Final Public Report

Philipp Koehn, Vicent Alabau, Michael Carl, Francisco Casacuberta, Mercedes García-Martínez, Jesús González-Rubio, Frank Keller, Daniel Ortiz-Martínez, Germán Sanchis-Trilles, Ulrich Germann

Distribution: Public

CASMACAT
Cognitive Analysis and Statistical Methods
for Advanced Computer Aided Translation

ICT Project 287576 Deliverable

Project funded by the European Community
under the Seventh Framework Programme for
Research and Technological Development.

Project ref no.	ICT-287576
Project acronym	CASMACAT
Project full title	Cognitive Analysis and Statistical Methods for Advanced Computer Aided Translation
Instrument	STREP
Thematic Priority	ICT-2011.4.2 Language Technologies
Start date / duration	01 November 2011 / 36 Months

Distribution	Public
Contractual date of delivery	November 18, 2014
Actual date of delivery	January 7, 2015
Date of last update	January 7, 2015
Deliverable title	Final Public Report
Type	Report
Status & version	Final
Number of pages	49
Contributing WP(s)	All
WP / Task responsible	UEDIN
Other contributors	CBS, CS, UPVLC
Internal reviewer	
Author(s)	Philipp Koehn, Vicent Alabau, Michael Carl, Francisco Casacuberta, Mercedes García-Martínez, Jesús González-Rubio, Frank Keller, Daniel Ortiz-Martínez, Germán Sanchis-Trilles, Ulrich Germann
EC project officer	Aleksandra Wesolowska
Keywords	

The partners in CASMACAT are:

University of Edinburgh (UEDIN)
Copenhagen Business School (CBS)
Universitat Politècnica de València (UPVLC)
Celer Soluciones (CS)

For copies of reports, updates on project activities and other CASMACAT related information, contact:

The CASMACAT Project Co-ordinator
Philipp Koehn, University of Edinburgh
10 Crichton Street, Edinburgh, EH8 9AB, United Kingdom
pkoehn@inf.ed.ac.uk
Phone +44 (131) 650-8287 - Fax +44 (131) 650-6626

Copies of reports and other material can also be accessed via the project's homepage: http://www.casmacat.eu/

Contents

1 Executive Summary

The CASMACAT project built the next generation translator's workbench to improve productivity, quality, and work practices in the translation industry.

We carried out cognitive studies of actual unaltered translator behaviour based on key logging and eye tracking. The acquired data was examined for how interfaces with enriched information are used, to determine translator types and styles, and to build a cognitive model of the translation process.

Based on insights gained in the cognitive studies, we developed novel types of assistance to human translators and integrated them into a new workbench, consisting of an editor, a server, and analysis and visualisation tools. The workbench was designed in a modular fashion and can be combined with existing computer aided translation tools.

We developed new types of assistance along the following lines:

- Interactive translation prediction, where the workbench makes suggestions to the human translator how to complete the translation. We adapted the existing interactive machine translation paradigm by adding input modalities, especially electronic pens and basing the suggestions on better exploitation of novel statistical machine translation models.
- Interactive editing, where the workbench provides additional information about the confidence of its assistance, integrates translation memories, and assists authoring and reviewing.
- Adaptive translation models, where the workbench learns from the interaction with the human translator by updating and adapting its models instantly based on the translation choices of the user.

The project demonstrated the workbench's effectiveness in extensive field tests at a translation agency. In addition, we also reached out to the wider language service industry and online volunteer translation platforms. The outcome of the CASMACAT project was made available as open source software and resources to industry, academia, and to individual end users.

2 Project Context and Objectives

CASMACAT carried out cognitive studies of actual unaltered translator behaviour based on key logging and eye tracking. The acquired data was examined for how interfaces with enriched information was used, to determine translator types and styles, and to build a cognitive model of the translation process. Based on insights gained in the cognitive studies, CASMACAT developed novel types of assistance to human translators and integrated them into a new workbench, consisting of an editor, a server, and analysis and visualisation tools. The workbench was designed in a modular fashion and can be combined with existing computer aided translation tools. The CASMACAT project demonstrated the workbench's effectiveness in extensive field tests of real-life practice of a translation agency.

2.1 Scientific Objectives

While there have been significant improvements to machine translation technology, the vast majority of this work was previously targeted towards bulk translation that is good enough or fit for use. A user on the Internet is satisfied with a rough translation, if it fills her information need. Opposed to that is the demand for high quality translations by the marketplace: the translation of reports and announcements of multi-national organisations, marketing material and product descriptions of commercial companies, and many other localisation needs. Such high quality translations are still almost exclusively provided by human translators.

Productivity of human translators can be increased with computer aided translation (CAT) tools: translation memories are standard in the translation industry, but post-editing machine translation output is only slowly becoming an increasingly used practice. The current integration of machine translation technology into human translators' work processes is often done overly simplistic, breaks their work practices, and it is widely resisted.

Hence, the CASMACAT project carried out in-depth study of translator behaviour to tailor the tools to the requirements of translators, and not the other way around. The project built a novel workbench that increased the productivity of human translators by addressing their needs for the right type of assistance at the right time.

Prior to the project, there had been some progress in aiding human translators, but the vast potential of creating a new workbench for human translators had been mostly unfulfilled. The CASMACAT project believed that the transfer of methods from the statistical machine translation community can be of great benefit to the task of assisting human translators. Whereas the translation technology was ripe enough, design issues of the user interface and its acceptance by the translator had been widely neglected. The development of such tools must not simply follow technical possibilities, but it should be driven by a better understanding of the behaviour of human translators.

2.2 Scope

Human translation is performed by different types of translators, tackles different text types, and deals with different language pairs.

Translators — The project addressed the needs of user communities that range from professional translators to volunteer translators (participating in efforts such as TED, Global Voices or Wikipedia).

Text Types — Much of what professional translators translate is repetitive, technical material. Volunteer translators are more commonly interested in generally accessible material, while monolingual translators seeking out information which may be very technical in their own area of expertise.

Language Pairs — Statistical machine translation methods do not work equally well for all language pairs. Translating between syntactically divergent languages or translating into morphologically rich languages is more difficult. The CASMACAT project tested its workbench on different language pairs (involving English, Spanish, Danish, and German).

2.3 Workbench

The CASMACAT project developed a new open source workbench for human translators, in collaboration with the MateCat project, a parallel EU Framework Programme 7 STREP. All functionalities developed by the project were integrated in a web-based online service which may also be installed locally on the desktop of a translator (dubbed CASMACAT Home Edition). All new successful features developed in the project are implemented within the new system.

The availability as web service makes it easy to integrate the workbench into existing translation workflows. The important aspects are (1) integration with real-time interactive translation prediction systems, (2) novel editing possibilities with e-pen, (3) connection to eye-tracking hardware for detailed analysis, and (4) extensive logging facilities.

2.4 Cognitive Analysis

An important objective of the CASMACAT project was to gain insight into the cognitive processes involved in human translation. How large are the text segments actively considered by a translator (the whole text, individual sentences, or only subsentential segments of limited length)? What are the subtask that a translator spends most time on, e.g., understanding the source text, looking up unknown words, investigating lexical translations, syntactic restructuring of the sentence, ensuring fluency of the output? How does translation differ from well-studied simpler cognitive processes such as reading and text production? How do added assistance such as word alignment change the behaviour of the translator?

The cognitive analysis informed the design of the CASMACAT translation workbench in a range of ways. It determined what types of assistance are offered to the translator, what information should be displayed on the screen, and what information should be hidden as it would be distracting. The project evaluated different versions of the user interface in user studies using eye-tracking and other commonly used methods in cognitive analysis.

2.5 Advanced Computer Aided Translation

CASMACAT used well-established statistical methods and explored novel approaches in order to generate and disambiguate translation proposals. Dynamically generated translation options are sent to and visualised in an interactive translation assistance tool. Two basically different approaches to CAT, Interactive Translation Prediction and Interactive Editing, were developed, compared and evaluated and a cognitive model of the translator were developed to predict the translator's performance.

A novel reworking of the idea of interactive translation prediction (ITP) allowed for the construction of systems that produce high-quality results by placing a human operator at the centre of the production process. The IMT paradigm embeds a statistical MT engine within an interactive editing environment. The human serves as the guarantor of high quality; the role of the automated systems is to ensure increased productivity by proposing well-formed extensions to the current target text, which the operator may then accept, correct or ignore. Interactivity allows the system to take advantage of the human-validated portion of the text to improve the accuracy of subsequent predictions. This interactivity can be applied both to the basic machine translation and to the post-editing of the output of a machine translation system. In this new framework we developed:

- New models, search and machine learning criteria for translation prediction
- New different modalities to interact with the system
- Techniques that allow the system suggest the parts to be corrected
- Active learning and on-line adaptation to new scenarios and translators.

2.6 Field Trials

The CASMACAT workbench was integrated into the workflow of the professional translation agency Celer Solutions which tested the workbench under realistic work conditions. The project measured the effectiveness of its approach by carrying out user studies that measured the productivity of human translators.

CASMACAT fostered adoption of its methods by the translation industry by forming a user group of early adopters. The project also co-operated with translation communities who were able to make use of customised versions of the workbench. This allowed the collection of usage data under real-world conditions which were used for further evaluation of the system. The adoption of the CASMACAT workbench by various sectors of the translation community was an important outcome of the project.

2.7 The Consortium

The CASMACAT project brought together capable research groups with strong track records that have been approaching the problem of computer aided translation from different angles.

