exploit it for:

- academics preparing new courses, who can survey a larger amount of material more quickly and effectively
- students who wish to follow a topic of interest beyond the prescribed materials
- serendipitous use by anyone

We have built two tools to investigate how using Semantic Web and Linked Open Data (LOD) principles [1] could aid search and reuse. We cover the background and motivation behind the Open University's development of such tools (Section 2), then introduce the annotation tool 'Annomation' and ontology (Section 3), the browser software 'SugarTube' (Section 4), and the Linked Data sets and services they use (Section 5). We compare our work with related systems and approaches (Section 6) and conclude (Section 7).

## 2 Background

The Open University is one of the world's largest providers of tertiary education: it has over 200,000 students currently enrolled on courses at bachelors, masters, and

PhD level. Since its founding in 1969, the OU has pioneered innovative teaching methods. Most famously, it has produced a wide range of television programmes, often in collaboration with the British Broadcasting Corporation (BBC), both for its own students and the general population. Today, the OU continues to use new multimedia technologies, having moved from television through CD, DVD and now online delivery mechanisms.<sup>1</sup> The OU has also developed online teaching facilities like course forums, multi-participant audio environments for language learning, and is now moving to make greater use of the Semantic Web.

Although supported by local, face-to-face tutorial sessions, Open University courses are primarily taught at a distance and from the beginning they have used multimedia content to engage and educate students. Staff preparing an OU course will work with BBC archivists to find appropriate video material to illustrate certain points of the course. This is labour intensive: a text database of available footage (principally from the OU and BBC archives) is searched to identify dozens of hours of video which may be of interest. The course team will then view this video, manually noting sections of interest. When the final set of video segments are identified for the course (typically an hour or two of video will result), the information on unused material is lost. The OU and the BBC are currently making large investments in Semantic Web technologies, and the cataloguing of video material is an ideal candidate. We therefore developed prototype systems that enable OU course staff to mark up video in a Semantic Web or Linked Data style, and to subsequently navigate it and augment it with additional materials.

### 3 Annotation

Our first tool is Annomation,<sup>2</sup> which handles the input of annotations from users. Annomation provides a simple Web browser interface (see Figure 1) split into four main areas:

- a video player (top-left)
- a list of videos, and existing annotations for the current video (top-right)
- a set of controls for the video player, and input widgets to enter the annotations (across the centre)
- a set of panels to aid in finding suitable linked data annotations (at bottom)

In operation, the user first selects which of the videos to annotate. The videos of interest at the moment are transcoded from the original material by a BBC archivist at the request of the course team, and we serve them directly from the Annomation server, but the architecture supports using video served from anywhere on the Web. Having selected a video, the user will be presented with a timeline of any annotations previously made by herself and others: the user can skip to particular instants or durations noted in the annotations. The most basic

---

<sup>1</sup> <http://open.edu/itunes/>

<sup>2</sup> <http://annomation.open.ac.uk/annomation>

way to create an annotation is simply to pause the video at the appropriate point, enter a duration if appropriate, and add a Semantic Web/Linked Data URI. This is sent to the server, and the annotation recorded. Annotations can be optionally marked as applying to something visible in the video, audible in the soundtrack, or to the conceptual subject matter of the video. Since finding appropriate URIs is non-trivial, the fourth part of the interface is dedicated to helping the user find them. The details of the data sources are discussed in Section 5, but include the Dewey Decimal and Library of Congress classifications, geographic searches using interactive maps, suggestions based on natural language processing and entity extraction, and internal Open University classifications. The user can also peruse the URIs that have been previously used on other videos or by other users.

