

## Video Lectures Mashup – remixing learning materials for topic-centred learning across collections

Lyndon Nixon, MODUL University

[lyndon.nixon@modul.ac.at](mailto:lyndon.nixon@modul.ac.at)

Tanja Zdolšek, Jožef Stefan Institute

[tanja.zdolsek@ijs.si](mailto:tanja.zdolsek@ijs.si)

Ana Fabjan, Jožef Stefan Institute

[ana.fabjan@ijs.si](mailto:ana.fabjan@ijs.si)

Peter Keše, VIIDEA d.o.o.

[peter@viidea.com](mailto:peter@viidea.com)

### Abstract

In this paper, we introduce the VideoLecturesMashup, which presents 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 media technologies promoted by the MediaMixer project<sup>1</sup> – semantic annotation and media fragment URIs – we have implemented a first demo of VideoLecturesMashup.

### Keywords

VideoLectures.NET, e-learning, video lectures, learning materials, video repository, media fragments, semantic annotation.

### 1. Introduction

Currently the VideoLectures.NET portal (Figure 1) 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. Further to that, Videolectures.NET is tightly integrated in the three world scale communities of higher education institutions that are providing open access to free educational contents. The three communities OpenCourseWare Consortium (OCWC), Opencast and Knowledge4All combine altogether more than 600 HigherEd (mainly Universities) around the world including the first ten highest ranked Universities. This provides on one side a huge market and on the other a unique dissemination channel to reach world scale impact.

Visitors to VideoLectures.NET are looking to consume learning materials on specific topics of interest. 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

---

<sup>1</sup> <http://mediamixer.eu>; access all materials about MediaMixer technologies at <http://community.mediabox.eu>

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 “mash ups” 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 (e.g. video streams / TV channels on selected topics) and be integrated into other learning channel offers (mobile like iTunesU, or a virtual TV channel as a SmartTV application).

Figure 1. The VideoLectures.NET portal

Source: <http://videolectures.net>

The screenshot shows the VideoLectures.NET portal interface. At the top, there are logos for 'videolectures.net', 'World Summit Award', and 'WSA'. A navigation menu includes 'HOME', 'BROWSE LECTURES', 'PEOPLE', 'CONFERENCES', 'ACADEMIC ORGANISATIONS', 'EU SUPPORTED', 'ABOUT US', and 'BLOG'. A search bar is on the right. Below the navigation is a row of featured lectures with speaker photos and titles. On the left is a 'CATEGORIES' sidebar listing various subjects like Architecture, Arts, Astronomy, etc. The main content area is titled 'NEWS' and contains several articles, including 'The crew is off to Paris to collect the UNESCO award', 'United Nations and UNESCO to award VideoLectures.Net best educational product of the decade', 'New MIT OCW courses', and 'NIPS 2012: Posner Lectures, Invited talks, Workshops, all published!'. On the right is a 'NEWSLETTER' section with a subscription form and a 'RECENT EVENTS' section listing conferences like the 27th European Congress of Arachnology (ECA) and 7.01SC Fundamentals of Biology.

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. The mash up will be specifically addressable and hence bookmarkable/saveable for subsequent reference and viewing. The MediaMixer project is a Support Action funded by the EU to promote semantic multimedia and media fragment technology. Through supporting different use cases through to prototypical technological integration and proof-of-concept demonstrators, MediaMixer aims to show the potential value of the media technology in different industry domains, including the e-learning community via VideoLecturesMashup.

In this paper, we will look at the relationship of this work to prior efforts and the state of the art in e-learning and multimedia technology in the following Chapter 2. We will then introduce the design of the use case and the technology used in Chapter 3. The outcome, the VideoLecturesMashup demonstrator, is described in Chapter 4. Finally we look at future work and evaluation of the results in Chapter 5.