Copenhagen Business School brought in a leading research group with a long track record on translation process studies, while the University of Edinburgh also has broad experience in cognitive modelling of language processing. Edinburgh and the Polytechnic University of Valencia host leading research groups in statistical machine translation groups, as demonstrated by competitive performance of their systems in open evaluation campaigns and Edinburgh's widely used Moses toolkit, which was extended in this project. Both groups had also previously developed innovative computer aided translation tools. Celer Solutions has been on the forefront of deploying advanced innovative computer aided translation technology in their daily operations.

This is the first European project that brought together the leading groups spanning the whole range from cognitive modelling to statistical machine translation and computer aided translation technology research and development and finally deployment.

3 Tangible Outcomes

Besides a large number of publications and dissemination activities to advance and expand the field of computer aided translation, the project created three important resources that have direct and substantial benefit to many interested parties: (1) experimental data organized with in the CRITT TPR-DB database, (2) the workbench for use by translators and researchers, and (3) the CASMACAT Home Edition as a standalone comprehensive CAT solution with integrated machine translation training.

CRITT TPR-DB database. All experimental data gathered from field trials and major lab studies are organized in a standardized format in the Translation Process Research maintained by CBS. This database allows for large-scale in depth studies of human translation processes and thus serves as a basis to empirically ground the development of computer aided translation tools.

CASMACAT Workbench. Based on the workbench developed by the MATECAT project with the focus on professional translators, the CASMACAT project added its many novel types of assistance. This makes the workbench not only a benchmark open source implementation of these state-of-the-art methods, but also open them up for real world use in actual commercial deployments. The flexibility of the modular design also allows for the integration of the separate front end and back end modules into existing workflows.

CASMACAT Home Edition. To give the power of computer aided translation that exploits machine translation into the hands of individual professional translators, the CASMACAT project developed the CASMACAT Home Edition, a standalone version that can be installed on a single desktop computer. In addition to all the features of the workbench, the Home Edition also includes a machine translation training environment that allows the translators to build customized machine translation system based on personal translation memories.

The partners in this projects aim to continue work on these resources in the future.

4 Science and Technology Results

4.1 Work Package 1: User Interface Studies, Cognitive and User Modelling

This WP presents the empirical foundations for the development of the CASMACAT workbench. A series of experiments and field trials were run to establish basic facts about translator behaviour in computer-aided translation, focusing on the use of visualization options and input modalities while post-editing machine translation. Another series of studies deals with cognitive modelling and individual differences in translation production, in particular translator types and translation/post-editing styles.

A large part of the work relied on data collected in the CRITT TPR-DB (Translation Process Research database), which allows for large-scale in depth studies of human translation processes and thus served as a basis to empirically ground the development of the CASMACAT workbench.

4.1.1 Task 1.1: Post-editing

Using the data collected in the CRITT TPR-DB, a large-scale multi-lingual comparison of translation and post-editing behaviour was conducted to investigate and compare the behaviour of 68 different translators when translating and post-editing six English texts into four different languages: German, Spanish, Hindi and Chinese.

Through the analysis of key-logging and eye-tracking data, the main aim of this research was to evaluate human translators, performance with a view to assess different assistance possibilities for automated translation support. More specifically, this analysis aimed at explaining differences in the production time of Alignment Units (AUs) i.e. sequences of source-target correspondences.

Were as follows. We found that from-scratch translation always takes longer than post-editing. Furthermore, we found that keystrokes are an index of translation efficiency, the more keystrokes are produced the longer it takes to produce the translation. We also found that shifting attention frequently between different areas (target and source text, keyboard) is time-consuming and reduces translation efficiency. Word frequency is inversely correlated with production time; this tendency is more pronounced for student translators.

When we investigated the role of the number of different possible translations for a word, we found that high translation ambiguity only has a slow-down effect in post-editing, not in from-scratch translation. We also found that translation is slowed down if there is a large edit distance between the source and target text. Cross-value (distance of word alignments between the source and the target) only has a limited effect, in particular on post-editing.

4.1.2 Task 1.2: Interactive Translation

When studying the use of interactive translation prediction (ITP), we found that it seems to suit more experienced translators with a more positive attitude towards MT. We assume that these post-editors are also more open-minded to acknowledge and accept ITP suggestions, and thus learn how to make better use of the system in their translation work. A qualitative assessment of some post-editing sessions show that post-editors in their first encounter with the ITP mode often overwrite translation suggestions with an identical text. Presumably, once a (partial) solution for a translation problem has been found, translators (and post-editors) are engaged in a typing process which is difficult to interrupt, even if the suggested MT continuation is identical to what the post-editor had in mind. This behaviour seems to change as they become more experienced and familiar with the translation predictions.

4.1.3 Task 1.3: Translator Types and Translation Styles

As part of this task, we proposed a taxonomy of post-editing styles which – amongst others – include style 1: first read the target text and then refer to the source text before making changes in the target text, and style 2: first read the source text and then proceed to read the target text looking for changes needed in the MT output. The data gathered in the lab experiments suggest that these differences in post-editing styles are related to expertise of post-editors.

We ran three eye-tracking experiments in which participants had to spot errors in machine translated text and click on the erroneous word. Participants had access to the source and the target text in all experiments. In one of the experiments they also saw lines aligning the source and the target words. Two of the experiments used naive bilinguals (non-experts). One experiments used professional translators, who saw lines in half of the sentences, and no lines in the other half.

There was no effect of the presence of visual alignment on overall trial duration, the presence of lines did reduce the time that bilinguals spent looking at the translation; this was not the case for the professional translators. There was also a penalty associated with reading the target sentence first for bilinguals, which was not found for the translators. We also found that translators spent less time overall on their first reading of the Target sentence compared to the Source sentence; the opposite was true for the bilinguals. In terms of movements between sentences, the non-professionals had a lower rate of switching.

We also found some evidence that the cognitive load or task demand was higher for the non-professionals as pupil dilation increased significantly more for them when they read the Target sentence. This consequence of inexperience was more pronounced when the Target sentence was read first and even more so when there were visible alignment lines. In fact the professionals did not show a main effect on pupil dilation between reading the two sentences in each pair, whereas the bilingual participants did. Dilation measures for the professional group were also completely unaffected by the presence or absence of Visual Alignment. Combined, these findings could be considered as evidence that expert translators did not find the overall task as cognitively demanding as the nave translators did.

4.1.4 Task 1.4: Text Type

Whereas the most of the lab experiments and field studies used standard newspaper text, the CFT14 study used a medical text extracted from the EMEA corpus. In contrast to the news text, the medical text is very rich in terminology, segments on the source side as well as on the target side are on average shorter than in the news text. Translations of the medical text tend to be more literal than for news text. We found that the EMEA translations have less lexical variability then the news translations, as measured by the translation entropy. Translations are also syntactically closer to the source text: lower cross-values indicate greater syntactic similarity between the source and the target language.

4.1.5 Task 1.5: Cognitive Modeling

The aim of this task was to investigate the cognitive processes involved in spotting and correcting errors in translated text. We tested three groups of participants: monolingual speakers (who only saw the target text), naive bilinguals and professional translators (see Task 1.3). In general, we found that all three groups have similar detection rates, suggesting that monolinguals and naive bilinguals could be used to do a cheap first pass, correcting obvious mistakes and flagging ones that may need the original source text to rectify.

The following general conclusions about cognitive processes during MT post-editing can also be drawn. We found that participants spent more time on the translated sentence rather than the source, and this effect was enhanced for translations containing errors. The presence of an error resulted in an overall reduction in the number of times participants moved between Source and Target sentences. Another general effect was a reduction in switching if the Source sentence was read first, and this became stronger when an error was detected.

The pupillometry results were remarkably similar to the reading time findings. For all the experiments, a correctly identified error increased mean pupil dilation, and pupil size was larger when reading the Target rather than Source sentences, and there was a significant interaction between these two effects. Perhaps surprisingly, given the increase in saliency and perceptual complexity, the only effect of Visual Alignment on pupil dilation was a minor additive interactive effect for the bilinguals (see Task 1.3 above).

4.1.6 Task 1.6: User Modeling

This task tested whether the events that make up the translation process provide enough information for the individualization of post-editor profiles. By using machine learning models, we were able to not only identify the post-editors' profiles, but also cluster and discriminate between post-editors. We found that more experienced post-editors or post-editors with a more positive attitude towards post-editing are likely to adapt better to ITP than inexperienced post-editors who do not have a positive approach towards post-editing.

4.2 Work Package 2: Interactive Translation Prediction

Work package 2 focused on the basic aspects of translation prediction, i.e., to develop and implement novel models, protocols and search algorithms for rendering the interaction process efficient. Given the nature of this work package, work performed implied in some cases basic research, including the development and implementation of new search strategies, novel ways of user interaction, and less conventional SMT models being applied to ITP. The main goal of this work package were to supply the final user with a more flexible and efficient assistance system, be it by means of multi-modality, by introducing novel ways of interacting with the system, or by improving the quality of the translation completions provided by the underlying ITP system.

Work in this work package was divided into five different tasks.

4.2.1 Task 2.1: Search and machine learning criteria for prediction (M1-M18)

Task 2.1 dealt mainly with the core of the interactive prediction framework. Work was split into three main research directions: to analyse and implement novel error rates, to research novel decision rules, and to analyse prediction directly as a machine learning problem.

Concerning the novel error rates, a stochastic error-correction model was developed and integrated into the CASMACAT workbench. A common problem in ITP arises when the user sets a prefix which cannot be produced by the statistical models, and specific techniques need to be introduced into the ITP system so as to guarantee that any given suffix can be generated. This was tackled by means of stochastic error-correcting techniques. Previous related works used the concept of edit distance, which cannot be integrated into the statistical formulation of ITP. Laboratory experiments were carried out, measuring the effort required for a human hose to generate a high-quality translation, and positive results were obtained. Hence, the resulting stochastic error-correcting model was fully integrated into the CASMACAT workbench.