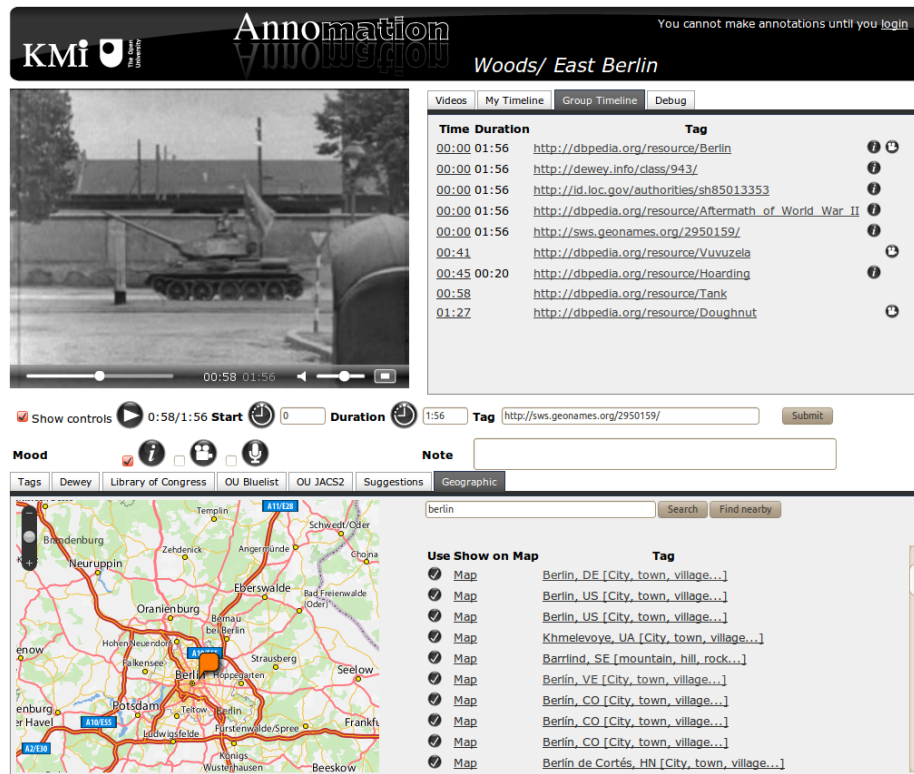
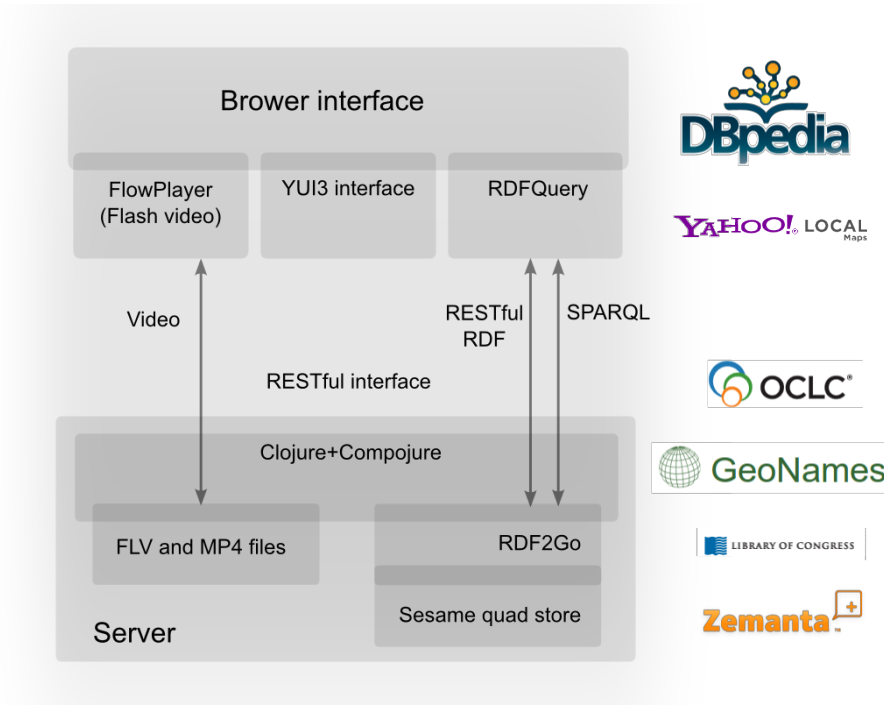


Fig. 1. The Annomation interface

We implemented the system (Figure 2) in the Clojure language, using Sesame as an RDF quad store, and RDF2Go 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. It is our intent that Annomation's RDF become part of the Linked Open

Data cloud. In the client we use Javascript with the Yahoo YUI library and the JQuery plugin RDFQuery, and FlowPlayer for video playback. We use OpenId for user identification and authentication.



