



CASMACAT: Cognitive Analysis and Statistical Methods for Advanced Computer Aided Translation

Philipp Koehn, University of Edinburgh, pkoehn@inf.ed.ac.uk

Michael Carl, Copenhagen Business School, mc.isv@cbs.dk

Francisco Casacuberta, Polytechnic University of Valencia, fcn@dsic.upv.es

Eva Marcos, Celer Solutions, eva.marcos@celersol.com

<http://www.casmacat.eu/>

Description

In its third year, the CASMACAT project has developed – in collaboration with the MATECAT project – a new open source workbench for translators that is deployed over the web and as a stand-alone tool. With insights from cognitive studies of translator behavior, new types of assistance have been developed and tested in field trials.

The cognitive studies of translator behavior with the CASMACAT workbench include

- identification of translator types and styles
- cognitive and user models of translation processes

The advances of the CASMACAT project include:

- interactive translation prediction with machine learning methods
- interactive translation prediction for syntax-based models
- active learning applied to translation tasks
- sentence and word level confidence measures
- synthesis of translation memories and machine translation
- word alignment visualization
- display of multiple translation options
- online learning (incremental updating of models)
- domain and user adaptation
- integration of e-pen as input device
- logging and replay mode
- integration of eye tracker for collection of user activity data
- visualization tools for logging data

Additional advances are currently under development:

- integration of paraphrasing and alternative translations on demand
- automatic reviewing

The tool comprising the CASMACAT workbench are currently integrated into a stand-alone version that runs on any computer.