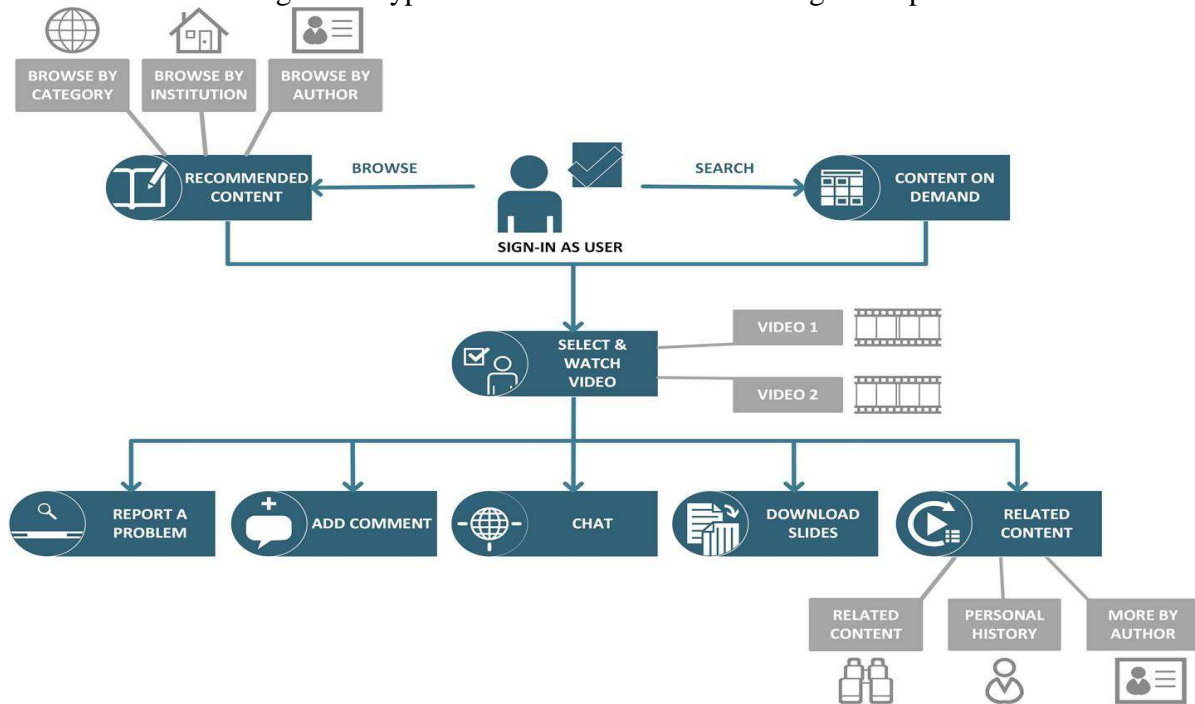
## **2. State of the art**

There are several hugely popular websites among scholars, students, professionals and the general public such as Google Talks, TED, Videolectures.NET, Yovisto, or tele-TASK that pursue recordings of rich media footage (RMF). RMFs are collections of multimedia materials used for presentations that are often supported by synchronised slides, embedded video, presenter descriptions and comments, user metadata including comments, ratings, related embedded and linked textual materials and also the transcript of the text in languages other than the original.

Users typically enter these video portals and find material of interest via search along different facets (authors, institutions, category). From a list of matched videos to their search query, they select and begin watching material based either on a ranking (perceived to be) performed by the site itself (i.e. top listed contents are typically assumed to be the most relevant in search results) or by preferences the users themselves have (e.g. picking out speakers whose name they recognise). The links returned in the search are to complete materials and the matching is based on textual retrieval techniques, such as string matching between the search term and textual metadata about the video lecture (title, description). Once a video is selected, it is streamed and viewed; typical actions available to the user when consuming the material would be to add comments or download contents. The only links from this material to others would be based on factors such as material by the same author, related material by collaborative filtering (other users who watched this video also watched...) or user's own history (together with this video, you previously watched...). Typical user workflow on learning video portals is illustrated below (Figure 2).

Portals are focused on improving the visitor's experience and thus, in the e-learning context, supporting the visitor's learning goals. Different aspects of use of e-learning video portals contribute to this goal, including personalisation, social networking or, the aspect we are focused on, improving the search for and access to appropriate learning materials. Improved search and retrieval of media assets from a larger collection has long been tied to the extraction of detailed metadata (descriptions) for those media assets which can be used in matching them to a search query [1][2]. There is a gradual shift taking place from unstructured, ambiguous forms of metadata (e.g. notes taken in natural language) to more structured metadata re-using controlled vocabularies and domain models [3], which supports more accurate media search and retrieval, alongside other applications [4].

Figure 2. Typical user workflow on a learning video portal



While the introduction of “semantic search” [5], where the media is described in terms drawn from a knowledge model called an ontology and the search term is equally converted into terms from the same model so that matches can be made not only on terms expressed as being synonymous but also via related, more specific or more general terms, can improve retrieval of appropriate learning materials, this still functions only at the level of the whole media object as long as the semantic description also is made only for the whole media object.

The multimedia community has long understood that every media object is potentially the sum of many distinct parts each of which may have its own, separate, meaning for a media consumer. Any audiovisual work is the sum of a video track, one or more audio tracks and zero or more text tracks (subtitles, commentaries...). A longer video can be considered the sum of many shorter, distinct videos (termed, e.g. “chapters”). An image is the sum of its regions, some of which standalone in showing a distinct concept. Hence media descriptions may also describe the salient parts of the media asset and apply descriptions individually to each part, such that a search/retrieval task may return only parts of a larger media asset that are relevant to the search term such as a spatial region of an image or a temporal segment of a video.