**Fig. 2.** Annomation architecture

### 3.1 Data representation

At the heart of our approach is the Linked Data that is created and stored in Annomation, and searched by SugarTube. Users' annotations are recorded in an RDF quad store, with each user having their own context (a named graph) to facilitate privacy and provenance. The data sets related to user accounts and videos are also stored in the same store, in a management context. Most entities described in the Annomation store, including videos, users, and annotations, are assigned generated, globally unique URLs so they can be named as part of the Linked Open Data cloud. We reused several existing RDF Schema [2] vocabularies, and introduced a few new terms specifically for Annomation. The existing vocabularies we reused are:

- Friend-of-a-Friend (FOAF) [3], for identifying users and their accounts.

- The Timeline ontology [4], for identifying temporal instants and durations on the video timeline.
- Dublin Core, in its RDFS form,<sup>3</sup> for meta-data such as a video’s title, and the author and creation time of each annotation.

To tie together the data described using these vocabularies, we designed our own small ontology:

```
@prefix : <http://annomation.open.ac.uk/annomation/ns/annomation#> .
@prefix tl: <http://purl.org/NET/c4dm/timeline.owl#> .
@prefix owltime: <http://www.w3.org/TR/owl-time/#> .

:Video a rdfs:Class .

:Annotation a rdfs:Class .

:Mood a rdfs:Class .

:topic a :Mood ;
    rdfs:label "Topic of conversation"@en .
:inVideo a :Mood ;
    rdfs:label "In video stream"@en .
:inAudio a :Mood ;
    rdfs:label "In audio stream"@en .

:annotates a rdfs:Property ;
    rdfs:domain :Annotation ; rdfs:range owltime:TemporalEntity .

:reference a rdfs:Property ;
    rdfs:domain :Annotation ; rdfs:range rdfs:Resource .

:mood a rdfs:Property ;
    rdfs:domain :Annotation ; rdfs:range :Mood .
```

We allocate a unique URI to every video, user, time, and annotation. In this way, we can identify a video URI as a **Video**, reference that URI from a **TemporalEntity** in the timeline ontology, and then create an **Annotation** to tie together the time with its associated video, the ‘mood’ of the annotation and the annotation itself. As an illustration, the annotation about the tank seen in Figure 1) is:

```
:fa031dd6-89cb-49ec-b9d9-43e9f6ad7b65 a anno:Annotation ;
    anno:annotates :dd5e5b59-7f83-4b30-b1a4-15f16781d9a2 ;
    anno:reference <http://dbpedia.org/resource/Tank> ;
    anno:mood anno:topic , anno:inVideo , anno:inAudio ;
    dc:creator :ae8d5a6d-f3d3-4b68-be3d-47160a26d497 ;
    dc:date "2010-09-16T10:20:15Z"^^xsd:dateTime .

:dd5e5b59-7f83-4b30-b1a4-15f16781d9a2 a tl:Interval ;
    tl:onTimeline :c2b91ed5-ba09-48ff-97f6-1fe0e0bb9c12 ;
    tl:at "PT01M00.000S"^^xsd:duration ;
    tl:duration "PT10.000S"^^xsd:duration .
```

