KantanMT: Cloud-based Machine Translation at your fingertips

In recent years Statistical Machine Translation (SMT) has established a dominant position among the variety of available machine translation paradigms. While research efforts have been mainly focused on improving the core SMT technology, i.e., the Moses toolkit and related pre- and post-processing techniques, KantanMT has focused on bringing this technology to the end user in an accessible and scalable manner.

KantanMT aims at two general tasks: (i) to allow users to efficiently build personalized, i.e., using their own data, SMT engines; and (ii) to use these customized SMT engines to translate user content.

The KantanMT platform is fully distributed on the cloud. It faces several challenges typical for distributed computer systems, for example, request handling and resource allocation. The architecture of the KantanMT platform addresses these challenges and provides efficient solutions.

This presentation addresses the design and implementation decisions that are required in order to construct an efficient distributed system such as KantanMT. In particular it covers the topics of (i) user interface; (ii) request handling and resource allocation and (iii) security and data protection.