Referring to, let alone accessing and playing back, a salient part of some media asset (hereafter referred to as a “media fragment”) in a commonly understood, Web friendly way remains a challenge today since many media sites, if they support media fragments at all, choose to use a proprietary syntax and need to implement a specific playback logic for those fragments. This situation is changing with the introduction of the Media Fragments URI specification by the World Wide Web Consortium (W3C) [6] which provides for a standardized, URL-friendly (the address format of the Web) syntax for referring to a fragment and the slow but steady support for

this syntax in Web browsers and media players. Not only can a Media Fragment URI be used as the subject for the description of a fragment of some media, but eventually it can be directly passed to a Web browser or media player for playback of that fragment to the consumer [7].

While semantic search for better retrieval of appropriate media assets and media fragments for focusing that retrieval to the part of the media which is most relevant both appear to be clearly beneficial to e-learning video portals, the extent of their usage in current portal offers is very limited. To our knowledge, only YoVisto ([www.yovisto.com](http://www.yovisto.com)) enables pinpoint content-based search via state-of-the-art video analysis technologies, e.g. shot boundary detection, video OCR, and automated speech analysis (ASR). Users are able to tag and comment video fragments for educational purposes. In this way, online courses can be complemented with exercises, discussions, student guidelines, as well as additional learning material. Moreover, Yovisto video metadata and search results are semantically annotated and published [8][9]. Yet the focus of the video annotation was on improved search and retrieval, not the re-use of the semantic information for re-mixes of lecture fragments across a topic, so the experience of VideoLecturesMashup could potentially next be applied to YoVisto material.

Hence VideoLecturesMashup is intended to go beyond the state of the art in e-learning video portals, demonstrate the value of the semantic media and media fragment technology, and act as an inspiration and example for other video-based e-learning offers to take up the technology and improve their effectiveness for online learners.

There are two main reasons why and how VideoLectures.NET profits from semantic multimedia technology enabling fragmented content in comparison with current approaches:

1. According to the VideoLectures.NET portal visitor analysis main part of the visitors are looking for very specific information inside the lecture and are not interested to watch the entire lecture. This is why the access to the interconnected video content fragments is important.
2. VideoLectures.NET is tightly integrated in the three main communities of higher education institutions that are committed to provide open access to open lectures at their sites. The main effort in these communities is to find out how to interlink multimodal contents across many multi-lingual sites. The technology to support semantically interlinked multimedia fragmented video content lies in the core of these attempts. Currently there are no feasible solutions to support this aspect.