Concerning the novel decision rules, an optimal decision rule following the Bayes decision theory was developed for the suffix search problem. ITP systems have traditionally used a *maximum-a-posteriori* (MAP) approach to the problem of searching for a good suffix. MAP is known to be an optimal decision rule when our goal is to obtain a suffix that perfectly matches the user's expectations. Unfortunately, current state-of-the-art SMT technologies incur in substantial errors, mainly due to practical approximations assumed in order to reduce the computational complexity. For this reason, ITP does not aim to produce perfect automatic translations but to allow the user to obtain the desired output with less editing effort. An efficient implementation of an alternative optimal decision rule was developed, based on the Bayes decision theory. Laboratory experiments were carried out, and showed that the novel decision rule developed was in all cases at least as good as the traditional decision rule. However, improvements were not decisive, and were mostly inappreciable for human translators. Hence, we decided that such improvements did not justify modifying the core of the ITP protocol, which could have possibly compromised future evaluations of the workbench. Work in this direction was published in a second tier JCR-indexed international journal.

Concerning prediction as a machine learning problem, the main goal was to replace the probabilities in the search graph of an ITP system by other scores computed according to a logistic model, in which diverse features can be considered and weighted according to their importance. However, experiments did not show great improvements over the canonical approach, and the added computation cost rendered the approach impractical in a real world setting. Hence, it was decided not to integrate this method into the CASMACAT workbench.

4.2.2 Task 2.2: Multi-modality in interactive translation prediction (M6-M24)

Task 2.2 dealt mainly with the introduction of multi-modal interfaces into the CASMACAT workbench. In this direction, two different approaches were studied, both using an e-pen device as interaction device: handwritten text recognition as input modality, and the use of e-pen gestures for assisting in the proof-reading process.

Concerning the use of handwritten text recognition as input modality, an e-pen text recogniser was fully integrated into the CASMACAT workbench. A main challenge that emerges when using an e-pen device for post-editing MT and ITP is how to deal with the fact that on-line *hand-written text recognition* (HTR) systems are not error free. This is essential, since robustness is key for users to accept the use of an e-pen system. This is possible by taking advantage of information derived from the ITP process so as to achieve a system-user synergy which ultimately boosts both e-pen accuracy and usability. In particular, the source sentence and the already validated prefix was considered for enhancing the accuracy of the e-pen device. Empirical experiments showed that such information was useful. Hence, e-pen interaction, allowing for context adaptation, was fully integrated into the CASMACAT workbench. Work in this direction was published in a first tier JCR-indexed international journal and in a high impact international conference (CORE A).

Concerning the use of e-pen gestures, a novel text-editing gesture set was developed. Although there is already a ‘de facto’ standard for gestures for proof-reading, the results with state-of-the-art gesture recognisers present high error rates. A straightforward solution to incorporate text-editing gestures was proposed, providing disambiguation from handwritten text, excellent accuracy, and an algorithm that is trivial to implement and runs efficiently in any device. Furthermore, the gestures are easy to remember and reproduce, although only 8 possible commands operating on single words are allowed. This was also integrated into the CASMACAT workbench, and was evaluated in the field trials in the context of proof-reading assistance. Work in this direction was published in a first tier JCR-indexed international journal and in a high impact international conference (CORE A).

4.2.3 Task 2.3: Prediction from active interaction (M13-M36)

Task 2.3 dealt mainly with the development of confidence measures with the purpose of determining which translations or partial translations required human revision. Traditionally, ITP systems assume the user to systematically supervise each successive translation generated by the system. The goal of this task was to provide the human translator with a tool that elucidates potentially incorrect translations, be it sentences or parts thereof. The proposed approach was based on the use of confidence measures, which were fully integrated into the CASMACAT workbench. However, such implementation was evaluated with real users in the second field trial, revealing that, even though users seemed to appreciate such tool, the performance of the provided solution was below their expectations. Hence, this functionality is present in the CASMACAT workbench, but disabled as default. Since Task 4.2 is very related to this one in the technological background, but the results were more promising, part of the effort initially planned for this task was devoted to task 4.2.

4.2.4 Task 2.4: Prediction from parse forest (M6-M18)

Task 2.4 dealt mainly with the adaptation of the ITP paradigm to deal with syntax-based translation models. In this direction, two different options were analysed: developing a fully-fledged hierarchical ITP system, and developing an ITP system based on the search forests produced by a syntax-based translation model.

Concerning the development of a fully-fledged hierarchical ITP system, a working prototype was developed. Although phrase-based translation models are still the most common approach within SMT, grammar-based models are starting to be widely adopted, specially for pairs of languages with very different syntactic structure, such as English-Chinese. In this task, a fully-functional grammar-based ITP system was developed. Such system was evaluated in laboratory experiments on two different translation tasks, yielding significant gains, in a human effort simulation, when dealing with Chinese–English, and less prominently in Spanish–English. However, given the experimental nature of the implementation, which was designed more as a proof of concept than as a fully-fledged system, we decided not to replace the CASMACAT ITP engine, which lies at the core of the CASMACAT workbench. Work in this direction was published in a high impact international conference (CORE A).

Concerning the development of an ITP system based on search forests produced by syntax-based translation models, different methods were developed for finding the best derivation in the parse forest that has the minimal edit distance with respect to the given prefix. It was found that conducting a top-down search in the parse forest is significantly faster than conducting a bottom-up search. Different refinements were proposed, leading to further improvements in response time. However, we did not achieve speeds needed for practical use, so we did not integrate this functionality into the workbench.

4.2.5 Task 2.5: New SMT models for interactive translation prediction (M1-M12)

Task 2.5 dealt mainly with the analysis of novel SMT models for their implementation within the core of the IMT engine. In this direction, two different models were analysed: phrase-based hidden semi-Markov models and finite-state models.

Concerning the phase-based hidden semi-Markov models, the main idea was to model the conditional translation probability assuming that a monotonic translation process has been carried out from left to right in segments of words or phrases. Even though the results obtained in terms or SMT evaluation metrics were competitive, and in certain circumstances even better than state-of-the-art phrase-based systems, work in this direction was not carried on to the final CASMACAT workbench because of the high computational cost involved. Given this issue, potential improvements achieved in translation quality might not justify the time required to produce the output, which could render the system seemingly unresponsive.

Concerning the finite-state models, a similar approach to the log-linear models in state-of-the-art phrase-based models was followed. However, a monotone translation approach was adopted because of theoretical restrictions. Even though the results achieved by the finite-state approach were comparable to those obtained by monotone Moses, the comparison with a fully-fledged, state-of-the-art SMT system with reordering was less convincing. For this reason, this work was not integrated into the CASMACAT workbench. Work in this direction was published in an international workshop.

4.2.6 Outstanding results

From the tasks detailed above, the outstanding scientific and technological results and those that were integrated into the CASMACAT workbench were:

1. Development and integration of a stochastic error-correction model, allowing the ITP system to account for potentially arbitrary suffixes (T2.1).
2. Integration of an e-pen interface into a CAT workbench. This technological result was twofold, since interaction allowed both competitive e-pen handwritten text recognition and an efficient set of gestures for text editing (T2.2).
3. Active integration was integrated into the CASMACAT workbench. However, the results achieved with the underlying confidence measures were not convincing enough for the final users (T2.3).
4. Development of a prototype for a hierarchical ITP system, yielding promising results (T2.4).

4.3 Work Package 3: Interactive Editing

Workpackage 3 aims at the development of new methods to assist the editing of machine translation output and refining human translations. A core element of these methods is to provide enriched annotations of source and target sentences with information based on statistical machine translation models.

The objectives of Task 3.1 and 3.2 have received wider attention in the research community, including the EU-funded research projects MATECAT (our "sister" CAT project) and QTLaunchpad (a research project aimed at quality translation). To collaborate community-wide on the problem of confidence estimation (now called more commonly quality estimation), the three projects organized a shared task at the Workshop on Statistical Machine Translation (WMT), which takes place under the umbrella of the Annual Meeting of the Association for Computation Linguistics (ACL).

4.3.1 Task 3.1: Sentence-level Estimate of Post-editing Work Effort

The goal of this task is to develop techniques to estimate the amount of effort required to fix automatic translations. The work for this task is based on our earlier research in sentence-level confidence measures. Sentence-level quality estimation (QE) is usually addressed as a regression problem. Given a translation (and other possible sources of information), a set of features is extracted and used to build a model that predicts a quality score. However, the feature sets contain a large number of noisy, collinear and ambiguous features that hinder the learning process of the regression models, e.g., due to the curse of dimensionality. Hence, we developed dimensionality reduction methods based on partial least squares regression.

The key results of the experiments on dimensionality reduction methods are: (1) methods based on a projection of the data are usually more effective than feature selection methods; (2) methods based on partial least squares regression are usually more effective than other methods; (3) the performance-wise ranking of dimensionality reduction methods is to a great extent independent of the chosen learning model; (4) a combination of partial least squares regression projection and support vector machines was the best performing setup in the experiments.

4.3.2 Task 3.2: Word-level Confidence Measures

Word-level confidence measures (CM) are usually formalized as a conventional pattern classification problem in which a feature vector is obtained for each word in order to classify it as either correct or incorrect. This point of view provides a solid, well-known framework, within which accurate two-class classifiers can be derived. The challenges of this approach are to find an appropriate set of features, and to learn accurate classification models. We extend previous word-level confidence approaches in several aspects, including the addition of new features, and the use of a novel classification model based on multidimensional statistical analysis.

Our results showed that features based on the Model 1 lexicon achieve the best performance, followed by those computed based on posterior probabilities from an N-best list of translations.

