Experiences in Building the *Let's MT!* Portal on Amazon EC2

Jörg Tiedemann
Uppsala Universitet
jorg.tiedemann@lingfil.uu.se

ABSTRACT
In this presentation I will discuss the design and implementation of *Let's MT!*, a collaborative platform for building statistical machine translation systems. The goal of this platform is to make MT technology, that has been developed in academia, accessible for professional translators, freelancers and every-day users without requiring technical skills and deep background knowledge of the approaches used in the backend of the translation engine. The main challenge in this project was the development of a robust environment that can serve a growing community and large numbers of user requests. The key for success is a distributed environment that allows a maximum of scalability and robustness. With this in mind, we developed a modular platform that can be scaled by adding new nodes to the different components of the system. We opted for a cloud-based solution based on Amazon EC2 to create a cost-efficient environment that can dynamically be adjusted to user needs and system load. In the presentation I will explain our design of the distributed resource repository, the SMT training facilities and the actual translation service. I will mention issues of data security and optimization of the training procedures in order to fit our setup and the expected usage of the system.

KEYWORDS: Let's MT!, Statistical Machine Translation, Distributed Computing.