FP7-ICT-2011-7 - Language technologies (STREP)  
Duration: 36 months (1 Nov 2011 – 31 Oct 2014)

<table>
<thead>
<tr>
<th>No</th>
<th>Name</th>
<th>Short</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>UNIVERSITAT POLITECNICA DE VALENCIA</td>
<td>UPVLC</td>
<td>Spain</td>
</tr>
<tr>
<td>2</td>
<td>XEROX SAS</td>
<td>XEROX</td>
<td>France</td>
</tr>
<tr>
<td>3</td>
<td>INSTITUT JOZEF STEFAN</td>
<td>JSI</td>
<td>Slovenia</td>
</tr>
<tr>
<td>3+</td>
<td>KNOWLEDGE FOR ALL FOUNDATION</td>
<td>K4A</td>
<td>UK</td>
</tr>
<tr>
<td>4</td>
<td>RHEINISCH-WESTFAELISCHE TECHNISCHE HOCHSCHULE AACHEN</td>
<td>RWTH</td>
<td>Germany</td>
</tr>
<tr>
<td>5</td>
<td>EUROPEAN MEDIA LABORATORY GMBH</td>
<td>EML</td>
<td>Germany</td>
</tr>
<tr>
<td>6</td>
<td>Deluxe Digital Studios Ltd</td>
<td>DDS</td>
<td>UK</td>
</tr>
</tbody>
</table>
Outline of the talk

• Motivation
• Goal and technology challenges
• Demo videos: project overview, intelligent interaction and Matterhorn integration
• Transcriptions and translations at the pilots
• Massive adaptation
• Intelligent interaction with users
• Integration and evaluation
Motivation

- Large on-line repositories of video lectures are being established: VideoLectures.NET, poliMedia, OCW and Opencast Matterhorn-based repositories, etc.
- Video lectures are neither transcribed nor translated due the lack of cost-effective tools.
- Transcriptions and translations are needed to make them accessible to speakers of different languages and people with disabilities. They would also facilitate further search and analysis functions.
- Starting hypothesis: current ASR and MT techniques are not far from achieving acceptable results in educational repositories.
Goal and technology challenges

Goal:
To develop innovative, cost-effective solutions to produce accurate transcriptions and translations in VideoLectures.NET, with generality across other Matterhorn-related repositories.

Technology challenges:
• Improvement of transcription and translation quality by massive adaptation.
• Improvement of transcription and translation quality by intelligent interaction.
• Integration into Matterhorn to enable real-life evaluation.
Demo videos

• Project overview
https://www.translectures.eu/project-summary

• Intelligent interaction with users
https://www.translectures.eu/web/progress

• Matterhorn integration
https://www.translectures.eu/web/progress
## Transcriptions and translations at the pilots

<table>
<thead>
<tr>
<th>Language</th>
<th>Nb. of lectures</th>
<th>Nb. of hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>9148</td>
<td>5900</td>
</tr>
<tr>
<td>Slovenian</td>
<td>741</td>
<td>490</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>9889</strong></td>
<td><strong>6390</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Language</th>
<th>Nb. of lectures</th>
<th>Nb. of hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish</td>
<td>7554</td>
<td>996</td>
</tr>
<tr>
<td>English</td>
<td>420</td>
<td>62</td>
</tr>
<tr>
<td>Catalan</td>
<td>91</td>
<td>11</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>8065</strong></td>
<td><strong>1069</strong></td>
</tr>
</tbody>
</table>

**VideoLectures.NET:** EnEs, EnFr, EnDe, EnSl, SlEn

**poliMedia:** (live) EsEn
Massive adaptation

- Massive adaptation of acoustic models
- Massive adaptation of language models
- Massive adaptation of translation models
Intelligent interaction with users

• Fast constrained search

• Intelligent interaction for transcription

• Intelligent interaction for translation

• Incremental training for translation
Transcription of 23 videos (3h) with no interaction
Transcription of 23 videos (3h) with no interaction

Supervised minutes
WER
18.8
Transcription of 23 videos (3h) with **batch interaction**

![Diagram showing the transcription process with 23 videos, each labeled with a CC symbol, and a chart indicating a WER of 18.4% with BI.]
Transcription of 23 videos (3h) with batch interaction

Supervised minutes
WER
18.0 BI
Transcription of 23 videos (3h) with **batch interaction**

Supervised minutes
WER
17.6 BI
Transcription of 23 videos (3h) with **batch interaction**

