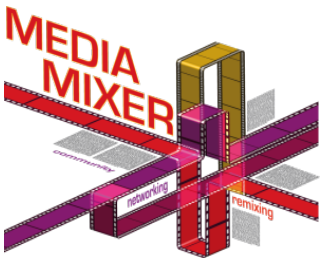
Institut "Jožef Stefan", Ljubljana, Slovenija

videolectures.net
exchange ideas & share knowledge



"VideoLecturesMashup"

Remixing Lecture Videos for Topical e-learning Across Video Repositories

A MediaMixer white paper
on media mixing technology for the e-learning community

Version 1.0	02 January 2014

Contact person:
Tanja Zdolšek (JSI)
tanja.zdolsek@ijs.si

Co-authors:
Ana Fabjan (JSI)
Peter Keše (Viidea)

Executive summary

VideoLectures.NET is an award-winning free and open access educational video lectures repository. The portal is aimed at promoting science, exchanging ideas and fostering knowledge sharing by providing high quality didactic contents not only to the scientific community but also to the general public.

Currently VideoLectures.NET hosts more than 16.000 video lectures from prominent universities and conferences. Most lectures are 1 to 1.5h long linked with slides and enriched with metadata and additional textual contents, which makes VideoLectures.NET a very efficient e-learning channel. On the other hand, visitors to VideoLectures.NET are looking to consume learning materials on specific topics of interest. However, typically visitors have limited time to find and watch the materials they want and the topics they search for may be of different parts of multiple learning resources rather than the subject of a specific complete learning resource. This problem can be solved with the MediaMixer technology. The integration into Videolectures.NET demonstrates the value of the technology while also itself benefitting from improved search and retrieval of fragments of its video assets. Users get an easier and quicker access to those different parts (fragments) in the form of a single, integrated presentation of learning materials. In turn this drives to more repeated access and the acquisition of new users, including in new contexts. E.g. dynamic provision of such learning resource “mash ups” would be particularly useful in mobile consumption contexts.

This paper presents the first use case demonstrator, based on a proof of concept realisation with MediaMixer technology. We present the demonstrator, outlining the path to its implementation and identifying the solution it offers to organisations in the e-learning domain. A link to an online UI or video of the demonstrator is given.

Table of Contents

Introduction

VideoLectures.NET is an award-winning free and open access educational video lectures repository. The lectures are given by distinguished scholars and scientists at the most important and prominent events like conferences, summer schools, workshops and science promotional events from many fields of science. The portal is aimed at promoting science, exchanging ideas and fostering knowledge sharing by providing high quality didactic contents not only to the scientific community but also to the general public. All lectures, accompanying documents, information and links are systematically selected and classified through the editorial process taking into account also users' comments.

Currently VideoLectures.NET 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.

VideoLectures.NET users are looking to consume learning materials on specific topics of interest. However, users 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). That is why the MediaMixer technology is needed as users would benefit from easier and quicker access to those different parts (fragments) 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, IPTV specialist channels).

This paper presents the first use case demonstrator, based on a proof of concept realisation with MediaMixer technology. We present the demonstrator, outlining the path to its implementation and identifying the solution it offers to organisations in the e-learning domain. A link to an online UI or video of the demonstrator is given.

VideoLecturesMashup - Remixing lecture videos for topical e-learning across video repositories

Summary of the use case

As use case partner, JSI integrated the MediaMixer technology into Videolectures.NET and hence demonstrated the value of the technology, while also itself benefitting from improved search and retrieval of fragments of its video assets. VideoLectures.NET users are looking to consume learning materials on specific topics of interest. Besides, they usually have limited time to find and watch the materials they want. Also, 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). Mediamixer technology allows users for easier and quicker access to those different parts (fragments). The **VideoLecturesMashup** is a dedicated channel on the VideoLectures.NET portal, which is capable of accepting a specific learning topic as input and produces as a result a mash up of fragments of learning materials from the site addressing that topic. These fragments are ordered in a meaningful way.

Identification of media content

A dataset from the VideoLectures.NET digital repository was used. It consists of ten videos, in particular five videos with slides and metadata and five videos with audio transcripts and metadata.