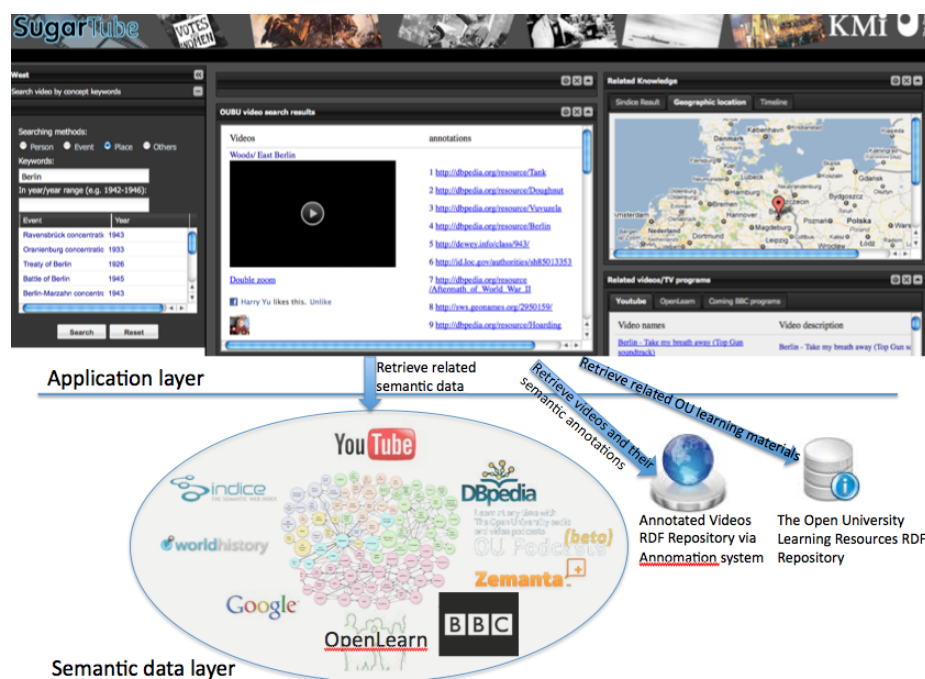
The `dc:creator` is the Annomation account of the user that made the annotation, and this is stored as FOAF and linked to the OpenId URI.

## 4 Browsing

To search the annotated videos, and explore related materials through Linked Data approaches, we developed the SugarTube browser (Semantics Used to Get

<sup>3</sup> <http://www.dublincore.org/schemas/rdfs/>

Annotated video Recording). The SugarTube interface and system architecture is illustrated in Figure 3. The SugarTube architecture has two layers: the application layer, responsible for presentation and interaction, and the semantic data layer which finds and integrates data that has been published in the LOD cloud by different organisations, including annotation data from Annomation. The Semantic Web data is obtained through RESTful services and SPARQL endpoints. The application layer allows the user to query and navigate the video search results using the semantic data (e.g. related knowledge, concepts, videos, and learning materials).



**Fig. 3.** SugarTube browser (top) and architecture (bottom)

SugarTube is designed to help the learner in navigating learning resources using Linked Data, so its first task is to connect the learner with appropriate concepts in the Semantic Web. It provides various search functions including those based on general keywords, or specific names of events, places, or people. For example, if learner types in a place name “Berlin”, SugarTube will invoke RESTful services provided by GeoNames and DBpedia<sup>4</sup> to get semantically relevant concepts and data about the city of Berlin such as the facts that it is the capital city of the country of Germany, has latitude of 52.52437° and longitude of 13.41053°, and that more RDF data can be found at the URL <http://dbpedia.org/page/Berlin>.

<sup>4</sup> These and other data sources will be discussed in Section 5.

In addition to specific searching by keywords, the learner can use more other searching methods: through geographical maps, entity recognition of natural language text supplied directly or from a URI, or by time. SugarTube implements the time search using the WorldHistory service, allowing the user to find related historical events and browse them in a graphical timeline format.