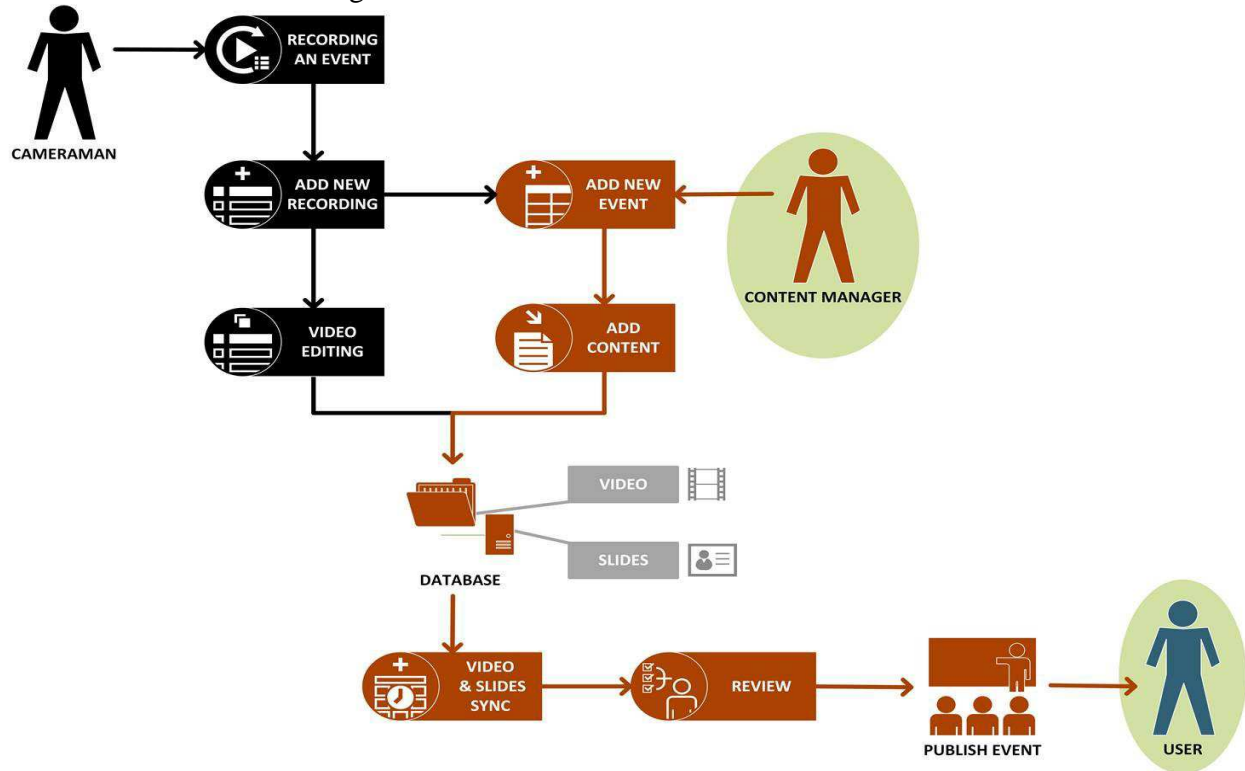
### **3. Use case design**

In this use case, the principle actors are the users (scientific community generating the learning materials with the support of their universities, and the learners who seek to access those materials in a suitable and appropriate fashion), aided by the content manager at VideoLectures.NET who is in charge of preparing the materials on the portal.

The below diagram (Figure 3) introduces the current principle actors in VideoLectures.NET with their activities in using the portal. In terms of benefiting these user types, we note that VideoLecturesMashup can help distribute more widely the materials available on the learning

materials portal by reaching learners who may not have otherwise found that material at all (since its primary subject is not the topic the learner may have searched for), promoting cross-disciplinary learning and promoting further the content of the providing institutions (NB. this requires that the mash up can still associate the selected fragments in the mash up with the content provider) and the work of the presenters (since presentation of their work to new communities may seed new applications or co-operations).

Figure 3. VideoLectures.NET basic workflow



In particular, the workflow for the user is simplified in VideoLecturesMashup by offering a direct intuitive access to a single learning channel built around the topic searched for, i.e.

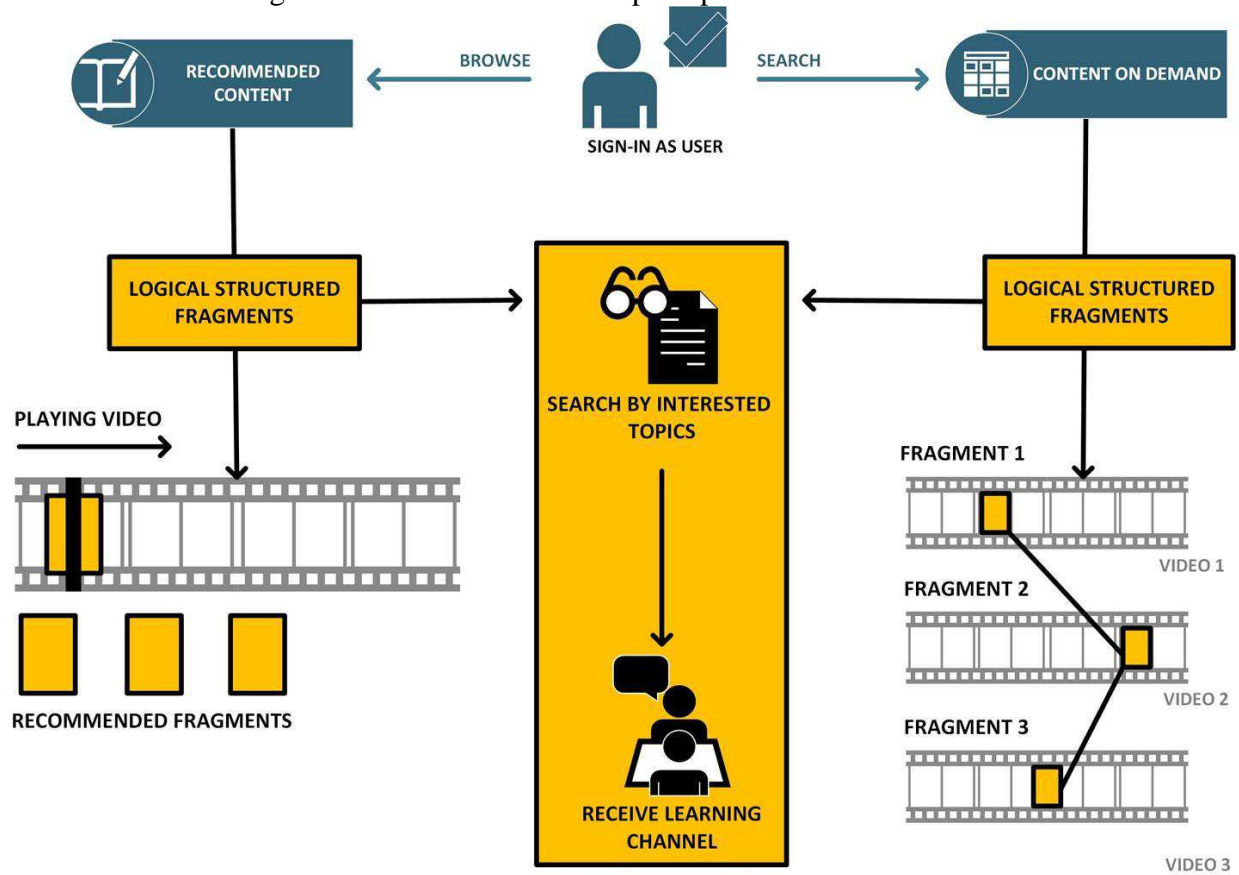
*search topic interested in ---> receive learning channel*