![Graph showing WER improvement over time](image)
Transcription of 23 videos (3h) with batch interaction

Supervised minutes
WER
16.8 BI
Transcription of 23 videos (3h) with batch interaction

Supervised minutes
WER
16.8 BI
Transcription of 23 videos (3h) with batch interaction

Supervised minutes
WER
16.5 BI
Transcription of 23 videos (3h) with intelligent interaction

Supervised minutes
WER
16.5 BI
18.5 II
Transcription of 23 videos (3h) with intelligent interaction

Supervised minutes
WER
16.5 BI
18.1 II
Transcription of 23 videos (3h) with intelligent interaction

Supervised minutes
WER
16.5 BI
18.0 II
Transcription of 23 videos (3h) with intelligent interaction

Supervised minutes
WER
16.5 BI
17.5 II
Transcription of 23 videos (3h) with **intelligent interaction**

![Diagram showing transcription progress and WER improvement over time.](Image)
Transcription of 23 videos (3h) with intelligent interaction

Supervised minutes
WER
16.5 BI
17.3 II
Transcription of 23 videos (3h) with intelligent interaction

Supervised minutes
WER
16.5 BI
17.2 II
Transcription of 23 videos (3h) with intelligent interaction

Supervised minutes
WER
16.5 BI
17.0 II
Transcription of 23 videos (3h) with intelligent interaction

Supervised minutes
WER
16.5 BI
16.5 II
Transcription of 23 videos (3h) with intelligent interaction

Supervised minutes
WER
16.5 BI
16.4 II
Transcription of 23 videos (3h) with intelligent interaction
Transcription of 23 videos (3h) with **intelligent interaction**

![Graph showing WER and supervised minutes]

- WER: 16.5 BI
- 15.7 II

**Supervised minutes**
Transcription of 23 videos (3h) with intelligent interaction

Supervised minutes
WER

16.5 BI
15.7 II
Transcription of 23 videos (3h) with intelligent interaction

Supervised minutes
WER
16.5 BI
15.6 II
Transcription of 23 videos (3h) with intelligent interaction

Supervised minutes
WER
16.5 BI
15.4 II

Graph showing the decrease in WER with increasing supervised minutes.
Transcription of 23 videos (3h) with **intelligent interaction**

Supervised minutes
WER
16.5 BI
15.2 II
Transcription of 23 videos (3h) with **intelligent interaction**

[Graph showing WER improvement over supervised minutes]

Supervised minutes
WER
16.5 BI
15.1 II
Transcription of 23 videos (3h) with intelligent interaction

Supervised minutes
WER
16.5 BI
14.5 II
Transcription of 23 videos (3h) with intelligent interaction
Transcription of 23 videos (3h) with **intelligent interaction**

Supervised minutes
WER
16.5 BI
14.3 II
Transcription of 23 videos (3h) with **intelligent interaction**

![Diagram showing transcription process with WER metrics](chart.png)
Transcription of 23 videos (3h) with intelligent interaction

- Supervised minutes
- WER
  - BI 16.5
  - II 14.0
Transcription of 23 videos (3h) with intelligent interaction

Supervised minutes
WER
16.5 BI
13.6 II
Transcription of 23 videos (3h) with **intelligent interaction**
Transcription of 23 videos (3h) with **intelligent interaction**

![Diagram showing transcription process and WER improvement over supervised minutes.](attachment:image.png)

- Supervised minutes: 
  - WER: 16.5 BI
  - WER: 13.4 II

---

**Note:**

- Each video is transcribed with **CC** indicating the transcription status.
- The graph illustrates the decrease in WER over supervised minutes.
Transcription of 23 videos (3h) with intelligent interaction

Supervised minutes
WER
16.5 BI
13.1 II
Transcription of 23 videos (3h) with **intelligent interaction**

![Diagram showing transcription process]

Supervised minutes

WER

16.5 BI

13.0 II
Transcription of 23 videos (3h) with intelligent interaction

Supervised minutes
WER
16.5 BI
12.6 II
Transcription of 23 videos (3h) with intelligent interaction

Supervised minutes
WER
16.5 BI
12.6 II
Transcription of 23 videos (3h) with intelligent interaction

Supervised minutes
WER

16.5 BI
12.4 II
Integration and evaluation

• Integration and evaluation at the pilots

• Try our tools and transcribe your videos

www.translectures.eu

Free service opened by UPV on 13 March 2014