Architecture & implementation

Fragment retrieval and presentation was implemented as a pluggable extension to the existing VideoLectures.NET technological 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 player codebases were extended, to complement existing lecture presentation experience with fragment functionality. A separate XML document (besides the existing SMIL XML timeline) is thus formed and provided to both players to describe fragments, which need to be played and/or highlighted. The fragment description will ideally be provided from other Mediamixer online web-services. However in our case, where these services were not yet available, we provided a fallback mechanism to use fragments that we could directly extract from our current data, namely, slide synchronization timeline and subtitle information. For this reason, we collected and gathered both the slide timeline and subtitle information in the same data structure, which we then use for search and retrieval.

Media Fragment creation

Internally, the fragments were created based on slide synchronization timeline, where a view of existing data was used, and from available subtitle files which were parsed, analyzed and stored into the database to provide an option of full text search and fragment retrieval. The portal’s search interface was extended with fragment functionality and videos that are being displayed also present a list of fragments based on the specified search term.

Media Fragment annotation

Time-fragments are currently being annotated with the available textual information, be it from slide titles or subtitle captions. In the future, we will rely on other MediaMixer services / SPARQL to provide additional annotation.

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.

Media Fragment retrieval

Fragment search and retrieval engine is designed as a pluggable architecture. The initial implementation is based on a simple text-search from the fragment database. Once the SPARQL data is processed and made available for the VideoLectures.NET dataset, we will replace or complement the existing fragment retrieval engine with the external SPARQL one. In addition, we are planning to improve the quality, of the internal text search results, by replacing the current built-in SQL text search with an external SOLR engine.

Demonstrator UI

Assessing the online demo the user gets a search engine and examples for search key words (*example*, *learn* or *structure*). For instance, when entering search key word ‘learn’ the user gets seven matches, referring to the videos as a whole.

First, we have one video indicated as an invited talk. On the left side of the page, users see metadata (name of the lecturer, title of the lecture, short abstract).

invited talk:
Joshua B. Tenenbaum: How to Grow a Mind: Statistics, Structure and Abstraction
The fields of cognitive science and artificial intelligence grew up together, with the twin goals of understanding human minds and making machines smarter in more humanlike ways. Yet since the 1980s they have mostly grown apart, as cognitive scientists came ...

On the right side, we have fragments listed – fragments within the stated video referring to the search key word ‘learn’.

- 6:26 - Learning from very few examples
- 36:37 - Learning from very few examples (1)
- 37:02 - Learning from very few examples (2)
- 37:15 - Learning from very few examples (3)
- 38:16 - Learning from very few examples (4)

Second set of results (6 matches) refers to video lectures. The layout is the same as mentioned by the invited talk video lecture.

lecture:

Luis von Ahn: Duolingo: Translating the Web with Millions of People
I want to translate the web into every major language: every web page, every video, and, yes, even Justin Bieber's tweets. With its content split up into hundreds of languages — and with over 50 percent of it in English ...

- 16:07 - Learning a New Language
- 21:51 - Learning With Real Content

Catherine L. Drennan: Lecture 1: The importance of chemical principles

- 12:47 - So, you'll be learning in this class a great preparation
- 16:19 - yourself in what you really enjoy learning about.
- 19:45 - Did you guys have to learn the common ions?
- 20:00 - I didn't learn them, and that was really bad because it kept
- 22:27 - taken was because I would learn this new principal in chemistry

Eric Grimson: Lecture 1: Goals of the course; what is computation; introduction to data types, operators, and variables

- 1:10 - to help everybody learn about computation, and that's
- 1:41 - We're going to try and help you learn how to think like a
- 5:37 - It's not bad to, if you like, learn from the skills of others
- 10:35 - them to help you learn.
- 11:11 - helping you learn the material.

John Guttag: Lecture 24: Course overview; what do computer scientists do?

- 0:34 - people do once they learn about computer science?
- 3:03 - By now you've all learned that these experiments rarely are
- 8:02 - As part of this thinking recursively we learn about
- 8:05 - reduction, we learned how to, say, reduce the problem of
- 8:48 - scientists learn, is how to figure out what's relevant and

Gunnar Rätsch: Introduction to bioinformatics
I will start by giving a general introduction into Bioinformatics, including basic biology, typical data types (sequences, structures, expression data and networks) and established analysis tasks. In the second part, I will discuss the problem of predictive sequence analysis with ...

- 0:00 - Machine Learning in Bioinformatics
- 43:07 - Challenges for Machine Learning

John Wargo: Lecture 1 - Course Overview: Science and Law
Professor John Wargo introduces the central question of the course, "Can law shape a sustainable future for ten billion people?" The purpose of the course is to examine the most important U.S. laws adopted over the past forty years, and ...

