

**EUROPEAN COMMISSION
INFORMATION SOCIETY AND MEDIA DIRECTORATE-GENERAL**

Information and Communication Technologies

Collaborative Project



Transcription and Translation of Video Lectures

Grant Agreement Number 287755

Concept

Online educational repositories of video lectures are rapidly growing on the basis of increasingly available and standardised infrastructure.

A well-know example is the **VideoLectures** web portal, a free and open access educational video lectures repository, and a major player in the development of the widely used Opencast Matterhorn platform for educational video management.

As in other repositories, transcription and translation of video lectures in **VideoLectures** is needed to make them accessible to speakers of different languages and to people with disabilities.

However, also as in other repositories, most lectures in **VideoLectures** are neither transcribed nor translated due to the lack of efficient solutions to obtain them at a reasonable level of accuracy.

Objectives

Our aim is to develop innovative, cost-effective solutions to produce accurate transcriptions and translations in VideoLectures, with generality across other Matterhorn-related repositories.

Three scientific and technological objectives:

- 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.

Results

The main result of ***trans*Lectures** is a set of cost-effective tools to produce accurate transcriptions and translations in **VideoLectures** and other Matterhorn-related repositories.

Apart from **VideoLectures**, we consider a smaller repository of Spanish video lectures, **poliMedia**, which is also part of the Matterhorn Community.

For transcription, we consider English and Slovenian in **VideoLectures**, which account for more than 90% of lectures, and Spanish in **poliMedia**.

For translation, we consider the language pairs: $en \leftrightarrow es$, $en \leftrightarrow sl$, $en \rightarrow fr$ and $en \rightarrow de$.

Impact

The most important expected impact of transLectures is in the area of **Education**, by providing cost effective solutions to produce accurate video lectures transcriptions and translations.

transLectures effort might become a big step in making educational repositories truly accessible both to speakers of different languages and to people with disabilities.

transLectures tools will be made available not only to **VideoLectures** and **poliMedia**, but also to the large community of educational institutions provided through K4A, the Matterhorn/Opencast Community and the OCWC international community (more than 600 universities).

ID card

- Funded under: [FP7-ICT](#)
- Objective: ICT-2011.4.2 - Language Technologies
- Total Cost: €4,491,143.00
- EC Contribution: €3,125,000.00
- Project Type: CP
- Call for proposals: 287755
- Execution: *From:* 11/2011 *To:* 10/2014

Participants

No	Name	Short name	Country	Exit Month	Exit month
1	UNIVERSITAT POLITECNICA DE VALENCIA	UPV	Spain	1	36
2	XEROX SAS	XEROX	France	1	36
3	INSTITUT JOZEF STEFAN	JSI	Slovenia	1	36
3+	KNOWLEDGE FOR ALL FOUNDATION	K4A	UK	1	36
4	RHEINISCH-WESTFAELISCHE TECHNISCHE HOCHSCHULE AACHEN	RWTH	Germany	1	36
5	EUROPEAN MEDIA LABORATORY GMBH	EML	Germany	1	36
6	Deluxe Digital Studios Ltd	DDS	UK	1	36