VideoLecturesMashup: using media fragments and semantic annotations to enable topic-centred e-learning

Lyndon Nixon[1], Tanja Zdolsek[2], and Ana Fabjan[2] and Peter Kese[3]

[1] MODUL University, Vienna, Austria
lyndon.nixon@modul.ac.at
[2] Jozef Stefan Institute, Ljubljana, Slovenia
tanja.zdolsek@ijs.si
[3] VIIDEA d.o.o., Kranj, Slovenia
peter@viidea.com

Abstract. In this demo, we present the VideoLecturesMashup, which delivers re-mixes of learning materials from the VideoLectures.NET portal based on shared topics across different lectures. Learners need more efficient access to teaching on specific topics which could be part of a larger lecture (focused on a different topic) and occur across lectures from different collections in distinct domains. Current e-learning video portals can not address this need, either to quickly dip into a shorter part focused on a specific topic of a longer lecture or to explore what is taught about a certain topic easily across collections. Through application of video analysis, semantic annotation and media fragment URIs, we have implemented a first demo of VideoLecturesMashup.

1 Introduction

Currently the VideoLectures.NET portal hosts more than 16.000 video lectures from prominent universities and conferences mainly from natural and technical sciences. Most lectures are 1 to 1.5h long linked with slides and enriched with metadata and additional textual contents. Videolectures.NET is being visited by more than 15.000 unique visitors from all over the world daily, which provides a very efficient distribution and dissemination channel.

However, visitors typically have limited time to find and watch the materials they want and the topics they search for may be orthogonal to the materials themselves (be the subject of different parts of multiple learning resources rather than the subject of a specific complete learning resource). Visitors would benefit from easier and quicker access to those different parts in the form of a single, integrated presentation of learning materials, which in turn could drive more repeated access and win new users, including in new contexts, e.g. dynamic provision of such learning resource mashups would be particularly useful in mobile consumption contexts (where the user typically has more limited time and a restricted browsing interface). These mash-ups could subsequently form a new distribution channel for VideoLectures.NET contents.

Hence we have proposed a use case in the MediaMixer project for the VideoLecturesMashup which will be a dedicated channel on the VideoLectures.NET portal capable of accepting a specific learning topic as input and producing as a result a mash up of fragments of learning materials from the site addressing that topic, ordered in a meaningful way.

2 Technology Used

Currently, the search and retrieval on VideoLectures.NET works on text matching over complete materials titles and descriptions. Not even the internal descriptions that are currently maintained (e.g. slide titles and contents) can be used in the site search. These internal descriptions (where content of the videos is tied to specific, mainly temporal, fragments) need to be more detailed, and the slide boundaries in the presentation (which can be calculated) linked to the correct temporal boundaries in the video (since a slide may be shown before the speaker starts to reference it, or they reference it before it is shown). This will require additional analysis processes being applied to the learning materials video to generate this annotation. This can usually not be included during recording, even signaling when the speaker refers to the next slide is difficult for a cameraperson probably not knowledgeable about the speakers subject. Thus in post-processing of audio, video and the associated slides, VideoLectures.NET must incorporate:

1. automatic textual transcription from speaker audio (ASR). We make use of the transLectures-UPV toolkit (TLK) [4] , an open source set of ASR tools for video lectures.
2. concept extraction from slides (not just titles but textual content extracted via OCR technology). We are looking at the use of the solution from the Hasso Plattner Institute Potsdam [5] which is already used in yovisto video search (`http://yovisto.com`).
3. video analysis e.g. identification of spatial fragments of video with the speaker, slides and other objects. Here we use a set of tools courtesy of the research centre CERTH [6].

This richer annotation uses semantic technology, since associating a spatial or temporal fragment [1] to semantic concept (rather than, e.g. a simple text label) gives additionally the possibility to link that fragment to that concepts synonyms or related concepts in a semantic search and retrieval system. A metadata schema for the annotations has been selected, as well as a choice of vocabularies which contain the relevant concepts and provide (semantic) links to related concepts (e.g. within a taxonomy or classification scheme). An appropriate repository was provided to store the resulting (semantic) metadata and allow for efficient indexing and retrieval by a search agent. It is used alongside the current storage

[4] `http://www.translectures.eu/tlk/`
[5] `http://www.yanghaojin.com/research/ACM-MM-GC-DEMO/`
[6] `http://multimedia.iti.gr/mediamixer/demonstrator.html`

solution with the use of shared unique IDs for learning resources to provide a link between data in both stores. Automatic analysis handles timing of slide changes in the video, for example, however manual correction may likely still be important for the results of automatic concept detection. It may be that the accuracy of the processes is sufficient to rely on it in user search - however irrelevant results may be less tolerated in a mash up situation. Given the need for specialist understanding of the topic, one future option is to incentivize the learning resource creator to correct the annotation of their learning resource. Another is to rely on crowdsourcing, whether Mechanical Turk or the learning resource viewers themselves. Given the availability of richer annotation of the learning materials, a semantic search and retrieval module is provided for the fragment selection. Given the association of media fragments to semantic concepts, this module is able to match the input topic to concepts in the annotations via the use of appropriate ontologies (logical models of how different concepts relate to one another). There are three core functions performed by such a module:

1. the input topic is internally modelled as a semantic concept;
2. the annotated learning resources are internally indexed in terms of the concepts they are associated to, and
3. the module is able to calculate a match via semantic proximity between the concepts in the input topic and the concepts in a learning resource fragments annotation.

