

# Visual and Statistical Analysis of VideoLectures.NET

**Erik Novak**, Inna Novalija

Artificial Intelligence Laboratory

Jožef Stefan Institute

---

Si-KDD, October 10, 2016

# VideoLectures

## Goals:

- Analyse of viewer behaviour
- Detect which lectures are interesting for the users
- Detect similarities between lectures

videolectures.net  
exchange ideas & share knowledge

World Summit Award  
United Nations Educational, Scientific and Cultural Organization  
Institut "Jožef Stefan" Ljubljana, Slovenia  
UNESCO Chair on Open Technologies for Open Educational Resources and Open Learning

Welcome Erik Trac (my) | [SIGN OUT](#)

1012 events, 14098 authors, 18956 lectures, 21985 videos

HOME • BROWSE LECTURES • PEOPLE • CONFERENCES • ACADEMIC ORGANISATIONS • EU SUPPORTED • BLOG • ABOUT US

SEARCH

**WATCHING NOW**

- Znanstveni slam
- Deep Learning Summer School
- Gaussian Process Basics
- Tutorial on Neural Network Optimization
- Distributed, Real-Time Bayesian Learning in

**SHOW CHAT**

**CATEGORIES**

- Architecture (247)
- Arts (394)
- Astronomy (75)
- Biology (556)
- Business (1212)
- Chemistry (213)
- Computers (522)
- Computer Science (12591)
- Earth Sciences (28)
- Environment (191)
- Events (436)
- Health Sciences (33)
- Humanities (546)
- Life Sciences (33)
- Mathematics (636)
- Medicine (473)
- Military (36)
- Philosophy (170)
- Physics (376)
- Regional (13)
- Science (650)
- Social Sciences (2211)
- Sports (39)
- Technology (934)

**NEWS**

**KDD 2016 update: Applied Data Science and Research Sessions Talks Online Now** Sept. 27, 2016

Be sure to check them out here.

**Public Lecture on Gravitational Waves: A New Astronomy** Sept. 21, 2016

In her lecture Sheila Rowan, a professor in the School of Physics and Astronomy at University of Glasgow, explains the nature of gravitational waves, describe what sources out in the Universe can produce them, explain how they are detected and what the future of this new era in astronomy might look like.

**NEW! International Society for Bayesian Analysis (ISBA) 2016 videos** Sept. 5, 2016

ISBA 2016 World Meeting is a unique event where the Bayesian community gathers together to discuss recent advances and the future of our profession, at the same time looking back to our roots and traditions, following the footsteps of those who laid the foundations of where we are now.

**First KDD 2016 videos are here!** Sept. 1, 2016

We are very excited to announce that KDD 2016 keynote talks and plenary panel videos are online now! Videolectures.Net will host the whole conference, so stay tuned for more!

**Fresh from Montréal: Deep Learning Summer School 2016 online now!** Aug. 24, 2016

We start the day with another great addition to our video collection, go and check it out at Deep Learning Summer School.

**ESWC 2016 videos online!** Aug. 23, 2016

After a summer break Videolectures.Net team is back with some exciting new content. First on our recently published list are the 13th Extended Semantic Web Conference (ESWC) videos. Enjoy the knowledge and stay tuned for more!

**NEWSLETTER**

Subscribe to our newsletter to receive digest of activity

Your Email:  [Subscribe](#)

**RECENT EVENTS** [MORE](#)

**IHPCSS 2016 - Ljubljana**

International HPC Summer School (IHPCSS), Ljubljana 2016

The seventh International HPC summer school 2016 is a unique event where leading European, Canadian, Japanese and American computational scientists ...

**SSJLK 2016 - Ljubljana**

Drugačnost v slovenskem jeziku, literaturi in kulturi

52. seminar slovenskega jezika, literature in kulture: Drugačnost v slovenskem jeziku, literaturi in kulturi, Ljubljana 2016

Kdo je drugačen in kdo je tisti, ki mu drugačnost pripisuje? Ali je biti drugačen pozitivno ali negativno poslanstvo? Zakaj ...