4.3.3 Task 3.3: Rules from Translation Memory

The goal of this task is to generate translation rules from fuzzy matches from a translation memory, and use them in our statistical machine translation system. Translation memory tools attempt to match the source sentence in stored translations. If an exact match cannot be found, then a so-called fuzzy match is retrieved, i.e., a sentence that most closely resembles the input sentence currently under consideration. In addition to a translation memory lookup, we proceed to a recombination phrase where the mismatching part is replaced with other translations of source words. Viewed from a statistical machine translation perspective, the goal of translation memory is to find a very large translation rule, that combines and competes with other more general rule to produce the most probably translation. We reformat the fuzzy match as a hierarchical phrase rule and use it in the decoder alongside other rules. This approach gives superior performance over traditional translation memory lookups and pure statistical machine translation.

4.3.4 Task 3.4: Visualisation of Word Alignment

We aim to assist translators in their work by highlighting alignments between source and edited translation. This consists in visually highlighting on the computer screen the words or phrases that are currently the focus of the translator's attention, both within the source text and within the translation. The goal is to guide the translator's eye towards the parts of the text they are currently focusing on.

We developed three variants of this idea:

- Highlighting words at the caret position: The idea is that the main focus is the word that is currently being translated. By highlighting the source word that corresponds to the frontier of the constructed translation, the translators can quickly refer to the content that needs to be worked on next.
- Highlighting words at the mouse position: The translator may want to check, which source words correspond to the target words that have been already written. By moving the mouse over the translation, source words are highlighted. This variant is also useful for post-editing.
- Coverage-based word alignment visualization: While constructing the sentence, not only the source words corresponding to the frontier of the translation are highlighted, but also the source words that have been already translated are shaded off. This draws attention to parts of the source sentence that have not been translated yet, and may have been skipped.

4.3.5 Task 3.5: Display Multiple Translation Options

This tasks developed two distinct types of assistance that suggest alternate translation choices to the translator: Translation options in context and a translation option array.

Translation Options in Context: The user may select any source phrase, and the assistance shows occurrences of the phrase in a database of translated sentences (the parallel corpus on which the machine translation system was trained), aligned with their translation in the target sentence. The sentence pairs are grouped by the translations of the phrase, ordered by the frequency of each translation. We this *bilingual concordancer* in the third field trial qualitatively. About half of the searches carried out by translators led to results that were used in their translations.

Translation Option Array: A common problem of translation is to find the right words to express the meaning of the original. The idea behind showing a number of possible translations to the translators is based on the big difference between active and passive vocabulary: speakers of a language can recognize a much larger number of words than they are able to recall when needed. The idea of a translation option display is to show up to 5 translations of input words and phrases, so the translator can glance at them for inspiration. The translation option display also allows interaction. By clicking on words and phrases, these are automatically inserted into the editbox. We first displayed translation options in several rows below the editbox. However, since this display often stretched beyond the visible page, we refined this into a vertical slider display. Having to manually slide the translation option display creates additional effort of the translator, which may be too burdensome and lead to disuse. Given word alignment information, the translation option display should be centered around the words to be translated next, or just translated.

4.3.6 Task 3.6: Authoring Assistance

Work on authoring assistance in the reporting period focused on the development of a paraphrasing method that suggests reformulations of any part of the translation. In the edit area, the user may mark up any sequence of words, and then request paraphrases. These are displayed in a pop-up window and can be selected to replace the marked content.

We explored several methods to generate paraphrases. The search graph can be mined for alternative translation of the corresponding source phrase considered by the decoder. Or, in a purely monolingual approach, synonymous words and sub-phrases could be used to construct an alternative formulation. In both cases, a language model helps finding paraphrases that are fluent in content.

We built a web service which provides a paraphrasing back-end for any web-based application, and finally integrated that service as a component into the CASMACAT workbench.

4.3.7 Task 3.7: Automatic Reviewing

Reviewing of the first draft human translation by a second translator is an important task to ensure high quality translations. We introduced the use of an e-pen as a novel input modality. The motivation for editing by pen instead of keyboard is that the main editing actions — such as deleting and replacing content — involves less typing but more moving around in the document and selecting particular points for corrections that consists of few characters. The use of an e-pen requires hand writing recognition, which we integrated into the CASMACAT workbench alongside a larger editing window for e-pen editing.

We tested the use of an e-pen as an editing device in the reviewing phase of the third field trial. We developed our own hand-written text recognition (HTR) system. However, this performed worse than an off the shelf commercial system. We improved our HTR system by an adaptation techniques. The findings of the field trial do not support the hypothesis that use of e-pen is more efficient that the traditional use of keyboard and mouse.

The main goal of the task was to develop methods that aid the reviewer. We focussed on methods that predicted missing and added content. These methods are based on word alignments between the source and the human translation — unaligned content words suggest unmapped content. We evaluated this method in two ways. Given the reviewing data from the third field trial, we checked if our method predicts the additions and deletions (ignoring substitutions) of the reviewer. Our methods perform better than a basic baseline, but may not be good enough for practical use. Secondly, we carried out a subjective evaluation on English-German translations created by volunteer translators using the volunteer translation platform. We manually checked the flagged missing and added content. While most of the flagged errors were false positives due to alignment errors (mainly due the deficiency of the word alignment method to align German verbs, one-to-many alignments, and unknown words), the system flagged errors that otherwise would have been hard to detect.

4.4 Work Package 4: Advanced Translation Models

The main objective of this WP was to deal with adaptation in interactive translation prediction (ITP). Work in WP4 has been led by UPVLC and is composed of three different tasks, namely, Task 4.1 focused on online learning for ITP, Task 4.2 on active learning for ITP and finally, Task 4.3, focused on domain and user adaptation.

4.4.1 Task 4.1: Online Learning for Interactive Translation Prediction (M1-M24)

Task 4.1 was planned for the first two years of the project (months 1 to 24). During the first year, previous work by UPVLC on online learning was extended, providing a detailed mathematical framework that describes how to incrementally update the parameters of a state-of-the-art log-linear model for statistical machine translation. For this purpose, a set of sufficient statistics is maintained for each of the components of the log-linear combination. One of the key points of our proposal is the use of the incremental EM algorithm to estimate HMM-based word alignments, which are required to generate the word alignment matrices from which phrase pairs are extracted. The proposed techniques are able to process new training pairs in hundredths of a second, allowing us to efficiently add new n-grams to the language model and new phrase-pairs to the phrase table. In addition to this, the log-linear weights are also modified by means of different update rules. To demonstrate the feasibility of online learning to learn the parameters associated to the feature functions of the log-linear model, laboratory experiments were carried out using two translation tasks that have been previously used to report ITP results in previous European research projects, namely, the Xerox and the EU tasks. In all cases, the proposed ITP system was able to incrementally update its statistical parameters from scratch of from previously estimated models in real time. Additionally, we also tested the above mentioned update rules for the log-linear weights using the well-known Europarl and News-Commentary corpora. The majority of the experiments were carried out in an SMT scenario where the translations are generated in a fully-automatic way. In this case, mixed results were obtained, with two of the proposed update rules providing significant positive results. By contrast, preliminary experiments with ITP were not positive. The work on online learning to adjust the weights of the log-linear combination was published in the Pattern Recognition Journal.

During the second year of the project, the experimental properties of the online learning techniques developed during the first year were carefully studied. As it was explained above, previous experimentation was focused on demonstrating the feasibility of the application of online learning to the SMT framework. However, some crucial aspects of online learning were not studied, such as its performance with respect to a conventional, batch learning algorithm, or the impact of update frequency in the system performance. The new experiments allowed us to clarify such aspects. Specifically, experiments carried out using the Europarl corpus showed that online learning is able to achieve comparable or even slightly better results than those obtained by means of batch learning. Additionally, the presented results also showed the strong impact of update frequency in the performance of the ITP system. Moreover, according to the results, the best system should be capable of updating its models in a sentence-wise manner, highlighting the great potential of online learning to reduce the user effort in translation tasks. On the other hand, experiments reported during the first year of the project were executed on corpora of a small or medium size, mainly due to the computational requirements of processing larger corpora. A significant part of the work carried out within Task 4.1 during the second year of CASMACAT was focused on creating the software tools that are required to process large corpora as well as the integration of such tools in the CASMACAT Workbench. On the other hand, work concerning the online learning of the log-linear weights was also extended during this period. As it was mentioned above, initial ITP experiments did not produce the same positive results obtained in a conventional SMT setup. To deal with this problem, the previously proposed algorithms for online learning of the weights were redefined, and some encouraging results were observed by obtaining weight samples from a Gaussian distribution and hence yielding different wordgraphs. Work in this direction led to two different M.Sc. theses.

In spite of the fact that Task 4.1 was scheduled for the first two years, additional work was conducted during the last year of the project. The reason behind that is the great importance of online learning in the experiments with real users that were developed during the third field trial. Specifically, the work was focused on incorporating the online learning techniques developed and tested during the project into a publicly available version of the Thot toolkit for statistical machine translation (https://github.com/daormar/thot). The toolkit was presented in a research article submitted to an international conference. The online learning techniques implemented in Thot were also selected for their integration into the CASMACAT Workbench and tested during the third field trial, receiving very positive ratings from the professional translators involved in the evaluation. In addition to this, a research article on online learning for statistical machine translation has been submitted to the Computational Linguistics Journal.

On the other hand, a new, dynamic phrase table was implemented on the Moses back-end that constructs phrase table entries on demand by sampling available word-aligned, parallel text. In contrast to conventional phrase tables that are pre-computed and cannot be updated easily, the dynamic phrase tables can add new data with very low latency (under a second, usually) and utilize it immediately (this constitutes an alternative approach to the phrase model updating techniques based on sufficient statistics mentioned above). In a static scenario (no data added), the performance of the new phrase table is on par with conventional phrase tables. In terms of speed, the phrase tables actually out-perform the standard "binary" phrase table implementation in Moses. This work was performed crossing three FP7-funded projects under way at the University of Edinburgh: CASMACAT, MATECAT, and ACCEPT.

