

# Let's MT!

# META SHARE NET



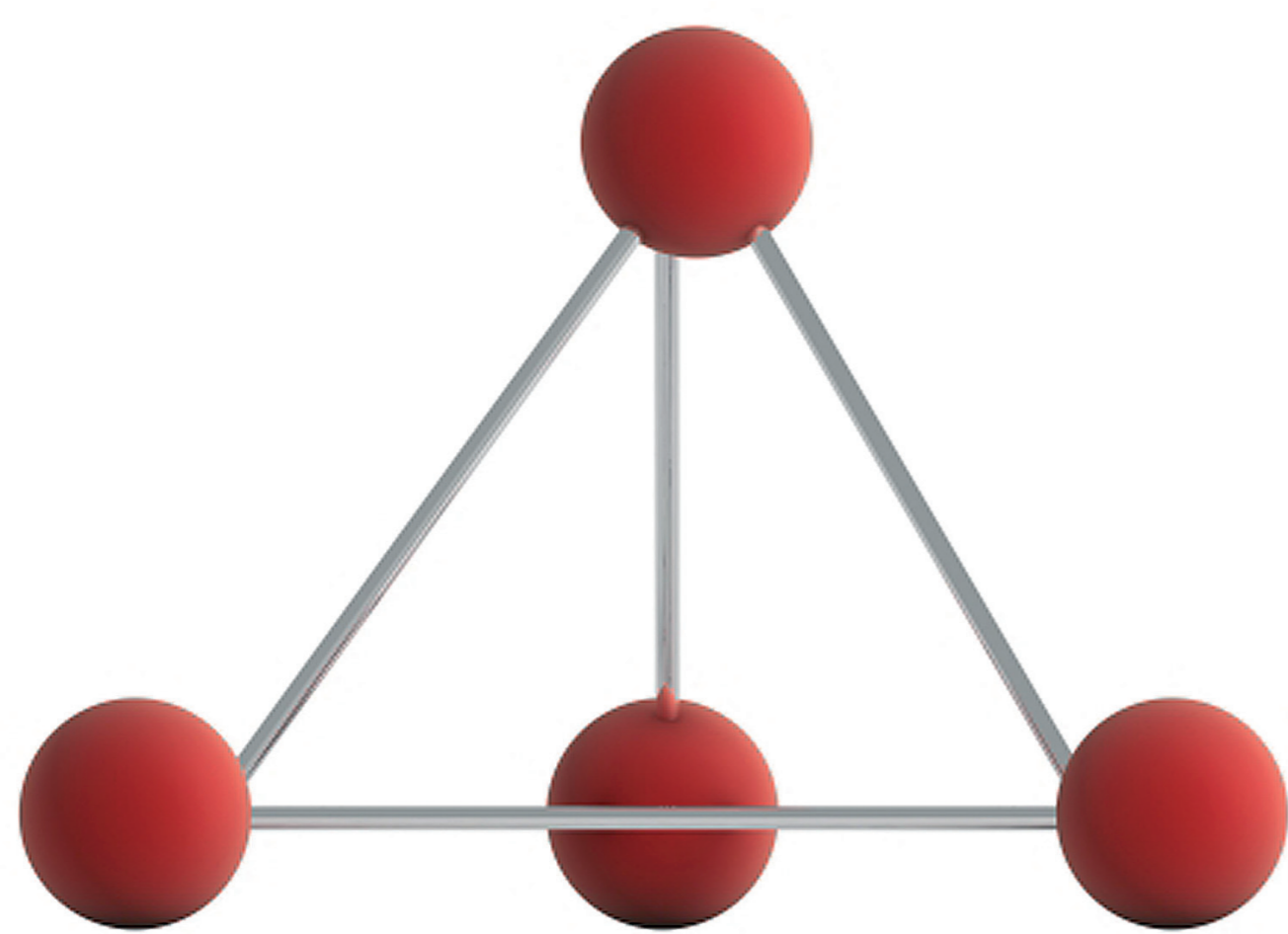
The research within the project LetsMT! leading to these results has received funding from the ICT Policy Support Programme (ICT PSP), Theme 5 – Multilingual web, grant agreement no 250456.

## Platform for Online Sharing of Training Data and Building User Tailored Machine Translation



### Main project objective

The core objective of Let'sMT! project is to provide innovative online MT services through sharing of parallel corpora provided by users, with emphasis on less-covered languages and specialized domains.



### Results

LetsMT! has reached the following results:

- **full version** of the translation services online;
- **widget** for free inclusion in a web page to provide its translation using LetsMT! services;
- **browser plug-in** for Internet Explorer and Mozilla Firefox for using LetsMT! services;
- **application** for mobile devices using LetsMT! services;
- **integration** of services in web pages;
- **integration** of services in computer aided translation software.

### Target users

The LetsMT! project provides MT solutions for European citizens and businesses allowing more efficient usage of multilingual content. The LetsMT! platform targets actors in:

- **localisation & translation industry** – facilities for training MT systems on their data and generating custom MT systems to be used by localisation providers or their clients;
- **users of business and financial news** – free and instant MT services, with an emphasis on less-covered languages;
- **holders of language resources** – easily building of MT services using their specific parallel data.