**KDD 2016 - San Francisco**

22nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), San Francisco 2016

KDD 2016, a premier interdisciplinary conference, brings together researchers and practitioners from data science, data mining, knowledge discovery, large-scale data ...

**Deep Learning Summer School 2016**

# Outline

---

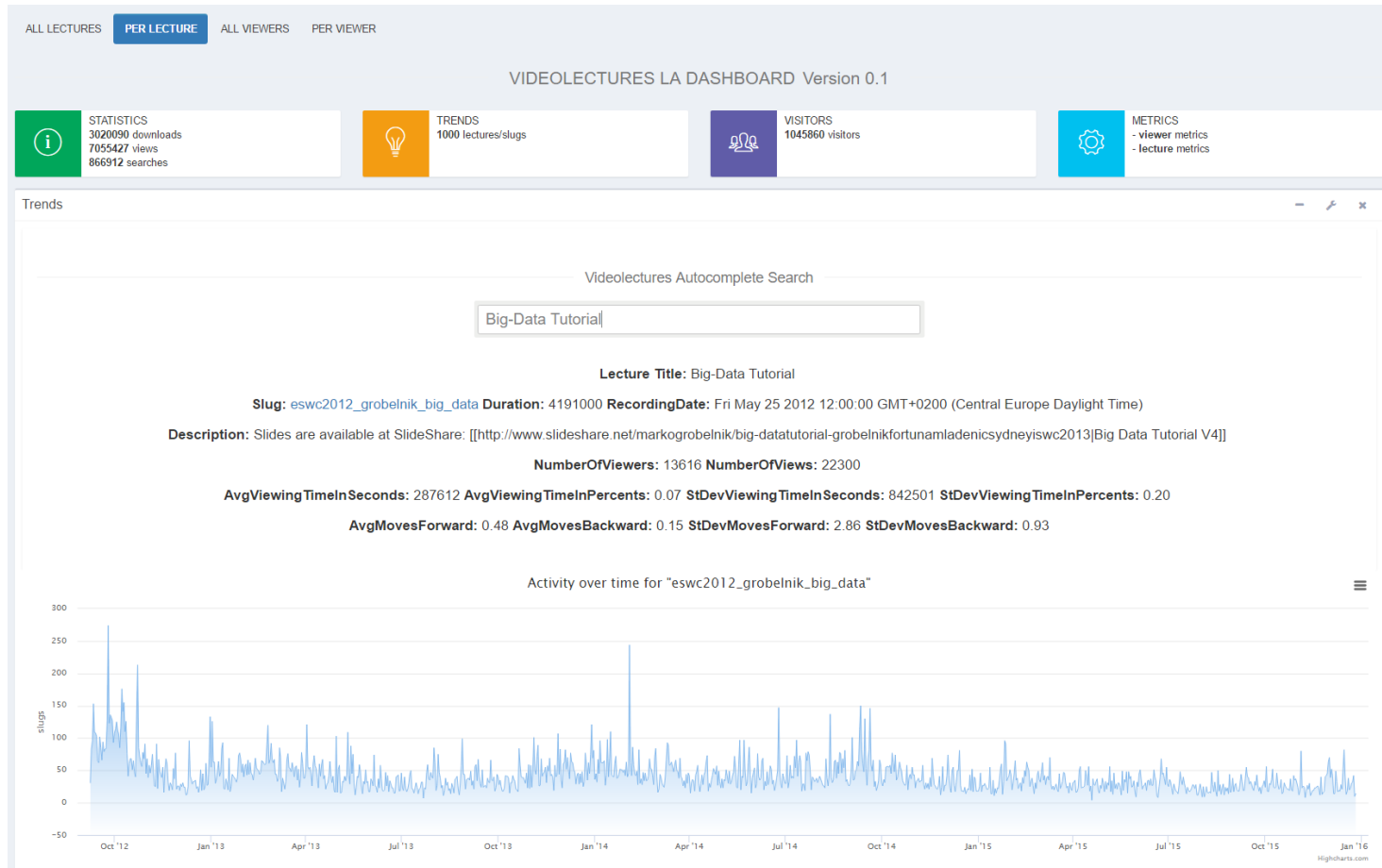
## Learning Analytics Dashboard

- Data processing
- Data analysis

## VideoLectures Explorer

- Methodology
- Tool components

# Learning Analytics Dashboard



# Learning Analytics Dashboard

Data processing

---