4.4.2 Task 4.2: Active Learning in Interactive Translation Prediction (M13-M24)

Task 4.2 was scheduled for months 13 to 24 and was focused on developing an active learning (AL) scenario to boost the productivity of ITP technology by modifying the behaviour of the ITP system in two different directions. On the one hand, we do not require the human expert to exhaustively supervise all translations. Instead, we propose a selective interaction protocol where the user only supervises a subset of informative translations. This selective supervision is similar to the active interaction protocol studied in Task 2.3 in WP2 but applied at the sentence-level. On the other hand, we replace the batch MT model typically used by ITP systems by an incremental SMT model, and more specifically, the one developed in Task 4.1.

The potential productivity improvements of our proposal are twofold. On the one hand, user effort is focused on those translations whose supervision is considered most informative. Thus, we maximise the utility of each user interaction. Additionally, we also define an implicit upper bound to the amount of human supervision effort which is a very useful feature in practice. On the other hand, MT models are continually updated with user feedback. Thus, they are able to learn new translations and to adapt their output to match the users preferences; which prevents the user from making repeatedly the same corrections. The specific implementation details of the proposed AL scenario have been published in an international conference (European Chapter of Association for Computational Linguistics) and in an international journal (Pattern Recognition Letters).

Results of our AL scenario in comparison to a similar scenario with no MT model updating showed a large leap in productivity. In fact, we were able to obtain twice the translation quality for the same human effort. Moreover, the proposed AL scenario allows us to adjust the behaviour of the system between a conventional automatic translation scenario where the human expert supervise zero translations and an online learning scenario where all sentences are supervised by the user. Thus, we can adapt our system according to the requirements of the task or the amount of available resources so as to reach an adequate trade-off between human effort and expected quality of the generated translations.

In spite of the fact that Task 4.2 was not active during the third year, the strong potential of active learning to reduce the user effort of ITP systems led us to conduct complete user effort studies testing this feature. In such studies, the overall productivity of the CASMACAT Workbench incorporating AL was compared with that of ITP with and without online learning.

4.4.3 Task 4.3: Domain and User Adaptation (M13-M30)

In the first year of this task, extensive work was conducted in terms of developing both the theoretical framework and the practical implementation of Bayesian predictive adaptation. Under this framework, model parameters are viewed as random variables having some kind of a priori distribution. Observing these random variables leads to a posterior density, which typically peaks at the optimal values of these parameters. The benefits of this approach are that the parameters are biased towards the optimal values according to the adaptation set, while avoiding over-training towards such set by not forgetting the generality provided by the training set.

Furthermore, re-estimating the systems parameters from scratch may imply a computational overhead which may not be acceptable in certain environments, such as SMT systems configured for post-editing or ITP, in which the final human user is waiting for the translations to be produced. Work in this direction led to part of a Ph.D. thesis and to a research article currently submitted to the Computer Speech and Language Journal.

During the last year of the project, the bulk of the work conducted was focused on developing and comparing sentence selection strategies for the purpose of optimising system training time, and also for SMT model adaptation. To this end, the sentence selection strategy developed in CASMACAT (infrequent n-grams selection) was compared with a state-of-the-art sentence selection strategy based on cross-entropy (which is actually the one used in the MATECAT project). In the experiments reported, infrequent n-grams achieved better performance than cross-entropy.

On the other hand, during the last 6 months, work carried out on Task 4.3 has been focused on testing different domain adaptation techniques that have been reported as successful in the literature, including language model interpolation and sentence selection. The details about this work are shown in Deliverable 4.3. Part of the work carried out on Task 4.3 during this last year has included a collaboration with the MATECAT project proposing a new domain adaptation technique combining components developed by UPVLC from CASMACAT and FBK (Fondazione Bruno Kessler) from MATECAT resulting in a joint research paper that will be submitted to a first level conference.

Since the results obtained with the sentence selection strategies were not conclussive, adaptation in the CASMACAT framework was conducted by adapting the log-linear weights by means of discriminative ridge regression, proposed in Task 4.1, and adaptation of the log-linear features was conducted by means of the online learning techniques also developed in Task 4.1.

4.4.4 Outstanding Results

The most important results obtained in Work Package 4 include the following:

1. Design and development of online learning techniques for SMT, allowing us to update the features of the log-linear combination and their corresponding weights (T4.1).
2. Release of the publicly available Thot toolkit (https://github.com/daormar/thot), which implements a fully-fledged SMT system including the above mentioned online learning techniques (T4.1).
3. Active learning was successfully applied to the field of SMT, defining a new translation framework in which the system decides which translations require supervision from the user. The supervised translations are used to continually update the SMT system models (T4.2).
4. Techniques to perform Bayesian adaptation. These techniques allow to adapt model parameters according to an adaptation set while avoiding the problem of overfitting (T4.3).

4.5 Work Package 5: Integration

The objective of WP5 is to integrate the interactive translation system with user interface and to develop the CASMACAT workbench. Three CASMACAT prototypes have been developed, one by each year of the project: the 1st prototype was developed completely by the CASMACAT team; the 2nd prototype was the merging of the tool with MATECAT interface and the integration of the interactive machine translation; and the 3rd prototype was an extension of the 2nd prototype with additional features. This work package was divided into nine different tasks.

4.5.1 Task 5.1: Specification of Requirements

The CASMACAT Workbench allows users to enter documents in a source language, and then receive assistance in translating them into a target language. The assistance is based on information from machine translation systems. The implementation languages have been defined at an early stage of the project (D5.1). We defined an API that resembles that of Google Translation but adapted to the needs of CASMACAT including word segmentation, word and sentence confidence measures, alignment information, translation options, and specific interactive text prediction and handwriting recognition APIs. The GUI Server this will be PHP running on Apache and HTML with JavaScript using the JQuery engine, whereas the back-end is a python socket.io server that connects with the IMT servers. The MT and IMT engines themselves are mostly written in C++. The database is in MySQL database for all components using a database. The browser extensions are implemented in C++.

4.5.2 Task 5.2: Graphical Interface

The interfaces of MATECAT and CASMACAT were merged, the MATECAT code has been adapted and extended with the CASMACAT features. The GUI consists of a couple of views designated to different tasks (translation, upload files, list documents and replay) and different optional features (interactive machine translation, visualization options, research and replace, translation options tables, biconcordancer and e-pen) as shown in Figure 1. A publication in a conference was written to decide what visualization features suite better and a publication in a journal was written to compare the conventional post-editing with the interactive machine translation.

4.5.3 Task 5.3 E-pen Interaction

E-pen interaction is regarded as a complementary input rather than a complete replacement of the keyboard. We have extended the CASMACAT UI with the necessary components to enable e-pen gestures and handwriting. On the one hand, the HTR server is responsible for decoding the user handwriting into digital text. The HTR server is based on iAtros, an open source HMM decoder. The online HTR recognizer an retrieve can arbitrary number of alternative transcriptions of the user's handwriting. Then, the user is able to select one of the alternatives to replace the original recognition. On the other hand, MINGESTURES is a gesture recogniser specialised in simple gestures for text editing. MINGESTURES runs in the browser as a javascript library, which makes it very responsive. All this work has been implemented with the technology developed in Task 2.2 in WP2. Furthermore, the e-pen mode has been tested in the 3rd field trial in the reviewing phase, as it can be more comfortable to use for a small amount of corrections. More details on such evaluation are described in Task 3.7 in D3.3 as a tool to help reviewers.

4.5.4 Task 5.4 Logging Functions

The user activity data is logged for cognitive analysis and user modelling (to be used in WP1). The logs contains all necessary information to fully reproduce a translation session. The logging information includes the configuration information, source, initial and final translation, the GUI interaction as well as time-stamped activities of the translator, including gaze movements, keystrokes, mouse and e-pen activities. The logging information is kept in the database and an XML file can be downloaded for an easier process of the analysis. All the data gathered in this way is available in the TPR database. For more information see D1.4.