This workflow may be particularly useful where the user is using a device where browsing is more restricted (TV, mobile) or is seeking to quickly access video which they can begin to consume (time restricted to browse all search results and select which they want). In comparison, there is less direct control over the selected material, which makes the relevance of the selection even more important, and some options may be desirable to offer the user in an interface (e.g. only use recordings made within a certain time period). VideoLecturesMashup user workflow is illustrated below (Figure 4).

The technical implementation of this use case is based on the adoption of new multimedia technology for analysis, annotation and publication of the learning video materials (as referenced

in Section 2) in such a way that the video portal can offer users topic-based search and media fragment retrieval for the play-out of learning “remixes”.

Figure 4. VideoLecturesMashup simplified user workflow



Currently, the search and retrieval works on text matching over complete materials titles and descriptions. Not even the internal descriptions that are currently maintained by VideoLectures.NET (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 speaker’s subject. Thus in post-processing of audio, video and the associated slides, VideoLectures.NET needs to incorporate:

- automatic textual transcription from speaker audio (ASR). We make use of the transLectures-UPV toolkit (TLK)<sup>2</sup>, an open source set of ASR tools for video lectures.
- concept extraction from slides (not just titles but textual content extracted via OCR technology). We are looking at the use of the HPI solution<sup>3</sup> which is already used in YoVisto.
- video analysis e.g. identification of spatial fragments of video with the speaker, slides and other objects. Here we have access to a set of tools courtesy of the research centre CERTH<sup>4</sup>.

This richer annotation will use semantic technology, since associating a spatial or temporal fragment to a semantic concept (rather than, e.g. a simple text label) gives additionally the possibility to link that fragment to that concept's synonyms or related concepts in a semantic search and retrieval system. A metadata schema for the annotations needs to be 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 needs to be provided to store the resulting (semantic) metadata and allow for efficient indexing and retrieval by a search agent. It could be used alongside the current storage solution with the use of shared unique IDs for learning resources to provide a link between data in both stores.

Automatic analysis should be able to handle 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 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.

Assuming the availability of richer annotation of the learning materials, a semantic search and retrieval module needs to be 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 set of semantic concepts;
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

---

<sup>2</sup> <http://www.translectures.eu/tlk/>

<sup>3</sup> <http://www.yanghaojin.com/research/ACM-MM-GC-DEMO/>

<sup>4</sup> CERTH demonstrator for results of concept detection and shot segmentation over selected e-learning video lectures can be seen at <http://multimedia.iti.gr/mediamixer/demonstrator.html>



The results list will contain not complete resources but fragments in terms of temporal divisions of the learning resources video. Rather than presented as a list to the user they can also be played out as a single video stream (while the interface may allow for browsing options, e.g. jump forward or back between fragments). This requires that VideoLectures.NET incorporates on both its media server and its embedded video player the necessary support for the Media Fragments specification.

#### **4. Outcome: VideoLectures Mashup**

Fragment retrieval and presentation was implemented as a pluggable extension to the existing VideoLectures.NET technology platform. For this, specific interfaces had to be built into the platform itself to provide placeholders, where the MediaMixer technology can be plugged-in, in order to extend and complement existing site functionality. In addition, several components were created and extended to provide fragment support.

