

Quality Translation In Half The Time

Lingtech has replaced its translation system with a new statistical machine translation system (SMT) developed in collaboration with PhD. student Jakob Elming and associate professor Dan Hardt at the CBS – Copenhagen Business School. Quality is high and speed is even higher.

The more you use it, the better it gets – said of the new translation system being put into operation by Danish language company Lingtech.

“This is an example of a new breed of systems which are able to learn – called 'intelligent systems'. The program learns from uploaded text – and it learns more when the final, edited version of the translation is uploaded. This means that the program improves all the time,” says Matthias Buch-Kromann, Research Group Manager at the Department of International Language Studies and Computational Linguistics at the CBS.

Sabine Kirchmeier-Andersen is the former head of the Department of Computational Linguistics and is now the director of the Danish Language Council. She is delighted that the system has now been put into production:

“It is a really good tool, which I am very pleased to see has come this far. A machine learning system can be trained to draw conclusions based on large amounts of text. Combine it by coding additional linguistic analysis, and it will be incredibly powerful,” she says.

Although there is room for improvement, feedback on the basic version from test companies was so positive that the program was put into production and developed for commercial use.

“It is a known technology which we have developed. But only in recent years have we been able to get this far, as computer capacity was previously insufficient to handle the large volumes of text which need to be uploaded to train the program,” says Jakob Elming, Ph.D. student at the CBS and co-founder of LanguageLens.

The translation process involves preparing the program by uploading large volumes of parallel text, e.g. previous translations into Danish and English. Once this is done, the program translates new text based on the previous texts.

The system checks the possible translations available to it for each segment of the sentence and selects the most probable translation based on the context of the sentence.

This means that both words and grammar are checked giving a better result than that normally seen for machine translations.

“I am very surprised at the quality. It is much faster, and the quality is also very good if you compare it to previous methods within machine translation,” says Birgit Pichat, CEO of Lingtech, which participated in the development.

Lingtech prepares the system by uploading large text volumes in one batch. Once this is done, the user only has to select which file to translate and then press “Translate”. This is done using a web browser. At first, the system is able to translate Word files and HTML text as this is where the greatest need is.

Contact:

Birgit Pichat, CEO, Lingtech:

Direct: 33 86 38 75

Mobile: 23 24 28 46

Mail: birgit.pichat@lingtech.com

Morten Esmann, Head of Technology, Lingtech:

Direct: 33 86 38 71

Mobile: 20 95 24 21

Mail: morten.esmann@lingtech.com

Dan Hardt, LanguageLens:

Direct: 38 15 31 26

Mobile: 28 52 62 12

Mail: dh.isv@cbs.dk

Jakob Elming, LanguageLens

Direct: 38 15 33 84

Mobile: 25 88 37 79

Mail: jel.isv@cbs.dk