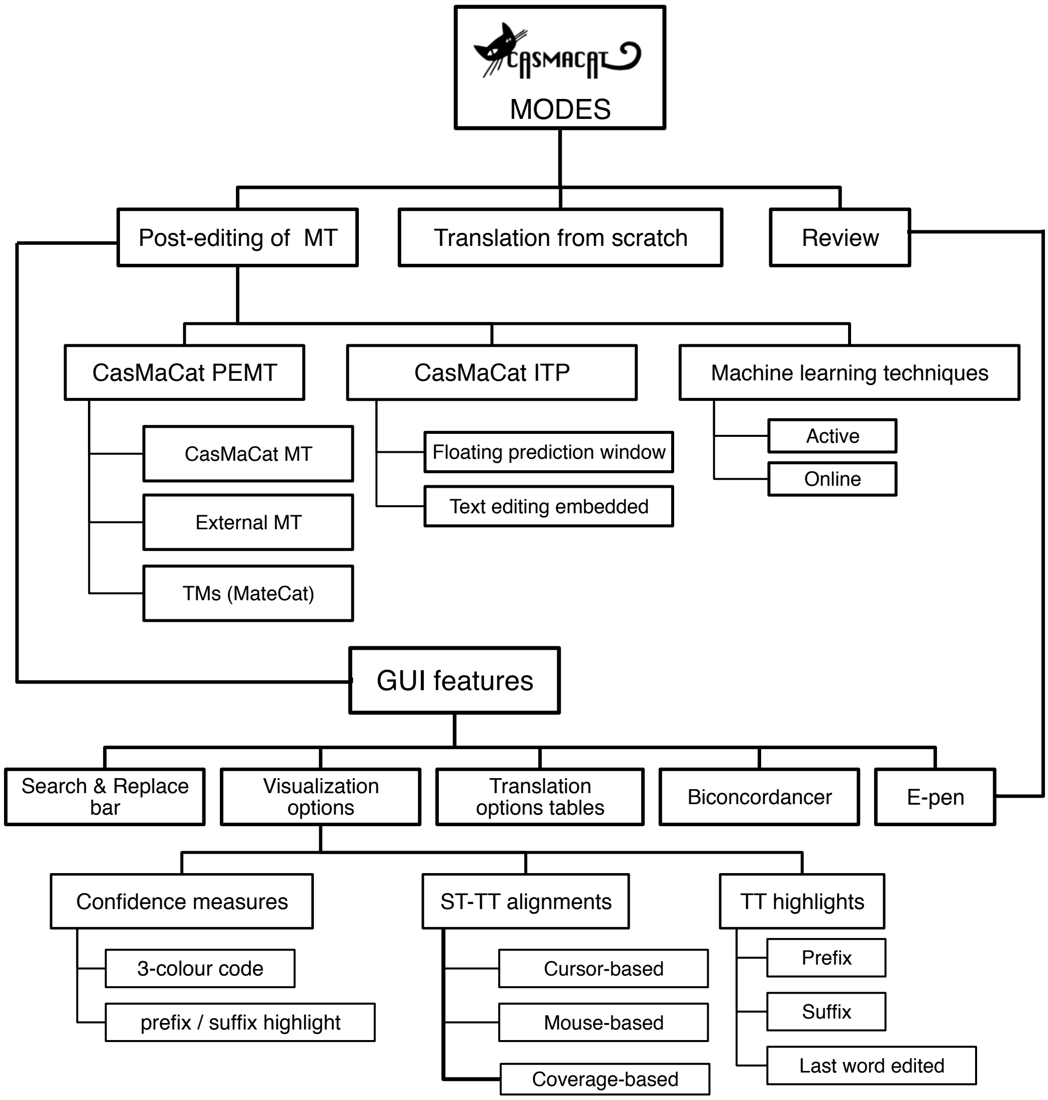

Figure 1: Different options of configuration in CASMACAT

4.5.5 Task 5.5: Machine Translation Server

The CASMACAT server connects to the open source Moses or Thot machine translation systems. It offers an API to request translations and translation options in real-time, together with their translation confidence scores and alignment tables. The API also receives information on the translators modifications of the translations so as to inform the underlying MT system about the changes and to enable online learning of its resources. Within the CASMACAT, we bring together two machine translation systems, Thot and Moses.

Additionally, the home edition of CASMACAT gives to the user the ability to install the workbench on their own machine.

4.5.6 Task 5.6: Manual Gaze-to-word Alignment

A tool has been implemented for aligning and correcting erroneous fixations so as to manually map them on the words that are likely to be fixated. The tool permits to create a corpus of high-quality gaze-to-word mappings, and can be used as a training set of automatic gaze-to-word mappings algorithms. The TPR-DB provides a set of experiments with manually re-mapped files.

4.5.7 Task 5.7: Automatic Gaze-to-word Alignment

We have implemented two algorithms for automatic gaze-to-word alignment. The manually re-mapped fixations were used to evaluate the precision of these algorithms. The algorithms and their evaluation are included in two publications.

The 3rd CASMACAT prototype includes a new function for automatic gaze-to-word alignment. The conversion of gaze data to screen position is handled by the browser plug-in and then sent through the browser to be caught by JavaScript, which is then able to decipher the screen position, converting it into a caret position. In addition to the characters found from the gaze data, two additional possible candidates of Gaze-to-Word Alignment are also calculated by the JavaScript algorithm returning the character just above and below the character. Such candidates are used for further disambiguation during the replay view.

4.5.8 Task 5.8: Replay Mode for User Activity Data

The replay mode allows the user to select or rewind a segment in a translation session, plotting each activity of the translator, including a visualisation of their gaze movements and marking the fixated words. The replay mode is instrumental for qualitative analysis of the translation sessions and for retrospective interviews, allowing translators to recall and comment their own translation sessions.

The replay view loads the CASMACAT interface into an iframe and remote-controls it with the data from the log file. Replay View provides gaze-to-word alignment visualizations. Gaze-to-word alignment visualizations display the information gained thanks to the automatic gaze-to-word alignment algorithm. This is visualized by three coloured boxes (magenta, yellow and cyan) surrounding the characters in focus. These boxes are clickable, so that the user can choose to change the golden standard character if necessary (e.g. to fix driftings). The golden standard is displayed with a green border.

4.5.9 Task 5.9: Visualization of Translation Processes

The Translation Process Research Database (TPR-DB) allows for plotting translation sessions in the form of translation progression graphs. However, translation progression graphs only visualize a small fraction of the information that is contained in the TPR-DB. A paper has been published with an overview of the features that are contained in the database and their current visualization possibilities.

4.6 WP6: Evaluation

The goal of work package 6 (WP6) is to study the impact of the innovative approaches implemented in the workbench. To do that, we carried out a series of experiments where the workbench is used by human users in different real-world environments.

4.6.1 Task 6.1: Field Trials at Translation Agency
Task 6.2: Analysis of Translator Feedback and Activity Data

We carried out three annual field trials of the CASMACAT workbench at Celer Soluciones SL. These generated user activity data and machine translation output with edits that were essential data resources for methods developed in other work packages. Additionally, we have conducted two preliminary field trials to evaluate different features of the workbench. The goal of these preliminary trials was to decide which of the features were mature enough to be tested in the official field trials.

Next, we describe, in chronological order, the different preliminary and official field trials carried out during the project. We provide a brief description of each trial and a summary of its results and the lessons we learned from it.

First Field Trial (June 2012)

Description The goal of the first field trial was to identify, analyze and consolidate the needs of translation professionals with regard to post-editing services with a view to improving the CASMACAT workbench. The main aspiration was to bring the CASMACAT workbench prototype closer to humans by providing constructive input that will enhance software engineers' understanding of the operational context of the system. The system tested by the users was a very limited baseline version of the tool that allows for post-editing machine translation.

The method selected in this study was a semi-structured interview with each participant since it was the type that best met our needs. For the structured part of the interview a series of predefined questions were asked to each participant, so all respondents could cover the same issues. The interviewers used a standardized interview schedule with set questions with were asked to all respondents. The questions tended to be asked in a similar order an format to make a form of comparison between all answers possible. However, there was also scope for pursuing and probing for novel, relevant information, through additional questions often noted as prompts on the schedule. The interviewers frequently had to formulate impromptu question sin order to follow up leads that emerge during the interview.

Results The interviews revealed a great number of specific needs concerning functionality mainly concerned to the post-editing editor and its basic functions. According to the interviewees, any post-editor workbench should be a tool to automatize most of the repetitive and tedious editing work that fixing MT outputs involve. Post editors often reported frustration and disappointment when having to go over and over again fixing the same kind of errors withing the same project.

The three key concepts corresponding to a post-editing tools appeared to bear the greatest importance to post-editors, as they emerged again and again throughout the interviews: flexibility, simplicity and ease of access.

- *Flexibility* seems to be attained when the system offers the possibility to the user of customizing and defining the settings for most of its features. This way it can be adaptable to each user's personal needs. It is also achieved by through the option of enabling or disabling certain features.
- *Simplicity* is perceived in different ways by different user groups. However, all post-editors seem to agree on the idea that simplicity can be achieved by organizing all features under the umbrella of a single simple process that supports the post-editing work-flow from the beginning to the end.
- *Ease of access* related to the affordability of the system, not only in terms of purchase cost, but also in terms of upgrade, support and training costs.

Preliminary Field Trial (March 2013)

Description The preliminary field trial study was performed before the second CASMACAT field trial. It enabled us to decide on the elements and configuration of the second prototype of the CASMACAT workbench before running the second CASMACAT field trial. For the purpose of evaluating the visualization of advanced ITP features, four different configurations of the workbench were tested:

PF1: basic intelligent autocompletion (ITP)

PF2: ITP + confidence measures (active interaction)

PF3: ITP + limit suffix length

PF4: ITP + search and replace + word alignments + prediction rejection

System PFT1 was a baseline system for ITP including only basic intelligent autocompletion. Systems PFT2 to PFT4 included the intelligent autocompletion feature (ITP) together with some advanced features. The main goal of this research was to measure user satisfaction when performing post-editing tasks using different workbench features. In this context, we were interested in knowing whether translators find such features useful while post-editing MT outputs.

Results Five participants were each asked to interact during two hours with the CASMACAT prototype across the four configurations described above post-editing MT outputs from English into Spanish (the same language pair that was going to be tested in the second field trial). At the end of each session, they answered a questionnaire about user satisfaction and post-editing views. Two major findings emerged from these experiments:

- The workbench should be able to identify when the translator edits back from right to left in order not to trigger further ITP suggestions that could change some of the editing work already done by the post-editor. Based on this observation, the UI was instrumented to keep track of which words had been modified by the user and, then, they should not be modified by the automatic ITP suggestions.
- All of the five post-editors agreed that it would be useful to be able to enable or disable ITP at their convenience. Based on this request, the button ITP was added to the GUI in order to be able to work with or without ITP at the segment level.

The most popular feature in this configuration of the CASMACAT workbench was *limit suffix length* and the least popular was *confidence measures*.

