Polish Spoken Language Translation

EU-BRIDGE Partner
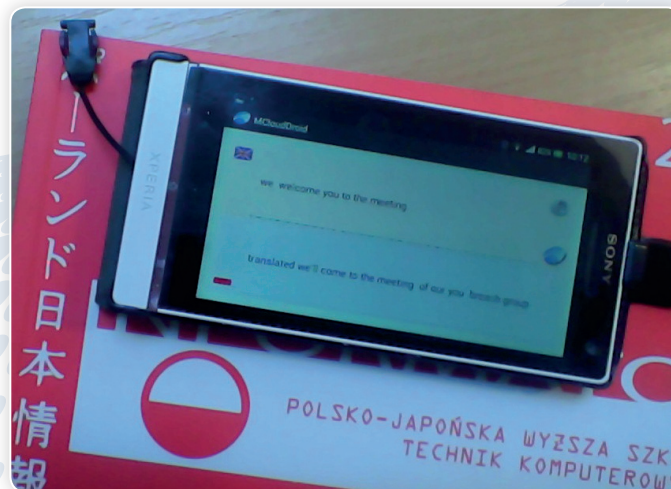
KIT, Pervoice, FBK, UEDIN

Contacts

Koszykowa 86
02-008 Warszawa
Poland
Krzysztof Marasek
kmarasek@pjwstk.edu.pl

Description and Exploitable Knowledge

Polish, one of the West-Slavic languages, due to its complex inflection and free word order forms a challenge for both automatic speech recognition (ASR) and spoken language translation (SLT). Seven cases, three genders, animate and inanimate nouns, adjectives agreed with nouns in terms of gender, case and number and a lot of words borrowed from other languages which are sometimes inflected as those of Polish origin, cause problems in establishing vocabularies of manageable size for translation to/from other languages and sparseness of data for statistical model training. Despite of ca. 60 millions of Polish speakers worldwide a number of public available resources for preparation of SMT system is rather limited, making the progress slower compared to other languages.

EU-BRIDGE project helped us develop our skills and collect more resources necessary to accomplish the task of Polish SMT, specifically for Polish/English language pair. Our SMT systems have since been successfully applied to various domains, including lecture and broadcast news, parliament speeches, tourist assistance, movie subtitles and medicine-related documents. We developed own BLSTM/DNN ASR decoder, Polish speech synthesis, mobile clients (Android) and specialized tools for the Polish language (text normalizer, corpus alignment, phonetic transcription, speech detection).

Android Client

Infrastructure

Depending on application our systems work on dedicated server (e.g. in telecommunication) or on the cloud (for mobile applications). We are currently support MCloud service architecture and the system developed by the U-STAR consortium compatible with ITU F.745 and H.625 recommendations.

Application Sectors

Telecommunication, tourist, medicine, entertainment

Technical requirements

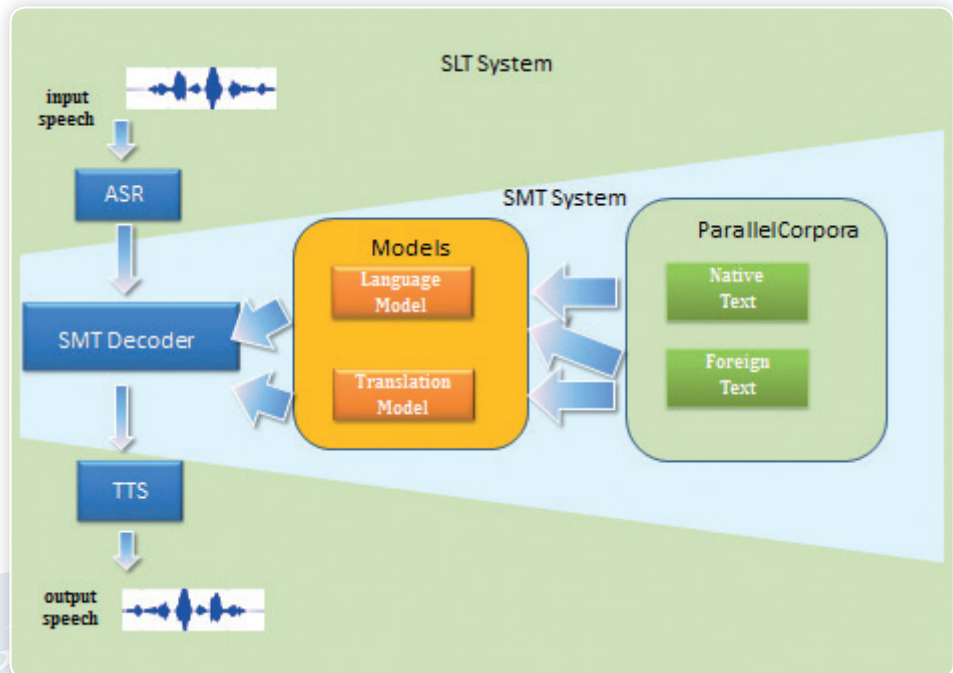Fast internet Connection or dedicated SLU server.

Terms of Availability

Can be inquired at the Polish-Japanese Institute of Information Technology

IPR Protection

Polish-Japanese Institute of Information Technology, EU-BRIDGE Consortium

EU-BRIDGE Partners (involved)

KIT, Pervoice, FBK, UEDIN

SLT-Architecture