1 Publishable summary

The objective of the MateCat project is to improve the integration of machine translation (MT) and human translation within the so-called computer aided translation (CAT) framework. CAT tools nowadays are the dominant technology in the translation industry. They provide translators with text editors that can manage multiple document formats and suitably segment documents into meaningful units for translation. Most importantly, CAT tools provide integrated access to translation memories, terminology databases, concordance tools and, more recently, machine translation (MT) engines.

A translation memory (TM) is essentially a database of translated segments. During translation, the CAT tool queries the TM to search for exact or fuzzy matches of the current source segment. The corresponding translations from the repository are proposed to the user as translation suggestions. Once a segment is translated, its source and target texts are added to the TM for future queries.

The integration of suggestions from an MT engine as a complement to TM matches is motivated by recent studies in the literature which have shown that post-editing (PE) MT suggestions can substantially improve the productivity of professional translators. The MateCat project leverages the growing interest and expectations in statistical MT by advancing the state-of-the-art in three directions:

- *Self-tuning MT*, i.e. methods to let MT seamlessly adapt to a specific translation project;
- *User adaptive MT*, i.e. methods to quickly adapt statistical MT in response to user corrections and feedback,
- *Informative MT*, i.e. MT that provides additional information about the translation in order to enhance users' productivity and work experience.

Research along these three directions has converged into a new generation of CAT software, which is both an enterprise level translation workbench currently used by several thousand professional translators, and an advanced research platform for advancing the state of the art in MT and PE by integrating new MT functions, running PE experiments and measuring user productivity. The MateCat Tool, which is distributed under an open source license, combines features of the most advanced systems (commercial software like the popular SDL Trados Workbench as well as free software like OmegaT) with new features. These include: *i*) an advanced API for the MT Platform *Moses*, customizable to languages and domains, *ii*) ease of use through a clean and intuitive web interface that enables the collaboration of multiple users on the same project, *iii*) concordancer, terminology databases and support for customizable quality estimation components and iv) advanced logging functionalities.

Progress in MateCat has been systematically measured through extensive field tests that evaluated the utility and usability of the new MT features. Field tests were run with professional translators performing real translation projects with the MateCat Tool.

Progress report

Research and development

During the third year of the project the consortium focused on consolidating all research and development activities in view of the final field tests and the release of the final version of the MateCat Tool, the MT software and benchmarks.

Work on self-tuning MT focused on consolidating and evaluating alternative adaptation methods involving continuous space models and suffix array data structures. In particular, adaptation with continuous space models has been also evaluated with a specific field tests of Year 3.

Work on user-adaptive MT developed, implemented and evaluated several original on-line adaptation methods, based both on generative and discriminative models (D2.1). In particular, generative methods based on cache models were integrated in the Moses engine and finally in the MT server. The whole pipeline for on-line learning was optimized for efficiency, given the close to real-time needs of the MateCat Tool, and for concurrent use by multiple users.

Work on informative MT focused on robust quality estimation methods capable to adapt users and tasks and on techniques to extract useful terminology for the translators before they start translating the document (D3.2).

The MateCat Tool was finalised and put in production starting from Spring 2014. Thanks to extensive use by hundreds of translators, many bugs were discovered and fixed and several minor improvements in the user interface were introduced. The third version of the CAT tool has been released in September 2013 (D4.3) and presented at the end of October 2014 at the AMTA, TAUS, and Localization World conferences, all co-located in Vancouver. In the same venue, the MateCat User Group Event (MS6) was organised in which the MateCat Tool was officially introduced to selected representative of the translation industry.

Field test

During Summer and Fall 2014, several field tests were run to measure the impact of the new MT functions developed in WP1-3 on translation quality and user productivity (MS5, D5.5). A

first field test aimed at evaluating the impact of self-tuning MT, user-adaptive MT and informative MT on user productivity. In particular, the test compared user productivity with the MateCat Tool under two contrastive working conditions, differing in the employed MT engine. In the first condition, translators post-edited MT suggestions from a state-of-the-art domain-adapted Moses engine, while in the second condition, translators received suggestions from an domain-adapted MT engine also integrating: (i) *project adaptation*, i.e. the MT engine adapts at the end of each working day from all the available translation project specific data, (ii) *online adaptation*, i.e. the MT engine continuously adapts from each single segment post-edit produced by the translator; (iii) *quality estimation*, i.e. the MT engine supplies quality estimation score along with each suggestion; and (iv) *terminology support*, i.e. terminology is automatically extracted from the document and incorporated in the MT engine. The field test was run on two translation directions (English-French, English-Italian), two domains (information technology and legal), and with 4 professional translators for each direction-domain combination, for a total of 16 translators. As recommended by the project reviewers an additional field test was also run on a new domain (TED Talks) and one translation direction (English-French).

As in previous field tests, two key performance indicators were used to measure user productivity: (i) the post-editing effort, namely the amount of corrections made by the translators on the MT suggestions, and (ii) the time to edit, that is the average number of words translated per hour. Results of the field-tests show overall significant improvements on productivity on the official MateCat tests, and also show that the adaptation methods we developed are robust and apt to cope with documents that contain less repetitive content. Subsequent lab test of MT quality based on human judgements confirmed that adaptive MT methods do favourably impact both on MT quality and even on post-editing quality. Finally, specific field tests were carried out on the informative MT components, namely the MT quality estimation and terminology extraction modules, and on the self-tuning MT in isolation. All field tests produced positive results thus proving the utility of the tested components.

Dissemination

During the third year, the MateCat consortium has further increased its dissemination activities and significantly promoted the project outcomes among the industrial and user communities. The consortium was present at several international conferences and workshops, i.e. the EACL 2014 Workshop on Humans and Computer-assisted Translation, the 9th ACL Workshop on Statistical Machine Translation (WMT 2014), the 9th Language Resources and Evaluation Conference (LREC 2014), the annual meeting of the Association for Computational Linguistics (ACL 2014), the 25th International Conference on Computational Linguistics (COLING 2014), the 4th COLING International Workshop on Computational

Terminology (*Computerm*), the Machine Translation Marathon (MTM), the 11th biennial conference of the Association for Machine Translation in the Americas (AMTA), the AMTA 2014 Workshop on Interactive and Adaptive Machine Translation (IAMT), the TAUS Annual Conference 2014, the Localization World Conference, the Conference on Empirical Methods in Natural Language Processing (EMNLP), and the Italian Computational Linguistics Conference (CLIC). During the same period the consortium also published four journal articles, one in *Intelligenza Artificiale* and three in *Machine Translation*. The project technology has been also publicly showcased at the COLING, AMTA, TAUS and Localization World conferences. Finally, a User Group Event has been organised, co-located with AMTA and TAUS, and attended upon invitation by about 70 representatives of the translation industry. The event gave a short overview of the research project and officially launched the MateCat Tool website

Expected Results and Impact

During the third year the project has continued to delivered remarkable good results both on the research and the industrial sides. New developments on machine translation technology, although not all yet fully deployable in an industrial setting, have shown great promise and raised interest in the scientific community. MateCat together with the Casmacat project have attracted much attention from researchers worldwide and we expect that this will stimulate in the future further progress around the research topics launched by these projects. From the industrial side, the MateCat Tool has now found its own way as a commercial product, which is indeed a great success. The tool has been put in production since February 2014 and is currently used by thousands of translators. The strategy of the industrial partner for the near future is to significantly enlarge the user basis of the tool, which is offered as a free web-service with an option of re-routing translation project to Translated.