As VideoLectures.NET is platform agnostic and is using either Flash or HTML5 technology to provide optimal experience on all devices, both Flash and HTML5 media player codebases were extended, to complement existing lecture presentation experience with fragment functionality. A separate XML document (besides the existing SMIL XML timeline used in synchronizing video and slides during playback) is thus formed and provided to both players to describe fragments, which need to be played and/or highlighted, in terms of Media Fragments URI specification.

#### **Media Fragment Creation**

Internally, the fragments are based on the results of shot segmentation analysis, which identifies when the video focuses on a different object during the talk. However, initial experiments indicate that this splits the talks too much since a switch between, e.g. the speaker, the blackboard and the audience, may occur while the same topic is still being discussed. Thus we now look at fragmentation based on the slide synchronization timeline, which is indicative of when the lecturer shifts between discussion on different topics within their talk, while the shot segmentation process will be refined by identifying continuous speech between shots as an indicator of staying in the same topic and will prevent a shot boundary from being generated. This provides us with more continuous shots within the video pertaining to distinct topics. Thus we will experiment with both options in this prototyping phase to analyse which gives the better results when watching the remixes.

#### **Media Fragment annotation**

Temporal fragments are annotated with concepts extracted from the available textual information, be it from slide text or subtitles. Available subtitle files or slide transcriptions were parsed, analyzed using an entity recognition service and the resulting annotations stored. An entity recognition service identifies entities of interest within texts and marks those entities with an identifier, to provide for disambiguation of the word or term. We use the NERD service which aggregates several online entity recognition services (<http://nerd.eurecom.fr>) and have focused our entity annotation on results from DBPediaSpotlight and TextRazor, both of which use Wikipedia articles to identify the entities uniquely. The entity identifiers (URLs) are normalized to use DBPedia, which is a structured metadata conversion of information from Wikipedia and acts as a global online knowledge model [10]. By using DBPedia, we could merge additional information about the concepts extracted from the learning material metadata into the metadata

store by retrieving statements about the concept from the Web. For example, DBPedia offers a rich categorization taxonomy and individual entities are given both types (classes of thing) and subjects (categories they belong to). Concepts are also linked to equivalent concepts as well as labels in various languages, so that both search term synonyms and multilingual search can be supported.

### **Media Fragment management**

VideoLectures.NET is a large and constantly updated website that must provide reliable service on day-to-day basis. Bringing the MediaMixer technology into everyday production system requires a fully dynamic and online support of all proposed services. Currently, all fragment management, creation and analytics are being run offline on a non-production database. The aim of the project is to enrich the internal interfaces of MediaMixer with a set of signals. These signals will either call external MediaMixer analytical services, each time a set of site content is added or updated and vice versa; update the site content when MediaMixer services provide additional or updated fragment and annotation information. Annotations are stored in a metadata repository referring to video assets on the production server. While media fragment support can be added to a media server to ensure only the fragment under consideration is delivered to the client (and not the entire media item, which can be considerably larger - consider the difference between a 2 hr lecture and a 5 minute snippet), since media delivery is part of the production architecture this has not been implemented.

### **Media Fragment retrieval & playback**

The fragment search and retrieval engine is designed as a pluggable architecture. The initial implementation was based on a simple text-search from a text index generated out of the subtitles or slide transcripts and stored in Apache SOLR. Once the semantic descriptions are generated based on the concept extraction from the text, this is processed and made available for VideoLecturesMashup via the semantic data repository which supports “semantic search”. Queries can automatically consider synonyms and terms in other languages, and term matches can be programmed to take in “conceptually close” terms using the additional descriptions sourced from DBPedia, e.g. 'if another term belongs to the same DBPedia category as the search term, match this term with a lower ranking'. The result of the query is not the terms matched, but the media fragments in the repository which are described with those terms. This list of media fragments can be explored one by one in a regular results list but we plan to provide for linear playback of the fragments after each other similar to a topic-specific learning TV channel in a next release.

### **Demonstrator UI**

The first version of the VideoLecturesMashup demonstrator shows the retrieval of media fragments based on user's search.

Accessing the online demo the user sees a search bar and can conduct a search on keywords for their topic of interest (example, learn or structure). For instance when entering as search keyword 'Learning' the user gets twelve matches, referring to the videos as a whole. 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 'Learning'.

