TraMOOC Project Overview Presentation

Translation for Massive Open Online Courses

Overview Presentation

Translating Europe Forum 2016
“MOOCs: The future of online training?”
Table of contents

- TraMOOC in a nutshell
- Project Motivation
- Project Objectives
- Work Description
- The TraMOOC Platform
- The TraMOOC Consortium

SUMMARY
- WHY?
- WHAT?
- HOW?
- RESULT?
- WHO?
### Project details

<table>
<thead>
<tr>
<th><strong>Partially funded</strong></th>
<th>European Union’s Horizon 2020 research and innovation programme under grant agreement No 644333</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Thematic priority</strong></td>
<td>Information and Communications Technologies (ICT)</td>
</tr>
<tr>
<td><strong>Topic</strong></td>
<td>Approved under 1&lt;sup&gt;st&lt;/sup&gt; call of the ICT priority in the strategic objective “Cracking the language barrier”</td>
</tr>
<tr>
<td><strong>Duration</strong></td>
<td>36 months (From 2015-02-01 to 2018-02-01)</td>
</tr>
<tr>
<td><strong>Budget</strong></td>
<td>Approximately 3M€</td>
</tr>
<tr>
<td><strong>Coordinated by</strong></td>
<td>Humboldt-Universität zu Berlin</td>
</tr>
<tr>
<td><strong>Consortium</strong></td>
<td>10 organizations from 6 European countries</td>
</tr>
</tbody>
</table>
TraMOOC in a nutshell

Objectives & expected impacts

• TraMOOC aims to make existing monolingual educational material available to speakers of other languages.

• The project’s vision is to tear down language barriers, thus providing previously excluded groups of people with new educational chances.

• The project results will be showcased and tested on a MOOC platform and on the VideoLectures.Net digital video lecture library.

• The core of the service will be open-source, with some premium add-on services which will be commercialised.

• The translation methodology is automatic and language-independent in nature and will be showcased for 11 indicative language pairs - 9 EU (DE, IT, PT, NL, BG, EL, PL, CZ and HR), and 2 BRIC languages (RU and ZH).
Main novelties

- Novel research in online and open education
  - Novel translation evaluation schemata
  - Added value to existing tools and resources in linguistics, natural language processing, text analytics, data mining and machine translation scientific communities
  - Topic identification of the source and translated text
  - Sentiment analysis on users’ posts on fora and social websites will be used for extracting users’ opinion on the translated material
Main novelties

- Novel exploitation activities:
  - Global positioning via UNESCO Chair on Open Technologies for OER and Open Learning
  - Creation of European Open Education Consortium
  - Deployment into the Opencast Matterhorn platform (800 institutions)
• TraMOOC in a nutshell
• Project Motivation
• Project Objectives
• Work Description
• The TraMOOC Platform
• The TraMOOC Consortium
MOOCs have been growing rapidly in size and impact.

This year, the number of universities offering MOOCs has doubled to exceed 400 universities, with a doubling of the number of cumulative courses offered, to 2400. It is estimated that 16-18 million students attend MOOCs worldwide.*

Yet the language barrier constitutes a major growth impediment in reaching all people and educating all citizens.

TraMOOC aims at tackling this impediment by developing high-quality translation of all types of text genres included in MOOCs (e.g. reading material, presentations, assignments, video lecture subtitles, MOOC forum blog text) from English into 11 European and BRIC languages, which constitute strong use cases, are hard to translate into and have weak MT support.
TraMOOC intends to address the needs of:

- **European and world citizens**, who want access to open online education that is not constrained by language barriers.
- **MOOC providers**, who wish to offer high-quality, integrated multilingual educational services.
- **Machine Translation developers**, who need a platform for promoting, testing and comparing their solutions.
- **Language Technology engineers**, who want access to accurate linguistic infrastructure with wide coverage, even for less widely spoken languages.
Table of contents

- TraMOOC in a nutshell
- Project Motivation
- Project Objectives
- Work Description
- The TraMOOC Platform
- The TraMOOC Consortium

SUMMARY

WHY?
WHAT?
HOW?
RESULT?
WHO?
The main objective of the project is to create a translation service for the educational material of MOOCs.

To achieve this goal, the project will pursue the following challenging scientific and technological objectives:

- Generation of high-quality machine translations
- Establishment of a new, appropriate, standardized, multi-level evaluation schema for determining the value of the produced translations
- Automatic bootstrapping of new resources for languages that are fragmentarily or weakly equipped with infrastructure
- Development of a machine translation process which will be language-independent
<table>
<thead>
<tr>
<th>Table of contents</th>
<th>28/10/2016</th>
<th>Valia Kordoni</th>
</tr>
</thead>
<tbody>
<tr>
<td>TraMOOC in a nutshell</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project Motivation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project Objectives</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work Description</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The TraMOOC Platform</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The TraMOOC Consortium</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SUMMARY**

**WHY?**

**WHAT?**

**HOW?**

**RESULT?**

**WHO?**
Work Description

Project Structure

SA-1
WP4: Machine Translation
WP5: Explicit Translation Evaluation
SA-2
WP6: Implicit Translation Evaluation
WP7: System Integration/Expandability/Updateability
WP8: System Viability/Exploitation/Commercialization
WP9: Dissemination and Diffusion
WP1: Management and Coordination

WP2: Architecture and Requirements Analysis
WP3: Data Collection and Infrastructure Exploration/Adaptation/Bootstrapping

Valia Kordoni

28/10/2016
TraMOOC System Architecture

TraMOOC’s architecture is comprised of four main components:
- User Interface / Web-based Service
- Translation Engines
- Evaluation Component
- Crowdsourcing Activities

Work Description

SUMMARY

WHY

WHAT

HOW

RESULT

WHO

28/10/2016

Valia Kordoni
Table of contents

• TraMOOC in a nutshell  
• Project Motivation  
• Project Objectives  
• Work Description  
• The TraMOOC Platform  
• The TraMOOC Consortium  

SUMMARY  WHY?  WHAT?  HOW?  RESULT?  WHO?
• For the first time…

  o A translation platform will offer translation services for 11 languages and will be open and available to everyone

  o A translation platform focuses on high quality machine translation, even though the targeted languages include weakly supported languages as well as languages that have proved to be hard to translate into in previous MT solutions

  o New services will be offered which focus on translating various types of educational MOOC text, which are currently not serviced by any of the existing translation platforms
The core of the service will be open-source with some premium add-on services such as:

- professional translation,
- transcription and subtitling services,
- support for additional MT target languages, and
- multilingual review services for students’ assignments by domain experts in the target languages.
Table of contents

- TraMOOC in a nutshell
- Project Motivation
- Project Objectives
- Work Description
- The TraMOOC Platform
- The TraMOOC Consortium

SUMMARY
WHY?
WHAT?
HOW?
RESULT?
WHO?
TraMOOC brings together a consortium of leading researchers, highly relevant industry organizations and leading use-case partners.

The partners’ diverse interests in machine translation, linguistics, text mining web analytics and crowdsourcing methodologies-related areas make the consortium ideally placed to tackle the challenges associated with TraMOOC.

The design of the scientific areas and the associated work packages have been arranged carefully to ensure maximum efficiency of input from each partner while maintaining a suitable distribution of responsibilities.