



transLectures

Transcription and Translation of Video Lectures



VideoLectures.NET & transLectures

Matjaž Rihtar
Josef Stefan Institute, Ljubljana



VideoLectures.NET

- Award-winning video lecture repository:
 - 16000 lectures (18500 videos)
 - Mostly in English (14000) & Slovenian (1600)
 - 12100 authors
 - 900 events
 - 13.4 million lecture views (from Sep 2013 to Mar 2014, 171 days)
- Objective:
 - Integrate results from **transLectures** project into VideoLectures.NET
 - Real-time serving of transcriptions/translations via **transLectures** Web service



Integration into VideoLectures.NET

- VideoLectures.NET before **transLectures**
 - 170+ videos with subtitles
- VideoLectures.NET after **transLectures**
 - 10067 new transcriptions
 - 9265 English
 - 802 Slovene
 - Transcribed 6558 hours 40 minutes of videos
 - 36367 new translations
 - 9632 English – German
 - 9632 English – Slovene
 - 9269 English – Spanish
 - 6987 English – French
 - 847 Slovene – English



Integration into VideoLectures.NET

- **transLectures** Web Service
 - Transcriptions and translations stored on separate server
 - New player on VideoLectures.NET connects to Web Service to get transcriptions & translations
 - Transcriptions & translations are stored in new DFXP format (XML)



Integration into VideoLectures.NET

- DFXP format
 - MPEG-4 Timed Text (TTML/DFXP) standard
<http://www.w3.org/TR/ttaf1-dfxp/>
 - Extended with our tags:
 - Confidence levels
 - Alternatives
 - Author (machine/human)
- New Player in use on
<http://translectures.videolectures.net>
- List of all currently transcribed videos
http://helium.ijs.si/lectures_with_trans.html



New Player on VideoLectures.NET

Switch off the lights

The screenshot shows a video player interface. On the left, a speaker is visible at a podium with a microphone. The podium has a sign that reads "Sheraton La Cote". A large play button is overlaid on the video. On the right, a slide is displayed with the following text:

Enhancing Workflow with a Semantic Description of Scientific Intent

Edoardo Pignotti, Peter Edwards, Alun Preece, Nick Gotts and Gary Polhill

Computing Science
University of Aberdeen

The Macaulay Institute, Aberdeen




Below the slide, there is a list of logos for various institutions and projects, including "The University of Aberdeen", "The Macaulay Institute", "Grid", and "M".

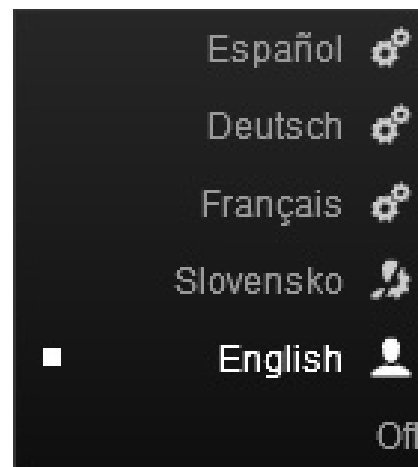
At the bottom of the video player, there is a progress bar showing "00:07 / 24:33" and a volume icon. A subtitle bar at the bottom of the video frame contains the text: "project from the University of Aberdeen and especially associates as we".

On the right side of the video player, there is a language selection menu with the following options: Español, Deutsch, Français, Slovensko, English, and Off. Each option has a gear icon next to it.



Identification of subtitles quality

- Using icons on CC pop-up window
 -  Completely human
 -  Mixed (partially supervised)
 -  Completely automatic
- Selection of languages





Identification of subtitles quality

- Using different colors of subtitles

Switch off the lights

Modeling Real Graphs using
Kronecker Multiplication

Jure Leskovec, Christos Faloutsos
Machine Learning Department

 **Carnegie Mellon**

parameters we want to estimate this parameters to generate the graph of let's take 100