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 6, 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 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:

- 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 “Thermodynamics” used in an English language lecture can still be found when the user searches for “Termodynamika” (Polish) or “Varmafræði“ (Icelandic).
- 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 “Thermic” 'disambiguates' to Thermodynamics or for “Thermo-dynamics” 'redirects' to Thermodynamics, based on the already available DBPedia metadata.
- 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 Thermodynamics happens to be categorized in Chemical Engineering, Concepts in Physics and a self-named category of Thermodynamics. That last category happens to have many other terms associated to it, such as László Tisza (who authored the key literature “Generalised Thermodynamics”), Vortex Tube (an object used in fluid dynamics) or the Joule-Thomson Effect (a key law of thermodynamics), hence we can associate video fragments mentioning these terms to a search for learning about thermodynamics<sup>5</sup>.

---

<sup>5</sup> A full list of terms in the category of Thermodynamics can be seen at <http://dbpedia.org/page/Category:Thermodynamics> (values of the property “is dcterms:subject of”)

Figure 5. Search results for the topic 'Learning'

videolectures.net exchange ideas & share knowledge  Manage

---

## Search results: Learning












-  **Sparse Methods for Under-determined Inverse Problems [1/3]**  
Rémi Gribonval, 2011 (904 views)  
t=36.62,43.66 t=55.17,62.12 t=1665.45,1672.1 t=3884.84,3890.18
-  **Kernel Methods [1/3]**  
Bernhard Schölkopf, 2011 (2432 views)  
t=127.8,131.74 t=601.12,604.27 t=802.72,808.51 t=913.37,920.9 t=1513.2,1517.61 t=1710.45,1716.15
-  **Bayesian Inference [1/3]**  
Peter Green, 2011 (3979 views)  
t=146.62,151.25 t=276.1,279.8 t=384.13,389.21 t=420.25,427.48 t=565.67,568.86
-  **The Biology of the Language Faculty: Its Perfection, Past and Future**  
Noam Chomsky, 2007 (745 views)  
t=389.5,397.48
-  **Sparse Methods for Under-determined Inverse Problems [2/3]**  
Rémi Gribonval, 2011 (904 views)  
t=417.57,426.08
-  **The Idea of Universality in Linguistics and Human Rights**  
Noam Chomsky, Elizabeth S. Spelke, 2005 (964 views)  
t=665.1,672.89 t=4156.35,4162.07
-  **Fast global convergence rates of gradient methods for high-dimensional statistical recovery**  
Alekh Agarwal, 2010 (398 views)  
t=1116.12,1121.85
-  **Bayesian Nonparametrics [2/3]**  
Yee Whye Teh, 2011 (3083 views)  
t=1429.07,1433.08 t=2590.0,2596.08 t=3152.13,3160.84 t=4464.04,4472.06 t=5072.35,5078.12 t=6263.26,6267.6
-  **Kernel Methods and Support Vector Machines [3/3]**  
John Shawe-Taylor, 2009 (2358 views)  
t=1848.48,1855.82
-  **Mining Large Graphs: Laws and Tools [2/4]**  
Jure Leskovec, 2007 (6608 views)  
t=1977.02,1982.26
-  **Bayesian Nonparametrics [3/3]**  
Yee Whye Teh, 2011 (3083 views)  
t=2441.8,2449.84 t=2549.15,2554.38 t=3386.8,3393.0 t=3393.56,3398.2 t=3475.31,3480.82 t=3522.42,3527.86

Figure 6. Media fragment playback and browsing

The screenshot displays the VideoLecturesMashup interface. At the top, there is a navigation bar with the 'videolectures.net' logo and 'MEDIA MIXER' branding. A search bar is located on the right. The main content area features a video player with a slide timeline. The video player shows a slide titled 'Structure of the course' with a bulleted list of sessions and their topics. Below the video player, there are three tabs: 'Overview', 'Slide timeline', and 'Video parts'. The 'Slide timeline' tab is active, showing a list of time points for different parts of the course. The 'Video parts' tab shows a grid of video thumbnails. The video player itself has a play button in the center and a progress bar at the bottom.

Location: EU Supported » PASCAL - Pattern Analysis, Statistical Modelling and Computational Learning » Machine Learning Summer School (MLSS 1, Bordeaux 2011)

## Sparse Methods for Under-determined Inverse Problems

MLSS 2011 - Bordeaux

Categories

- Top » Computer Science » Machine Learning
- Top » Computer Science » Digital Signal Processing

Rémi Gribonval, INRIA Rennes - Bretagne Atlantique September 2011

**Structure of the course**

- Session 1: Panorama
  - ✓ sparsity: compression, inverse problems, learning
  - ✓ introduction to compressed (random) sensing
- Session 2: Algorithms
  - ✓ review of main algorithms & complexities
  - ✓ success guarantees for L1 minimization to solve under-determined inverse linear problems
- Session 3: Deterministic vs Random dictionaries
  - ✓ compared success guarantees for different algorithms
  - ✓ robust guarantees & Restricted Isometry Property
  - ✓ explicit guarantees for various inverse problems