Second Field Trial (June 2013)

Description The second field trial of the CASMACAT workbench aimed at testing the second prototype of the tool based on the finding of both the first field trial in June 2012 and the pilot test carried out in March 2013. After completing an introductory questionnaire to collect metadata (age, gender, training, years of experience, expertise in post-editing, etc.), participants were introduced to the second prototype of the workbench featuring interactive translation prediction (ITP). This second field trial involved three tasks: 1) traditional post-editing (P), 2) basic ITP (PI) and 3) ITP with advanced visualization features (PIA). Post-editing guidelines were also provided as well as a detailed description of the characteristics of ITP which they would see implemented in either task 2 or 3. In order to collect user feedback, a series of interviews were also conducted at the end of these three tasks.

Based on the functionalities requested by the participants in the first field trial, the following extra features were included in the second CASMACAT prototype.

- Visual tracking of changes (through interactive machine translation features)
- Monitoring the post-editing progress by real-time word counters and progress bars
- Systematic search and replace
- Copying text from source to target segments
- Translation memory (TM) module as a result of the collaborative work done with the researchers of the MateCat project.
- Autowrite functions in the form of ITP (this feature was the independent variable of most of the analyses during this second field trial).

Results Productivity results for the second field trial showed that enabling interactivity has an effect on the number of insertions and deletions which the post-editor makes. According to the results, the PIA system required less operations than the rest of the systems, but only after removing from the overall task times pauses above 5/200 seconds. Another outstanding result to take into account when analysing the average post-editing times in this field trial was the learning curves of each system. While traditional post-editing (P) was typically well-known by translators, this was not the case for the other two systems featuring interactivity (PI and PIA). Despite the differences in performance, all participants improved productivity rates after more hours of interaction with the interactive prototype.

The results from the final interviews showed varying levels of satisfaction with regard to the different systems. Three out of the nine participants in this field trial claimed to be more satisfied with the interactive systems (PI and PIA) than with traditional post-editing alone (P). However, the rest claimed that they would need more time interacting with the tool in order to fully benefit from post-editing in an interactive environment. For this reason, an extra longitudinal study was conducted between April-May 2014 (see section 4.6.1).

Preliminary Field Trial (March 2014)

Description Prior to the third and final field trial at Celer Soluciones SL, another pilot test was conducted. A group of five participants volunteered to take part in this pre-field trial pilot post-editing from English into Danish. Thinking about which should be the final CASMACAT configuration that should be tested in the final field trial, three conditions were tested in the framework of this pre-field trial: i) ITP, ii) ITP with online learning (OL), and iii) ITP with active learning (AL). The two parameters used to analyse the user activity data collected in this pre-field trial pilot were: 1) *Speed*: total number of words translated divided by time in minutes. 2) *Technical effort*: total number of edits done by the participant divided by the number of translated words.

Results Results showed how OL techniques significantly improved translation speed (about 2.5 more words translated per minute). Regarding the number of keystrokes, results are not consistent: no significant difference was found for the two conditions involving OL/AL. Based on these results evaluating both OL and AL, the decision was made to further test the possibilities of ITP with OL in the context of third CASMACAT field trial.

Longitudinal Study (April-May 2014)

Description Based on the results and feedback collected in the framework of the second field trial, the decision was made to conduct an extra longitudinal study where participants were tasked to work with the CASMACAT workbench featuring interactive post-editing over a period of six weeks. The aim of this study was primarily to find out whether professional post-editors improved performance over time.

Five professional translators were recruited by Celer Soluciones SL to take part in the study. The experimental design involved 24 different source texts which were post-edited from English into Spanish over a period of six weeks (four texts per week). MT was provided by the CASMACAT server and the participants were asked to work under two conditions: i) traditional post-editing (P), i.e. no interaction is provided during the post-editing process, and ii) interactive post-editing (PI), i.e. interaction is provided during the post-editing process in the form of ITP.

Every week, post-editors worked in parallel on the same 4 source texts counterbalancing texts/conditions among participants in order to avoid any possible text/tool-order effect. During the first and the last week of the study, post-editors worked from Celer Soluciones SL while their eye movements were recorded using an eye-tracker. From week 2 to week 4, post-editors worked from home as they usually do when completing jobs for the company.

Results With respect to the main aim of the LS14 study, when investigating the production times over a period of six weeks interacting with ITP in the CASMACAT workbench, it was observed that post-editors become substantially quicker in the PI condition over time, while in the P condition (baseline) no significant change in *Kdur* effort was observed. In order to explain the drop in productivity for week 6 (the last week of the longitudinal study), a distinction was made between working from Celer Soluciones SL (weeks 1 and 6) and working from home (weeks 2-5). Making projections based on the data collected for a timeframe of 12 weeks (twice as much as the actual longitudinal), it is between weeks 9 and 10 that post-editors would become more efficient under the PI condition than under the P condition. While this is a hypothetical assumption, assuming a linear relationship between time spent working on the CASMACAT workbench and *Kdur*, this projection clearly shows a learning effect for the PI condition, which is absent in the P.

Third Field Trial (June 2014)

Description The main research aims of this third and final field trial were: i) To measure the productivity benefits derived from introducing online learning techniques during the post-editing process. ii) To assess how professional reviewers use the e-pen functionalities while reviewing from CASMACAT. This third field trial involved seven post-editors and four reviewers.

Two texts were involved in this third field trial. As in the case of the pre-field trial pilot study, the type of text involved in this third field trial was domain specific, i.e. medical specialised texts from the EMEA corpus (package leaflets for schizophrenic patients). They were approximately 4,500 words long comprising 131 and 141 segments respectively and they were pre-translated into Spanish by a SMT system and then loaded into the workbench for the participants to post-edit. In order to assess and compare the effects of enabling interactivity and online learning techniques, each participant post-edited two texts each under one of the following conditions: *Condition 1*: Traditional post-editing with no assistance during the process (P). *Condition 2*: Post-editing through ITP featuring on-line learning (PIO).

Results In order to measure the productivity benefits derived from introducing online learning techniques during the post-editing process, the amount of technical effort (i.e. the number of insertions and deletions needed to correct the raw MT output) was calculated for the two conditions. On the one hand, participants working under condition 1 (traditional post-editing) deleted 70.71 keystrokes and inserted 79.53 on average. On the other hand, post-editing with interactivity and on-line learning techniques made keyboard activity decrease as they inserted 68.73 keystrokes and deleted 36.94 on average. If we compare keyboard activity for both conditions, it can be claimed that there was a significant decrease in the number of insertions and deletions comparing both conditions. Since both texts were comparable in size and difficulty, this significant decrease in technical effort (typing) must be attributed to the benefits of online learning techniques during the post-editing process. Results with respect to user satisfaction when working with online learning techniques clearly show that all participants supported the integration of online learning techniques as they did not have to do repetitive work editing MT outputs and they could see their edits populated from segment to segment.

The results for the e-pen evaluating showed that e-pen was slower than keyboard on a segment level, but the differences could be acceptable. Unfortunately, e-pen recognition was perceived as not very accurate, causing many retries which made this condition less productive. On the positive side, e-pen gestures were appraised by the reviewers, in spite of minor problems related to font sizes. These results point out that productivity differences could be reduced to the minimum if the UI was redesigned and implemented in a tablet with e-pen support, and the accuracy of the recogniser was improved.

Additional Extra Field Trial (September-October 2014)

In addition to the field trials previously described, and given the interesting results of active and online learning in the 2014 preliminary field trial, we consider to extend such experimentation with additional users and a new language pair. The experimentation was conducted at the premises of Celer Soluciones SL. A group of eleven participants volunteered to take part in this pre-field trial pilot post-editing from English into Spanish. Three conditions were tested: i) ITP ii) ITP with online learning, and iii) ITP with active learning.

Data is being analyzed at this moment, so we do not have final results so far. Nevertheless, preliminary results are encouraging and seem to indicate similar conclusions as for the 2014 preliminary field trial, that online and active learning allow to improve the productivity of conventional ITP.

4.6.2 Task 6.3: Integration into Community Translation Platforms
Task 6.4: Collection and Release of Volunteer User Activity Data

Since the CASMACAT workbench is implemented as a web-based service, it is relatively straightforward to integrate it into web-based community translation platforms such as Global Voices[1]. These platforms emerge out of the volunteers desire to localize foreign language content to increase the information sources available to their linguistic community or to spread information about their native country to the outside world. Participating volunteer translators may also be motivated by the chance to improve their language skills. We will offer the workbench as a customized free service to such platforms in exchange for the right to collect and analyse user activity data. In addition to Global Voices, we have also integrated the CASMACAT workbench into TED talks[2] and Wikipedia[3].

Next sections describe the integration challenges for each of the platforms, and resume the volunteer user data collected so far.

[1] `http://www.globalvoicesonline.org/`
[2] `http://www.ted.com`
[3] `http://www.wikipedia.org`

Global Voices The news stories on the Global Voices website consist of a title and a main body broken up into paragraphs. Stories include hyperlinks and embed tweets, images, and videos. This requires the proper handling of tags. Since the tags will be identical for the original and the translated version their content has to be preserved and can be hidden from the translator. Quotations are handled in a special blockquote format, which preserves the quote in the original language alongside the translation.

Global Voices has a story editor, which is identical for original authoring of stories and translation of stories. The editor is a HTML textarea which uses a format similar to many Wikis. Alternatively, a WYSIWIG environment can be used. Volunteer translators are most likely interested in translating the most recent news stories that have been published. Hence, we offer on the opening menu of our support site a choice of recent articles in reverse chronological order (newest first). This list is based on a regular crawl of each of the language editions of the Global Voices web site.

The volunteer translator may click on the story title to see the original page, or click on "translate" to enter the CASMACAT workbench and work on translation of a story. In order to translate an article, the translator also has to enter an email address, which establishes a user account. There are no passwords or any further validation of the email address. Once a translator has entered an email address and started translating news stories, the main menu also displays a personalized list of news stories. In this list, in addition to links to the original story and the CASMACAT workbench, a third link allows the conversion of the translation entered into the CASMACAT workbench into the format used by the Global Voices content management system.

