



LetsMT!
**Platform for Online Sharing of Training Data
and Building User Tailored MT**
ICT Policy Support Programme: Theme 5 - Multilingual web
CIP-ICT-PSP.2009.5.1
Pilot Type B
Grant Agreement n°250456
<http://www.letsmt.eu>

List of partners
Tilde, Latvia (coordinator)
University of Edinburgh, Human Communication Research Centre, UK
University of Zagreb, Faculty of Humanities and Social Sciences, Croatia
University of Copenhagen, Centre for Language Technology, Danmark
Uppsala University, Department of Linguistics and Philology, Sweden
Zoorobotics, Netherlands
Moravia, Czech Republic

Project duration: March 2010 — August 2012

Summary

LetsMT! is developing a user-driven machine translation “factory on the cloud” to provide online services for the generation, customization, and running of statistical machine translation (SMT) systems. LetsMT! allows users to upload their parallel corpora to an online repository and generate user-tailored SMT systems based on user selected data.

Current mass-market and online MT systems are general in nature and perform poorly for domain and user specific texts. System adaptation is a prohibitively expensive service which is not affordable for smaller companies or public institutions. In particular the localization industry is not able to fully exploit the data they have.

To exploit the potential of existing open SMT technologies LetsMT! creates a novel online collaborative platform for data sharing and MT building. LetsMT! gathers public and user-provided MT training data and generates multiple MT systems by combining and prioritizing this data.

The core engine of the LetsMT! platform is Moses, which is adapted for rapid training and updating tasks. The LetsMT! platform hides the complexity of the Moses toolkit. It provides the infrastructure and an easy user interface, where users can upload data, train the SMT, and use it. User does not have to worry about the complexities of installation, compilation, configuration, and data processing.

In the first year of the project the core functionality has been developed (invitation-only BETA): processing of parallel translation memory data, generation of language and translation models, management of multiple SMT systems, translation website, translation widget and plug-in for Mozilla Firefox, integration in computer aided translation tool SDL Trados.