- 17:57 - different cases about what we've learned.
- 31:38 - food I think is something to take note of and to learn from.

Figure 1 shows the results of the search key word ‘learn’ as they appear in the online demo.

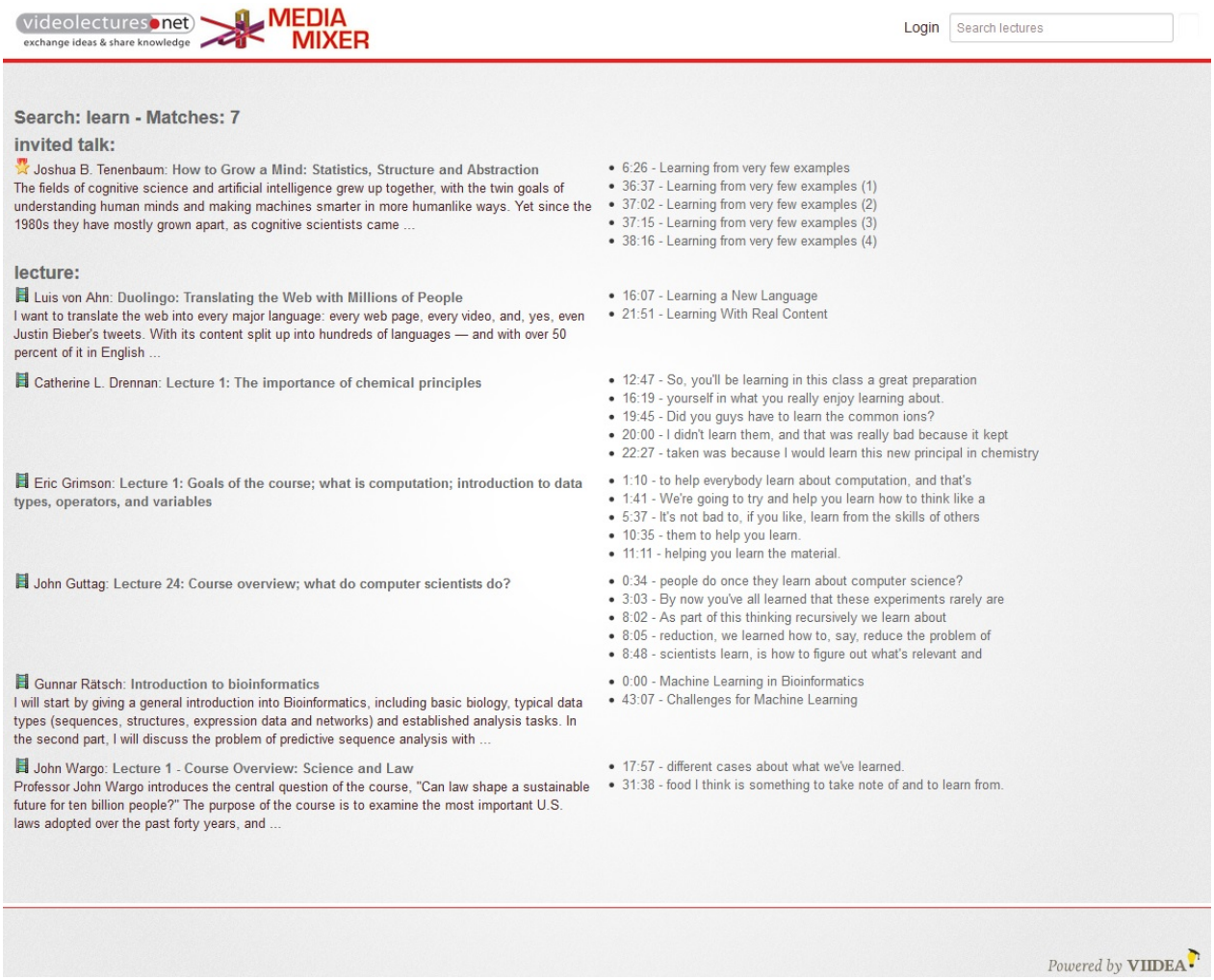


Figure 1: Results for the search key word 'learn'

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 2, 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 fragments with timing).