11.3GB of log files from VideoLectures portal

- Containing raw events from September 2012 until December 2015
- Log files contain: id, timestamp, session, log entry, lecture information etc.

We established main event types: *view*, *download* and *search*

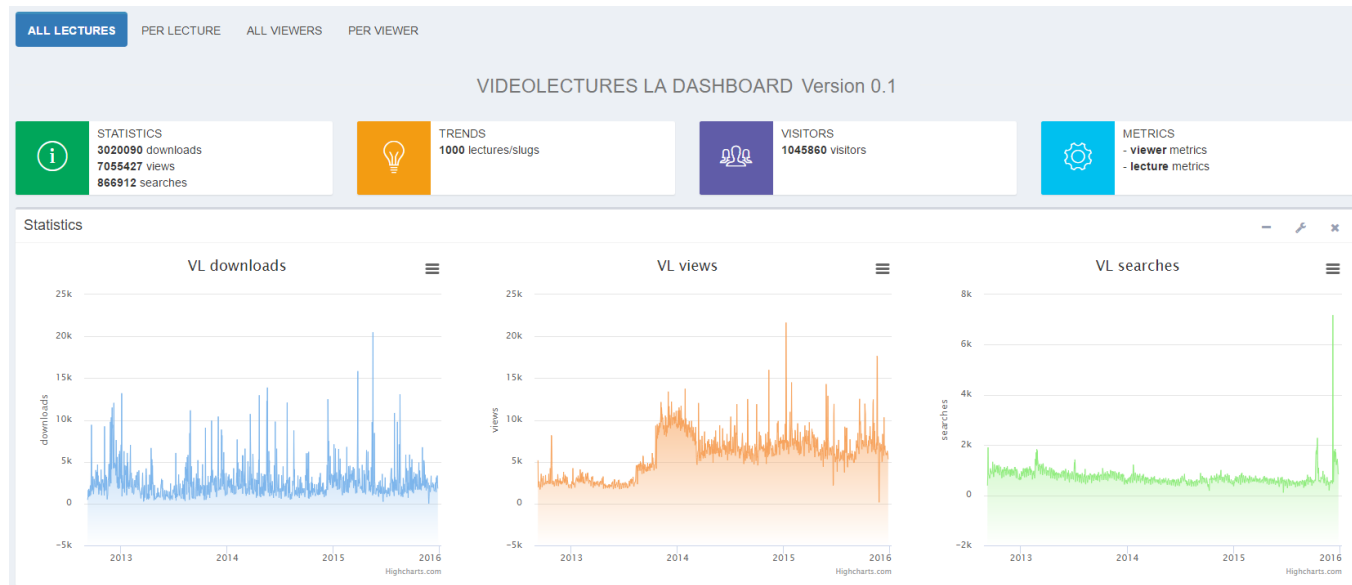
Collection of log files of the viewer behaviour

# Learning Analytics Dashboard

Analysis of log files

Four perspectives:

- aggregated perspective for **all lectures**
- perspective of a **single lecture**
- aggregated perspective of **all viewers**
- perspective of a **single viewer**