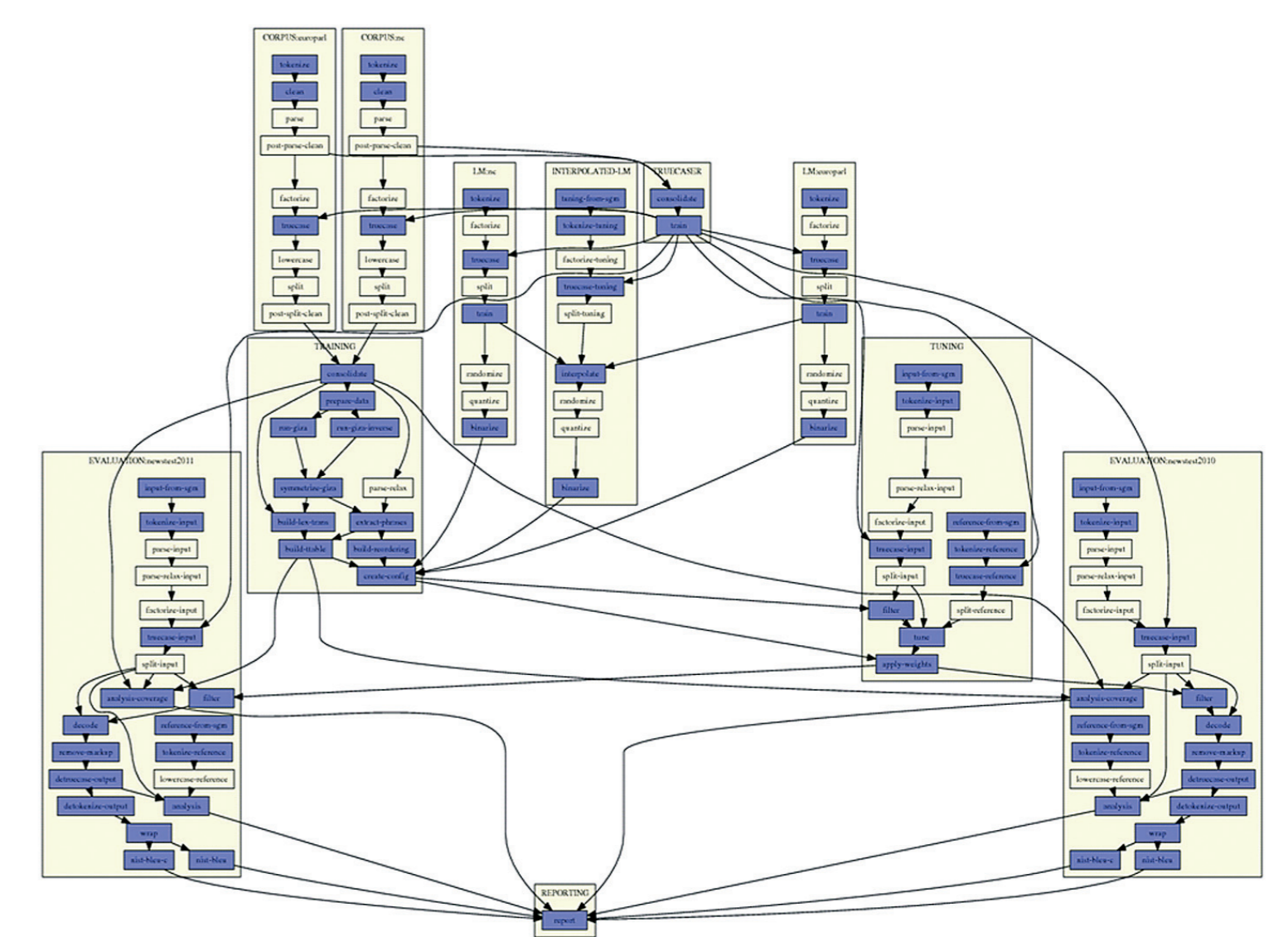
### Proposed services/solutions

Current MT systems are mainly built on the data accessible on the web, but this is just a fraction of all available parallel texts. Most of them still reside in the local systems of different corporations, public and private institutions, and on the desktops of individual users. Also, building custom MT solutions is a costly process that requires a lot of technical expertise; the results are proprietary and are typically not shared outside the organization that has invested in building a custom MT solution.

To fully exploit the huge potential of existing open MT technologies we have built an innovative **online collaborative platform for data sharing and MT building**. This platform is capable of holding public or user provided MT training data and generating multiple MT systems by combining this data.

### Technology

The core of the LetsMT! platform is the Moses SMT toolkit and Giza++. These are publicly available open source tools that are well known and widely used in statistical machine translation. The system also uses existing publicly available parallel corpora such as JRC-Acquis, Europarl, OPUS and it will allow users to upload and build parallel corpora using system incorporated aligners. This service will enable professional users to generate and employ customized MT services of higher quality based on the specific terminology and style required by their clients.