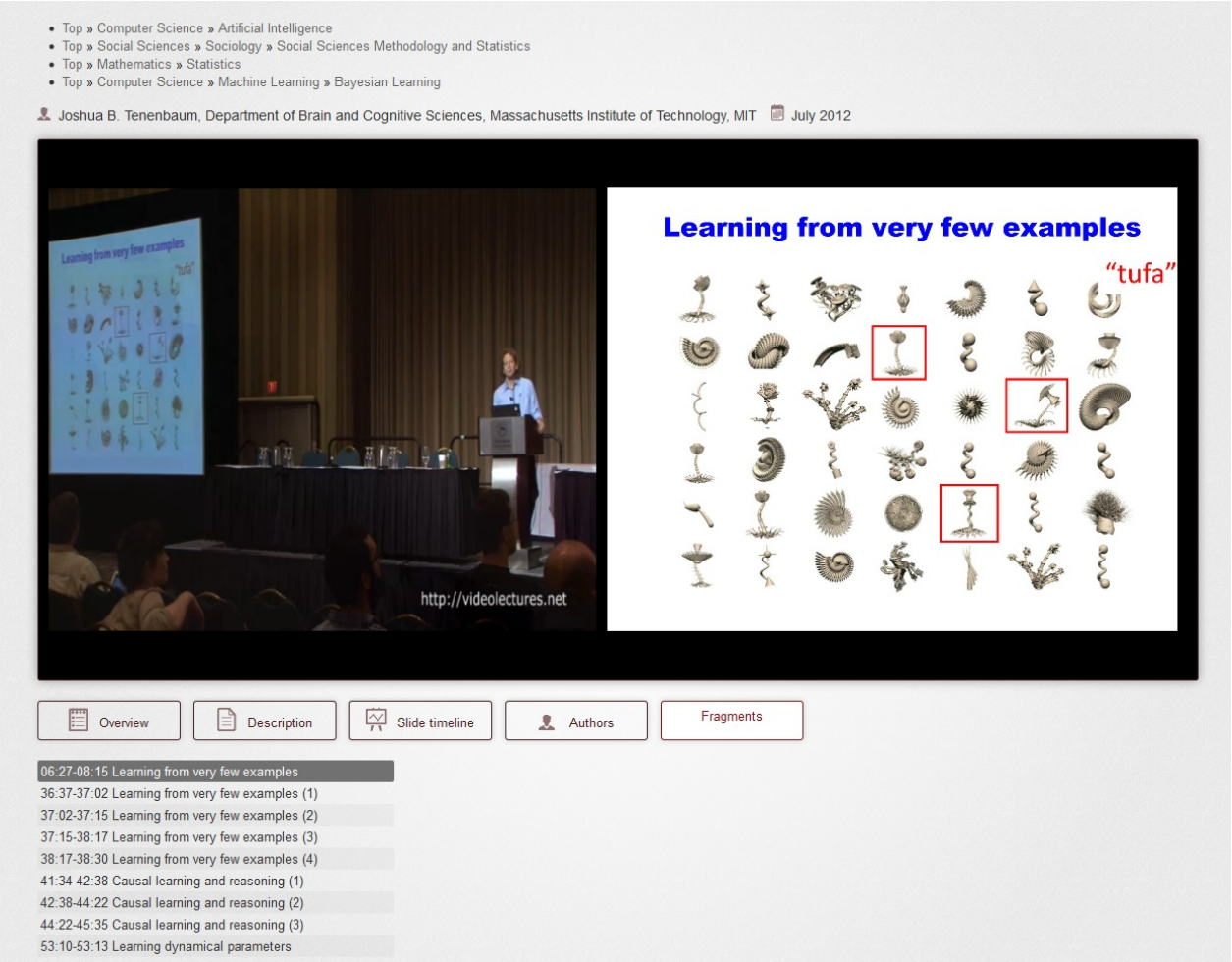


Figure 2: Layout of the demo features

Contact

Tanja Zdolsek, Jožef Stefan Institute, tanja.zdolsek@ijs.si
Gaber Cerle, Jožef Stefan Institute, gaber.cerle@ijs.si

Jožef Stefan Institute
Jamova cesta 39, Ljubljana, Slovenia

Links

- Video: available at **http://bit.ly/videolecturesmashup**
- Online demo: **http://mediamixer.videolectures.net/**
- For more information about mediamixing technology, see **http://community.mediamixer.eu**