# Learning Analytics Dashboard

Analysis of log files

## Lecture metrics:

Number of Viewers

Number of Views

Avg Moves Forward

Avg Moves Backward

Avg Time Spent

Avg Time in %

Std Dev Time Spent

Std Dev Time in %

Std Dev Moves Forward

Std Dev Moves Backward

Slug	Title	Number of Viewers	Number of Views	Avg Moves Forward	Avg Moves Backward	Avg Time Spent	Avg Time in %	St Dev Time Spent	St Dev Time in %	St Dev Moves Forward	St Dev Moves Backward
11nanodan2014_ljubljana		1457	4049	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
aaai2011_visser_humanrobot	The Bomb Squad: Securing a checkpoint with a Human-Robot Team	50	42	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
aaai2012_tenenbaum_grow_mind	How to Grow a Mind: Statistics, Structure and Abstraction	5475	12614	0.25	0.09	280037.32	0.07	863750.26	0.21	1.45	0.71
aaai2012_toronto		3359	11070	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
aaai2013_mooney_language_learning	Grounded Language Learning	1980	3019	0.28	0.17	405789.19	0.12	818533.66	0.25	1.14	0.78
aaai2013_washington		2147	7105	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
aaai_video_competition		1165	4167	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
academia		24309	89315	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
acai05_berthold_fl	Fuzzy Logic	17854	24501	0.57	0.28	278588.00	0.08	731806.50	0.22	2.69	1.89
act2013_sofia		1814	7191	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Similar metric for **viewers**.

# Learning Analytics Dashboard

Analysis of log files

## Lecture metrics:

Number of Viewers

Number of Views

Slug

Title

Slug	Title	Number of Viewers	Number of Views	Avg Moves Forward	Avg Moves Backward	Avg Time Spent	Avg Time in %	St Dev Time Spent	St Dev Time in %	St Dev Moves Forward	St Dev Moves Backward
11nanodan2014_ljubljana		1457	4049	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
aaai2011_visser_humanrobot	The Bomb Squad: Securing a checkpoint with a Human-Robot Team	50	42	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
aaai2012_tenenbaum_grow_mind	How to Grow a Mind: Statistics, Structure and Abstraction	5475	12614	0.25	0.09	280037.32	0.07	863750.26	0.21	1.45	0.71
academia		24309	89315	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
acai05_berthold_fl	Fuzzy Logic	17854	24501	0.57	0.28	278588.00	0.08	731806.50	0.22	2.69	1.89
acit2013_sofia		1814	7191	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Std Dev Time Spent