Having found relevant conceptual URIs in the Linked Data cloud, SugarTube can then allow the user to navigate the links. Of course, SugarTube can query Annotation for videos that are tagged using the URIs of interest, but it can also find related video material through YouTube, BBC and OpenLearn services. The user can be prompted with related concepts, Web documents, and other resources by SugarTube's analysis of neighbouring point LOD cloud, including RESTful RDF services. For example, from "Berlin", the Sindice Web service can provide further Semantic Web links connected to the related terms "West Berlin", "East Berlin" and "Berlin wall".

## 5 Data sources

Both Annotation and SugarTube make use of several services and data sources on the Web to help the user annotate and find video. Presently, Semantic Web (RDF) sources include:

- Dewey Decimal and Library of Congress classifications: 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,<sup>5</sup> and the resulting taxonomy is presented to the annotator as a browsable tree. The Library of Congress has published its entire classification system in RDF,<sup>6</sup> but this is much too large to present directly to the user. Instead, Annotation provides an interface to the Library of Congress keyword search service, which returns suitable RDF for the user to choose from.
- GeoNames<sup>7</sup> [5]: The GeoNames API is used to identify named locations according using keyword search, or perform reverse lookup to find named locations in a vicinity. The results are used to provide position and category information to allow visual display on maps and URIs to use for annotation.
- OU Linked Data<sup>8</sup>: Currently under development, this repository will extract, interlink, and publish in the LOD cloud material previously available in various disconnected institutional repositories of the Open University.
- Sindice<sup>9</sup> [6]: A semantic search engine, Sindice crawls and collates the Semantic Web (including microformats), and based on this provides services including keyword searching for linked data, and accessing cached fragments of the Semantic Web.

---

<sup>5</sup> <http://dewey.info/>

<sup>6</sup> <http://id.loc.gov/authorities/>

<sup>7</sup> <http://www.geonames.org/>

<sup>8</sup> <http://data.open.ac.uk/>

<sup>9</sup> <http://sindice.com/>

- Zemanta<sup>10</sup>: A service aimed at helping bloggers augment their posts with links and images, Zemanta analyses natural language text to identify various concepts and named entities, returning URIs to Linked Data hosted by, amongst others, DBpedia and Freebase.

In reusing the datasets and services listed above, we make use of another couple of ontologies: the Simple Knowledge Organization System (SKOS) [7], and the WGS84-based RDF vocabulary for geographical data [8]. We also use several non-semantic services including:

- BBC<sup>11</sup>: machine readable programme information.
- Maps: interactive Ajax mapping services from Google<sup>12</sup> and Yahoo.<sup>13</sup>
- OpenLearn<sup>14</sup>: Free access to a subset of the Open University’s course materials, along with metadata.
- YouTube data services<sup>15</sup>: Used in SugarTube to find videos by keyword.

## 6 Related work

There is significant prior work on video annotation tools, differing in two dimensions from the work presented here: they are either fat-client software rather than Web browser based, or do not create Linked Data annotations. An important early system was Vannotea [9] which relied on a dedicated client application, but enabled collaborative annotation, but the annotations were not in a Semantic Web style. M-OntoMat-Annotizer [10] did use Semantic Web annotations, linking them to annotations embedded using MPEG-7, but again was a fat client.

Videos are traditionally searched by syntactic matching mechanism (e.g. [11]), but recently, with more videos being annotated or tagged using in a Linked Data manner, researchers have begun looking at more Semantic Web oriented fashion. The two major approaches are the semantic indexing process and natural language analysis process. The indexing process assumes the video annotations are made from a fixed vocabulary that does not often change (e.g. [12]). Although this process can be efficient, the fixed ontology can introduce a gap between user’s knowledge and indexed annotations. In an educational environment, videos will often be annotated by different teachers or students who will each wish to apply different annotations to the same video in the context of different courses, groups, and key points. The natural language analysis process focuses more on adding semantics tags to the user’s search input (e.g. [13]). However, most of these approaches require some machine-learning mechanism to assist to dynamically add tags, usually restricting its application to smaller, closed domains of discourse.