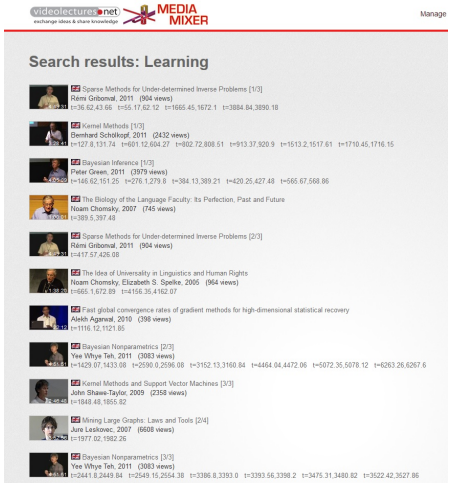
This semantic search module replaces in VideoLecturesMashup the text based search module used by VideoLectures.NET The results list no longer contains complete resources but fragments in terms of spatial or temporal divisions of the learning resources video. This required that VideoLectures.NET incorporates on both its media server and its embedded video player the necessary support for the Media Fragments specification [2].

3 VideoLecturesMashup demonstrator

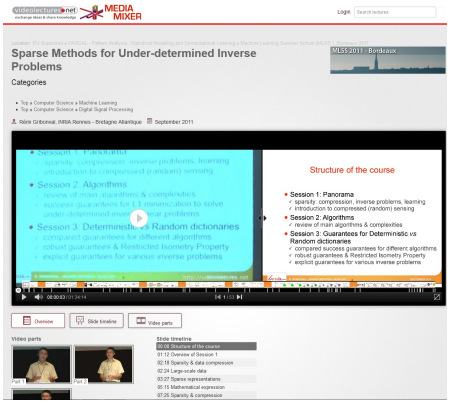
The first version of the VideoLecturesMashup demonstrator at `http://mediamixer.videolectures.net` shows the retrieval of media fragments based on user search.

Accessing the online demo the user sees a search bar and can conduct a search on keywords for their topic of interest. For instance, when entering as search key-word 'Learning' the user gets 12 video matches. For each video, users see a thumbnail and some metadata (title of the lecture, name of the lecturer, year of the lecture, number of views). Underneath, the fragments of the video which match the search term are listed, in this case we find a total of 35 fragments mentioning 'statistics' (Figure 1(a)).

The user can click on one of the listed videos or directly on the listed fragments to watch the video/fragments. For example, as shown in Figure 1(b), if a user clicks on the first video on the list, then the system will show the whole lecture title on the top, below the information of which categories the video is categorized in, and information about the lecturer. On the right, it shows



(a) User interface showing matching video fragments for 'Learning'



(b) Video fragment playback and information

Fig. 1: VideoLecturesMashup

a picture banner, which shows at which event the watched lecture was given. After all this metadata the VideoLectures.NET player, which composed of the usual VideoLectures.NET layout (video on the left and sync slides on the right), is presented. Below the player, five features are presented: overview (short description, slide timeline), description (longer description), slide timeline (all slide timelines, a result of the video with slides synchronization), authors (description of the lecturer) and fragments (list of the matched fragments with timing).

With the integration of semantic search, the following aspects become feasible for the learner:

1. Finding video fragments via multilingual search. Since DBPedia extracts metadata from Wikipedia in all available languages, it also stores links between resources across the different language pages. Thus the term 'Learning' used in an English language lecture can still be found when the user searches for 'Lernen' (German).
2. Finding video fragments across synonyms. Since DBPedia also captures the information of Wikipedia's disambiguation and redirection pages, it can associate a resource with other terms which have been considered synonyms or clarifications of that resource. Again, a search for 'Learn' disambiguates to the term 'Learning' or for 'Acquisition' (in psychology) redirects to 'Learning', based on the already available DBPedia metadata.
3. Finding video fragments on related subjects or topics. DBPedia has a very complete categorization scheme, putting almost all resources into one or more categories, which themselves are organised in a large taxonomy. We consider fragments about topics which belong to the same category as the topic the

user searched for as relevant. For example, the term 'Learning' happens to be categorized under Developmental psychology, Cognitive science and Intelligence. The category of Cognitive Science happens to have many other terms associated to theories about learning, such as the Semantic feature-comparison model, Dual-coding theory or Narrative inquiry, hence we can associate video fragments mentioning these terms to a search for 'learning'.

An explanatory video of the demo functionalities can be seen at `http://bit.ly/videolecturesmashup` (the UI shown is an earlier version of VideoLecturesMashup).

4 Conclusion

The current VideoLecturesMashup demonstrates the value of semantic multimedia and media fragment technology in enabling an e-learning video platform to offer learners a topic-centred path into parts of larger video lectures across various collections. As such, it provides a different structure to learning than current MOOCs which focus on individual courses which are curated with selected content from the outset. The learner's experience with re-mixes of materials needs further evaluation in terms of resulting satisfaction and further improvement of the visual interface will be part of this. [7].

Acknowledgments

This work was supported by the MediaMixer project, funded by the EU Framework Programme 7 (http://www.mediamixer.eu). MediaMixer offers a free community portal with access to materials about semantic multimedia technologies (http://community.mediamixer.eu)

References

1. R. Troncy, L. Hardman, J. van Ossenbruggen, and M. Hausenblas, Identifying Spatial and Temporal Media Fragments on the Web. W3C Video on the Web Workshop, 2007
2. D. Van Deursen, R. Troncy, E. Mannens, S. Pfeiffer, Y. Lafon and R. Van de Walle, Implementing the Media Fragments URI Specification. WWW 2010 Developers Track, Raleigh (USA), 2010.

[7] MediaMixer together with VideoLectures.NET ran a Grand Challenge at the ACM Multimedia 2013 conference to find new solutions for the temporal segmentation of video lectures. The winning proposal has an appealing visualisation of video fragment interlinking which we now examine as a potential UI extension for VideoLecturesMashup. Demo at `http://portal.klewel.com/graph/`