Std Dev Time in %

Std Dev Moves Forward

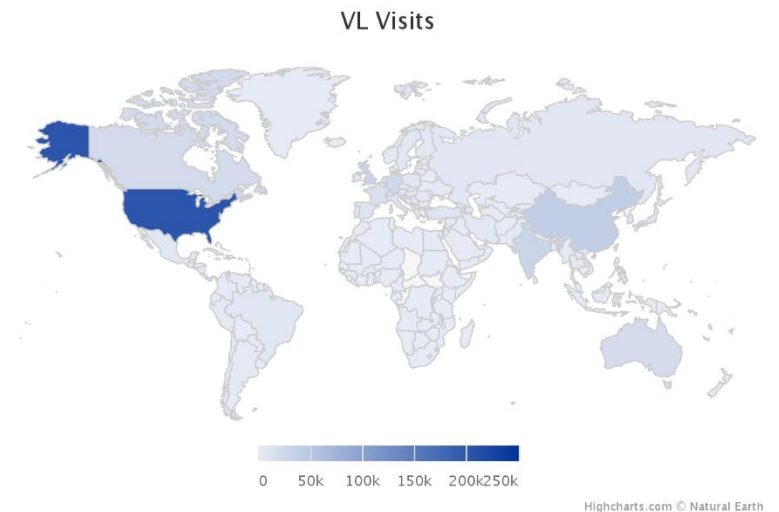
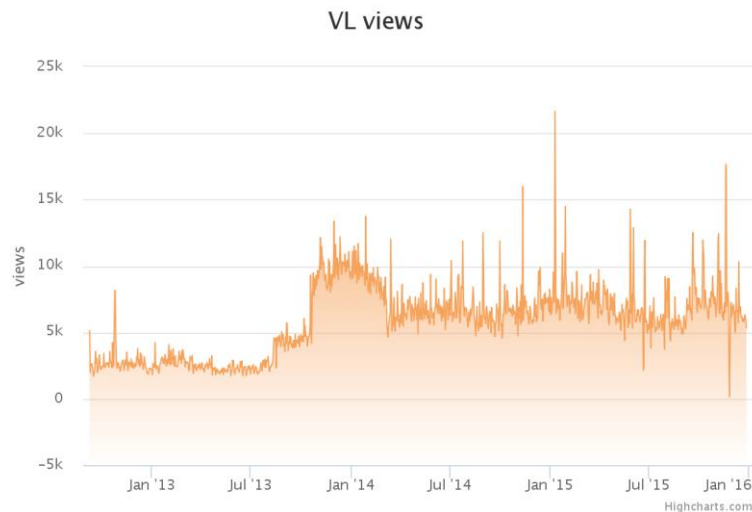
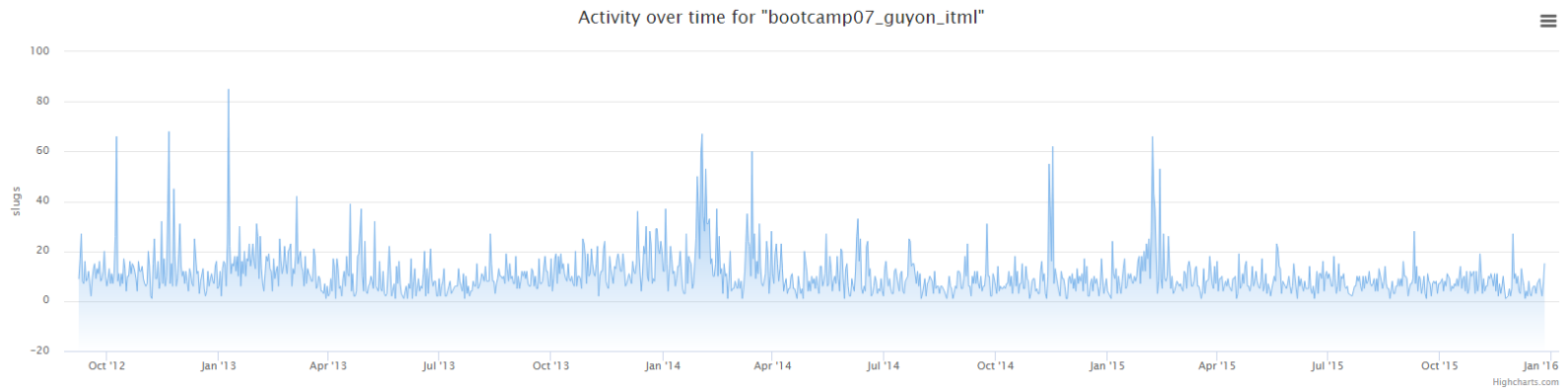
Std Dev Moves Backward

Similar metric for **viewers**.



# Learning Analytics Dashboard

## Statistical components



- ✓ Analyse of viewer behaviour
  - ✓ Detect which lectures are interesting for the users
- 
- Detect similarities between lectures

# VideoLectures Explorer

Search presenter, category or organization

Search
⌵

---

### Advanced Options

Category:	<input type="text" value="Type a category"/>	Organization:	<input type="text" value="Type an organization"/>
Presenter:	<input type="text" value="Type a presenter"/>	Country:	<input type="text" value="Type a country"/>
City:	<input type="text" value="Type a city"/>	Language:	<div>All ⌵</div>
# of views:	<input type="text" value="1000"/> max		

#### Searchwords

Presenter:  
Organization:  
Categories: Machine Learning  
City:  
Country:  
# of views: min: 1000  
Language:

#### Information

Number of found lectures: 464  
Total number of views: 1738084  
**Categories with number of occurrences:**  
Computer Science(464), Machine Learning(464), Bayesian Learning(66), Kernel Methods(49), Statistical Learning(32), Pattern Recognition(30), Reinforcement Learning(29), Clustering(24), Mathematics(22), Graphical Models(22), Neural Networks(20), Information Theory(18), Gaussian Processes(17), Deep Learning(15), Support Vector Machines(14), Computational Learning Theory(13), Data Mining(13), Unsupervised Learning(13), Markov Processes(11), Statistics(11), Principal Component Analysis(10), Structured Data(9), On-line Learning(9), Human Language Technology(8), Social Sciences(8), Monte Carlo Methods(7), Computer Vision(7), Dirichlet Processes(7), Semi-supervised Learning(6).

#### Lecture Information

Title: Support Vector Machines  
Presenter: Chih-Jen Lin  
Public: Yes  
Enabled: Yes  
Link: mss06tw\_lin\_svm

**Lecture title:** Machine Learning, Probability and Graphical Models  
**Presenter:** Sam Roweis  
**Organization:** N/A  
 The lecture is in English. The main categories of the lecture were **Computer Science, Machine Learning, Graphical Models**. It was published in M. 24.07.2006 and it's duration is 04:59:20. There have been **42947** views until 24.07.2016.

# VideoLectures Explorer

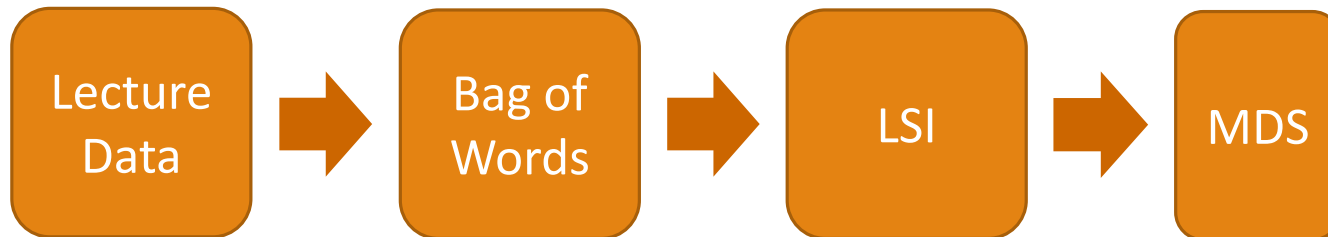
## Methodology

---

Approx. 23k lectures, keynotes, interviews and other events

- Containing lecture title, description, categories, number of views etc.

Drawing lectures into a two-dimensional space[1]



[1] B. Fortuna, D. Mladenić and M. Grobelnik. Visualization of temporal semantic spaces. In *Semantic knowledge management*, pages 155-169, Springer, 2009.

## Tool components

**Presenter:**  
**Organization:**  
**Categories:** Machine Learning  
**City:**  
**Country:**  
**# of views:**  
**Language:**

Number of found lectures: 3754  
Total number of views: 2470761  
Categories with number of occurrences:

**Lecture title:** Support Vector Machines  
**Presenter:** Chih-Jen Lin  
**Organization:** Department of Computer Science and Information Engineering  
**Description:** Support vector machines (SVM) and kernel methods are important machine learning techniques. In this short course, we will introduce their basic concepts. We then focus on the training and optimization procedures of SVM. Examples demonstrating the practical use of SVM will also be discussed. Basically we focus on classification....  
The lecture is in English. The main categories of the lecture were **Computer Science, Machine Learning, Kernel Methods, Support Vector Machines**. It was published in 27.07.2006 and it's duration is 03:54:31. There have been **33998** views until 24.07.2016.



# Conclusion

---

## Learning Analytics Dashboard

- Available: <http://learninganalytics.videolectures.net/>
- Analysis of viewer behaviour
- Detection of lectures interesting for the users

## VideoLectures Explorer

- Available: <http://explore.videolectures.net/>
- Exploratory tool for VideoLectures.NET
- Similarities between lectures

# Thank you for listening!

---

Any questions?