### Project partners

**TILDE:** Tilde SIA, Riga, Latvia

**UEDIN:** University of Edinburgh, The Human Communication Research Centre (HCRC), Edinburgh, UK

**FFZG:** University of Zagreb, Faculty of Humanities and Social Sciences, Department/Institute of Linguistics, Zagreb, Croatia

**UCPH:** University of Copenhagen, Centre for Language Technology, Copenhagen, Denmark

**UUP:** Uppsala University, Department of Linguistics and Philology, Uppsala, Sweden

**SEM:** Zoorobotics BV, Alphen a/d Rijn, Netherlands

**MOR:** Moravia, Brno, Czech Republic

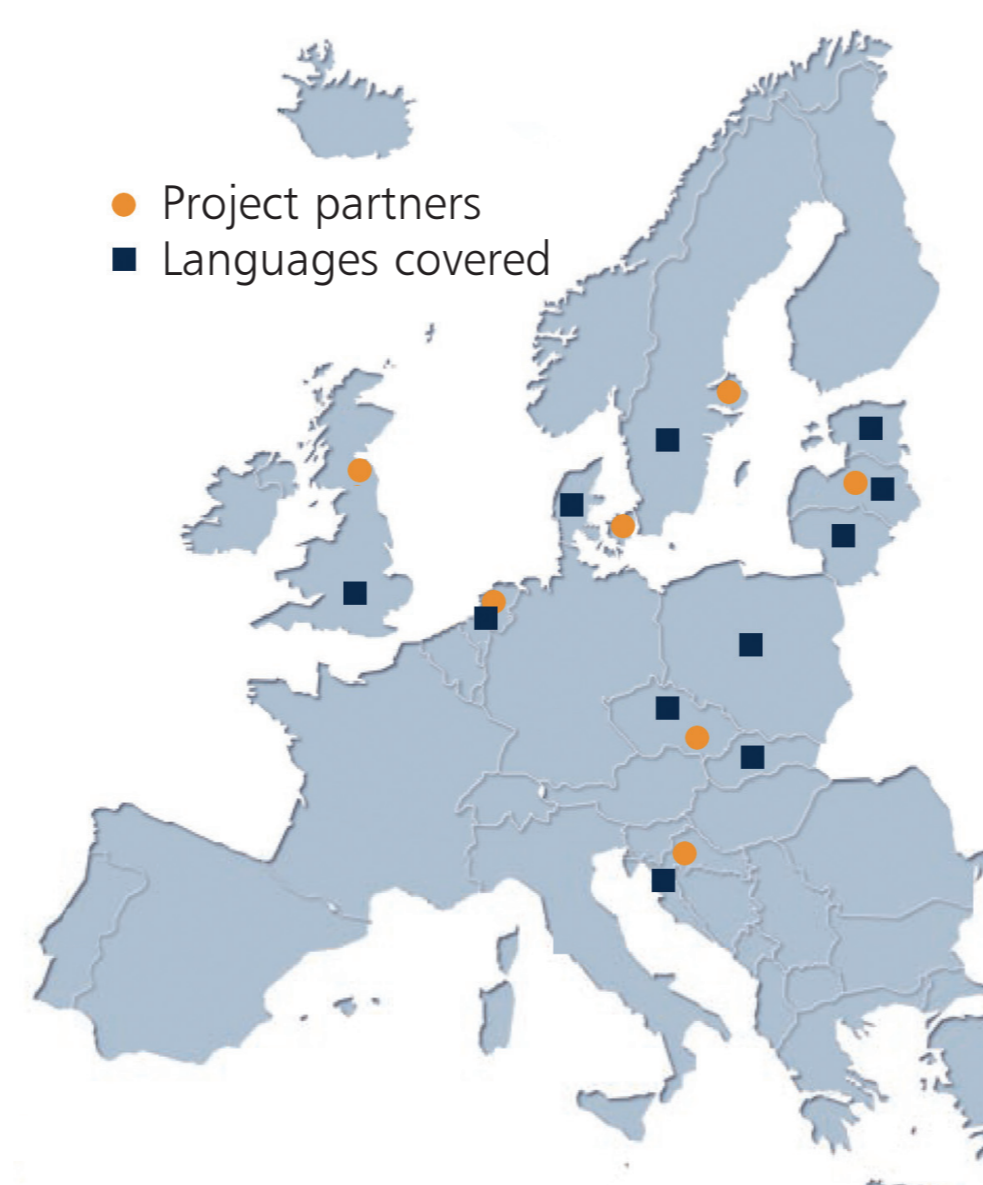
### Contact

Tilde, SIA  
75a Vienibas Gatve  
LV-1004 Riga  
LATVIA

Arturs Vasilevskis, project manager  
p: + 371 67605001  
f: +371 67605750  
e: arturs.vasilevskis@tilde.lv  
w: http://www.tilde.lv

LetsMT! results will also be available

META SHARE



[www.letsmt.eu](http://www.letsmt.eu)

A true cross-European project providing services that are inclusive and universal, but at the same time focusing specifically on a number of European languages that currently lack MT services of professional quality: Latvian, Lithuanian, Estonian, Czech, Slovak, Polish, Croatian, Danish, Dutch, and Swedish. The initial collection of corpora is focussed on parallel texts in these languages and English, as well as other feasible combinations.