<sup>10</sup> <http://www.zemanta.com/>

<sup>11</sup> <http://backstage.bbc.co.uk/data/BbcWebApi>

<sup>12</sup> <http://code.google.com/apis/maps/index.html>

<sup>13</sup> <http://developer.yahoo.com/maps/>

<sup>14</sup> <http://openlearn.open.ac.uk/>

<sup>15</sup> <http://code.google.com/apis/youtube/overview.html>



## 7 Conclusion

The Open University has a need to better catalogue and find material from its archives to reuse in teaching. More broadly, as the Internet makes it easier for students to find and use video as part of their education, it is important that we can provide tools to help them do so. We presented two tools that allowed academics to annotate educational video resources in a Linked Data approach, and to subsequently find appropriate video materials by navigating the Web of Linked Data.

The university will be trialing the Annotation tool in the production of a new history course, with professional historians providing the annotations. Future work will integrate other data sources such as video metadata including provenance and rights management, programme transcripts where available, address active collaboration in annotation, and add interfaces for different classes of users.

## Acknowledgements

This research was funded by the Open University's Open Broadcast Unit, and EU FP7 projects SOA4All and NoTube.

## References

1. Berners-Lee, T.: Linked Data (2006) Online at <http://www.w3.org/DesignIssues/LinkedData.html>.
2. Brickley, D., Guha, R.: RDF Vocabulary Description Language 1.0: RDF Schema. W3C recommendation, World Wide Web Consortium (W3C) (February 2004)
3. Brickley, D., Miller, L.: FOAF vocabulary specification 0.91 (2007) Online at <http://xmlns.com/foaf/spec/>.
4. Raimond, Y., Abdallah, S.: The Timeline Ontology (2007) Online at <http://purl.org/NET/c4dm/timeline.owl#>.
5. Vatan, B., Wick, M.: Geonames Ontology (2006) Online at <http://www.geonames.org/ontology/>.
6. Tummarello, G., Delbru, R., Oren, E.: Weaving the open linked data. In: Proceedings of the 6th International Semantic Web Conference. Volume 4825 of Lecture Notes in Computer Science., Springer (2007) 552–565
7. Miles, A., Bechhofer, S.: SKOS Reference. W3C recommendation (August 2009) Online at <http://www.w3.org/TR/skos-reference>.
8. Brickley, D.: Basic Geo (WGS84 lat/long) Vocabulary (2004) Online at <http://www.w3.org/2003/01/geo/>.
9. Hunter, J., Schroeter, R., Koopman, B., Henderson, M.: Using the Semantic Grid to Build Bridges Between Museums and Indigenous Communities. In: Proceedings of the GGF11—Semantic Grid Applications Workshop. (June 2004)
10. Bloehdorn, S., Petridis, K., Saathoff, C., Simou, N., Tzouaras, V., Avrithis, Y., Handschuh, S., Kompatsiaris, Y., Staab, S., Strintzis, M.: Semantic Annotation of Images and Videos for Multimedia Analysis. In: Proceedings of the 2nd European Semantic Web Conference. Volume 3532 of Lecture Notes in Computer Science., Springer (2005)

11. Wei, S., Zhao, Y., Zhu, Z., Liu, N.: Multimodal fusion for video search reranking. *IEEE Transactions on Knowledge and Data Engineering* **22** (2010) 1191–1199
12. Hakeem, A., Lee, M.W., Javed, O., Haering, N.: Semantic video search using natural language queries. In: *MM '09: Proceedings of the seventeen ACM international conference on Multimedia*, New York, NY, USA, ACM (2009) 605–608
13. Sebastine, S.C., Thuraisingham, B.M., Prabhakaran, B.: Semantic web for content based video retrieval. In: *Proceedings of the 3rd IEEE International Conference on Semantic Computing (ICSC 2009)*, 14-16 September 2009, Berkeley, CA, USA, IEEE Computer Society (2009) 103–108