00:19 / 26:10 1 / 30



User Evaluations

- Using **transLectures** Player
 - Separate site <http://helium.ijs.si/>
 - Students from Faculty of foreign languages
 - Several runs with few videos
 - With machine transcription (50 videos)
 - Without machine transcription (from scratch, 10 videos)
 - Collected feedback



http://helium.ijs.si/player/main.php

transLectures video player







File Edit View Favorites Tools Help

eBooks & eLearning Avax... [AVD] Emulator Skins of p... FileSnail - File search engi...

transLectures))) MAIN LOGOUT

COOPERATION cip

videolectures.net
exchange ideas & share knowledge

 <p>umetnina_bezlaj_poljub lang sl</p>	 <p>umetnina_zerovc_sskl lang sl</p>	 <p>zi09_dermelj_pzjn lang sl</p>
 <p>znanje_zagmajster_nrvp lang sl</p>	 <p>zv_ung_kolar_opkd lang sl</p>	 <p>ssjlk2011_smole_krompir lang sl</p>

trans Lectures)))

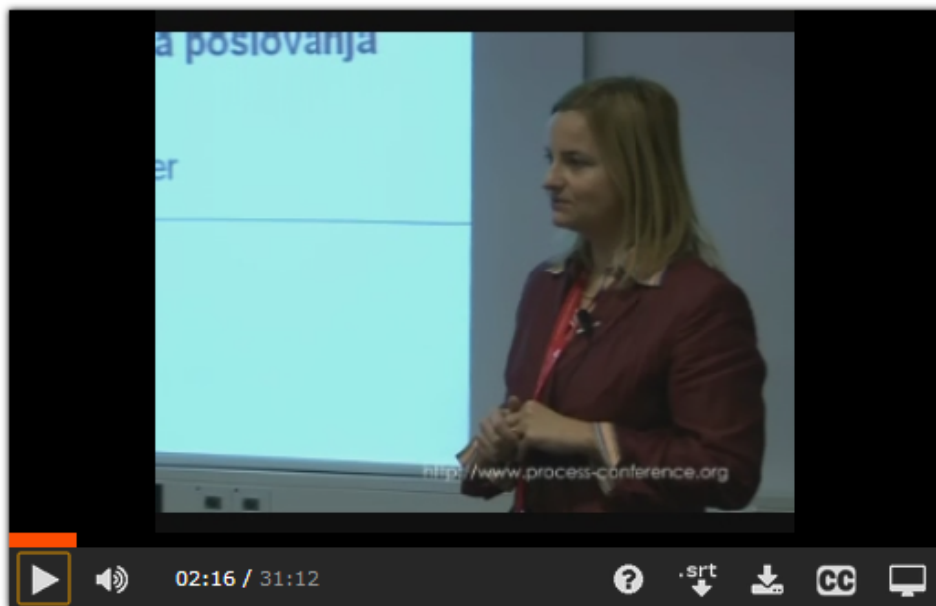
- Please **do not correct capital letters and punctuation marks and transcribe the words spoken literally.**
- You can supervise one video in multiple sessions, but please do a single session without any interruptions.
- Be sure to **click on the "Save Changes" button before leaving the editor, at the end of your work session.**
- You can save your work several times on the same session to avoid losing it (click on the

100%



User Evaluations

Editing video: mpp06_stemberger_puup (lang SL)



govorimo to na podlagi nekih
analiz ne pa samo tko na pamet in na podlagi posameznih primerov no
to je eden izmed namenov, drug namen pa je, da pravzaprav
poznamo tudi neko metodologijo, s katero lahko podjetje samo zase ugotovi, kje se pravzaprav trenutno
nahaja in potem na podlagi teh ugotovitev se lahko potem odloča, kako naprej
za zato ker podjetje, ki se recimo sploh še ne zaveda svojih procesov takšno podjetje prav gotovo
ne bo moglo postat čez noč procesno usmerjeno, ane, to je nekako dolgotrajen proces



Future Plans

- Transcribe all videos
 - A lot of computing power needed
- Include additional languages in translations
 - Other European languages
 - Chinese



transLectures

Transcription and Translation of Video Lectures



Thanks!

- My E-mail:
Matjaz Rihtar <matjaz.rihtar@ijs.si>
- Project coordinator:
Alfons Juan-Císcar <ajuan@dsic.upv.es>