00:00:03 / 01:34:14 1 / 53

Overview Slide timeline Video parts

Video parts

Slide timeline

- 00:00 Structure of the course
- 01:12 Overview of Session 1
- 02:18 Sparsity & data compression
- 02:24 Large-scale data
- 03:27 Sparse representations
- 05:15 Mathematical expression
- 07:25 Sparsity & compression

VideoLecturesMashup can be tested at <http://mediamixer.videolectures.net>. An explanatory video of the demo functionalities can be seen at <http://bit.ly/videolecturesmashup>.

## 5. Outlook: Evaluation

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. While it is clear that this approach can ensure a desired quality in the learning offer, it proves to lack flexibility for the individual learner and omits the value of learning material snippets that may be found in collections outside of the consideration of the MOOC. As such, VideoLecturesMashup is also designed to reflect another possibility with

online learning, when learning materials are richly annotated and thus could be remixed by learning systems to meet different learners' needs.

The value of such (automated) remixes lies of course heavily on the accuracy and completeness of the underlying annotations, which needs to be further studied to improve accuracy in automated annotation systems alongside examination of techniques to involve humans in correction of the annotations, e.g. crowdsourced via the learners themselves. While DBpedia provides a very complete, globally available and multilingual knowledge base for reference to extracted concepts and access to additional metadata about them, within the learning context, there is a need for more specialised domain models for specific subjects which could enable a better selection and ranking of learning materials for re-mixes. The learner's experience with re-mixes of materials needs further evaluation in terms of resulting satisfaction and is necessarily connected to trials of user interfaces and user experience since the learner needs to understand via the system how the offered mix of learning material fragments relates to their search term and promotes further learning about the topic. Visual interfaces can play a very important role in ensuring learners do not feel lost within the learning material re-mix and can intuitively understand the collection of materials they are browsing, where they are within the collection and how the materials relate to one another<sup>6</sup>.

Together with VideoLectures.NET we ran a Grand Challenge at ACM Multimedia 2013 to find new solutions for the temporal segmentation of video lectures. The winning proposal has an appealing visualisation of video fragment interlinking which we will examine as a potential UI expansion for VideoLecturesMashup.

## 6. 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>).

## 7. References

- [1] Polfreman, M, Broughton, V and Wilson, A (2008). Metadata Generation for Resource Discovery, AHDS Report.
- [2] Christel, M (2009). Automated Metadata in Multimedia Information Systems, in Synthesis Lectures on Information Concepts, Retrieval, and Services #2, Morgan & Claypool.
- [3] Bachmann, T (2009). Metadata explained: the evolution of schemas, taxonomies and ontologies in today's DAM systems, in Journal of Digital Asset Management 5:286-297.
- [4] Nixon, L, Dasiopoulou, S, Evain, J-P, Hyvönen, E, Kompatsiaris, I, Troncy, R (2011). Multimedia Broadcasting and eCulture, chapter in the book Handbook of Semantic Web Technologies, Springer. ISBN 978-3-540-92912-3.
- [5] Mäkelä, E (2005). Survey of Semantic Search Research, Seminar on Knowledge Management in the Semantic Web, Helsinki. <http://www.seco.tkk.fi/publications/2005/makela-semantic-search-2005.pdf>
- [6] Troncy, R, Hardman, L, Ossenbruggen, J van, and Hausenblas, M (2007). Identifying Spatial and Temporal Media Fragments on the Web. W3C Video on the Web Workshop.

---

<sup>6</sup> Demo at <http://portal.klewel.com/graph/>.

- [7] Deursen, D van, Troncy, R, Mannens, E, Pfeiffer, S, Lafon, Y and Walle, R van de (2010). Implementing the Media Fragments URI Specification. WWW 2010 Developers Track, Raleigh (USA).
- [8] Sack, H and Waitelonis, J (2006). Integrating social tagging and document annotation for content-based search in multimedia data, in the 1st Semantic Authoring and Annotation Workshop (SAAW 2006), co-located with ISWC 2006, Athens (GA), USA.
- [9] Waitelonis, J and Sack, H (2009). Augmenting video search with Linked Open Data, in I-SEMANTICS 2009, Graz, Austria.
- [10] Auer, S, Bizer, C, Kobilarov, G, Lehmann, J, Cyganiak, R and Ives, Z (2007). DBPedia: A Nucleus for a Web of Open Data. In 6th International Semantic Web Conference (ISWC).