In order to support the different formats involved in this platform (XLIFF format of CASMACAT and Wiki/HTML of the Global Voices web site), we built the following converters:

- HTML → Global Voices editor format
- Global Voices editor format → XLIFF
- XLIFF → Global Voices editor format

The XLIFF format also encodes meta information about tag contents, paragraph breaks, blockquotes, tweets, etc. This allows us to present in the CASMACAT workbench the text of the news stories, while ignoring these formatting issues as much as possible. Tags are just presented by their tag names, so a complex hyperlink becomes a simple `{a}` and `{\a}` markup.

Ted Talks The integration of the CASMACAT workbench with the existing translation tools for TED-Amara platform follows a similar pattern as the previously described integration of Global Voices. Talks for translation are associated with a translator via email address. Entering the email address gives a list of translation tasks in progress or completed, as well as a form to start a new translation task. Data between the TED-Amara platform and the CASMACAT workbench is exchanged in a text format. Specifically, once the translator decides to work on the translation of a talk, she proceeds with the following steps:

1. Download the original language subtitles in text format.
2. Upload the text file to the CASMACAT community translation platform. (involves automatic conversion of the text into XLIFF format)
3. Translate the subtitles with the CASMACAT workbench.
4. Download the translations from the CASMACAT workbench as a text file. (involves automatic conversion of the XLIFF into text format)
5. Upload the text file with the translations to the Amara platform.

Wikipedia Wikipedia uses a wiki format, similar to Global Voices, but much more complex since it also allows for hierarchical document structure, tables, and special information boxes. Translation of content for this platform requires clear distinction between functional elements and content, and the proper preservation and adaptation of the former. A particular issue is the translation of links since they link to other Wikipedia articles. The translation of a Wikipedia article should then point to a Wikipedia article in the translated language instead of the original language.

Wikipedia does not expect that articles are created through translation, but rather authored independently. Hence, there is no integrated support for translators. It does encourage translation of English articles into other languages, but does not prescribe any specific tool.

The integration of Wikipedia as a CASMACAT community translation platform is modeled very closely to Global Voices due to similarities in document format (marked up wiki) and edit interface (HTML textbox). Since English coverage is very good, we limit ourselves to translation from English to other languages (German, Danish, Spanish, French, Greek, and Portuguese). As a back-end, we currently use the machine translation engines as for Global Voices and, instead of providing a list of recent articles, we show the currently most popular articles in English that have not yet been translated into other languages.

The user may also enter the name of any other Wikipedia article to be translated. The more complex format of the wiki markup requires more dedicated methods for conversion of the wiki format into XLIFF and back. We directly convert from the wiki edit format, which is available by crawling the edit links.

Volunteer User Activity Data The integration of the Global Voices platform was finished first, and has been promoted to the Global Voices translator community, especially the translators working on translation into German, who requested additional language pairs (Portuguese-German and Spanish-German). The total number of sentence translation requests are summarized in the next table:

Language pair	en-da	en-de	en-el	en-es	en-fr	en -pt	es-de	es-en	fr-en	pt-de
Translations	43	287	150	86	22	23	4	33	42	10

Usage is currently still relatively light, but given that the support will remain available after the project is expected to increase.

5 Potential Impact

The European project of bringing together the nations of Europe in a ever increasing political, cultural, and economic integration faces the challenge of dealing with and preserving the linguistic diversity of the continent. Huge amounts of texts need to be translated constantly to meet the communication needs of the quilt of European linguistic communities. Even with improvements in machine translation technology, high quality translation requires the involvement of human translators, which is a scarce and costly resource.

Improvements in the quality of automatic translation in combination with productivity increases in computer aided translation technology will lead to higher-quality multilingual communication. This will help promote a European language policy that transforms the potential liability of linguistic diversity into a strong, creative asset by overcoming communicative barriers even for under-resourced languages.

Europe's competitiveness and cultural diversity depend heavily on our ability to communicate cross-linguistically in high-quality language produced at low cost. The European Commission translates more than 1000 pages per day. Traditional translation performed by human translators has no longer the capacity to meet the gigantic needs for multilingual communication at all levels.

The CASMACAT project had an ambitious agenda to cover the whole range from basic scientific research through applied research, development, integration, and deployment of the technology for daily use. Key to this agenda were various dissemination activities and the principle of releasing the project outcomes freely as open source software and resources.

5.1 Dissemination Activities

Besides publishing a very large number of research papers and presenting the work to academic conferences and workshops, the project organized a number of events to push the scientific, economic, and societal goals of the CASMACAT project. The events were planned to have wider impact beyond the immediate objectives of the project.

We organized the following events:

- Symposium on Eye tracking in Translation and Post-editing Studies at the Tobii conference EyeTrackingBehaviour 2012
- Workshop Eye tracking and Natural Language Processing at COLING 2012
- Workshop on Speech and Gaze in Translation, 2013
- Special symposium on Empirical approaches to gaze data analysis in reading, writing and translation at 17th European Conference on Eye Movements, 2013
- Tutorial on Empirical Translation Process Research at NLPCS 2013
- DGA workshop on Multimedia Information Processing 2013
- Workshop on Humans and Computer-assisted Translation at EACL 2014
- Workshop on Translation, Bilingualism and Translation Technology, 2014
- Workshop on Interactive and Adaptive Machine Translation at AMTA 2014
- 1st and 2nd Workshop on Future Directions in Translation Research, 2012, 2014
- Third and Fourth PhD Course in Translation Processes Research, 2013-2014
- Summer project on Translation Data Analytics, 2014
- 7th-9th Workshop on Statistical Machine Translation, 2012-2014, especially co-organizing shared talks on quality estimation
- 7th-9th Machine Translation Marathon, 2012-2014

We also reached out to end users of the CASMACAT workbench by forming and engaging a user group, and integrating the workbench with community translation platforms. We discuss the purpose of these activities in the next sections.

5.2 Scientific Impact

While statistical machine translation has evolved into a large research area with 400 publications per year and there is an established research field on translation process study, the connection of the two is only now gaining momentum. The CASMACAT project was instrumental in creating this momentum, by presenting the challenges of human interaction with machine translation to the machine translation community and by presenting the capabilities of interactive machine translation and other technologies to translators.

We organized events in Europe, the Americas, and Asia to reach a global audience of researchers. Some of the events were co-organized or at least co-ordinated with related EU projects, most notably MateCat, ACCEPT, and QT Launchpad.

In a number of events, we presented capabilities of eye tracking to the natural language processing research community — both to researchers interested in natural language processing and in translation process study. Some of these events were very focused, such as the PhD course on translation process research or the special symposium on empirical approaches to gaze data analysis in reading, writing and translation.

Within the context of the annual Workshop on Statistical Machine Translation, we organized the problem of confidence estimation (also called quality estimation) as a shared task for the research community, which led to many research groups taking up this task. Having a shared task with established benchmarks allows for better comparison of different approaches, and most importantly gives credibility to new approaches that do well in the competition.

To prepare the organization of more tasks within the area of interactive and adaptive machine translation in shared tasks with common data sets was one of the goals of staging the Workshop on Interactive and Adaptive Machine Translation in the final year of the project. Together with the MateCat project, we brought together researchers that have worked on new tools for translators in this space.

Also with the open source software that was released by the CASMACAT project, we expect to have started the research community on a path similar to one that supports statistical machine translation at this point: access to baseline systems that allow researchers to develop methods based on the state of the art, common benchmarks that allow for comparable results, and regular shared tasks that bring together the research community and provide a platform for outstanding new ideas and methods. While the emerging research field is still in an early stage, the CASMACAT project created large parts of the required infrastructure for it to flourish.

While we delivered a comprehensive novel workbench for translators, we covered such a large set of pieces of technology that we did not reach for all of them optimal mature solutions. There is a great space of follow-up work which has the potential to build on our work, and we hope that the research community, armed with the infrastructure that we provided, will rise to this challenge.

5.3 Economic Impact

The global translation market is estimated to be in the order of 10 billion Euros a year. Given that large amount of money spent on high quality translation and the reliance on costly translators, any productivity increase has a direct economic benefit. Note that this does not imply that the number of translators will decrease in the future. On the contrary, by lowering the cost of translation and the ever increasing pressures from a globalizing world mean that more content will need to be translated.

While the complexities of carrying out field trials (such as the large variance among translators) make it hard to draw definitive conclusions, we have shown that the developed interactive machine translation methods can lead to higher productivity after a training phase, and that our online and active learning methods also increase productivity. Some of the other technology (such as the bilingual concordancer, the translation option array, and the word alignment visualization — to name just a handful) are harder to evaluate in terms of pure productivity gains, but the qualitative feedback that we received from translators demonstrate that they are successful. So, the overall goal of increasing productivity of translators, and hence reducing the cost of translation has been achieved.

One of the main venues for the CASMACAT project to engage with commercial and governmental potential users are the annual conferences of the International Machine Translation Association, which organizes the AMTA, EAMT, and MT Summit conferences. The CASMACAT project had a strong presence at these conferences, typically with a booth that was open during the entirety of the conference and where we demonstrated the workbench and advertised the overall goals of the project.

Already during the proposal stage of the project, we directly contacted commercial and governmental users to make them aware of the technology that we envisioned in the CASMACAT project. Some of these joined our user group, alongside others that expressed interest in the project during its duration. The presence and visibility at industry events was very helpful in making these connections.

An important aspect of engaging the user group was the installation of customized versions of the workbench for any of them who was interested. We managed to have three such installations, which involve the building of machine translation systems, setting up the workbench on servers of the CASMACAT partners, agreeing on a configuration, and creating an open channel for exchange of feedback and addressing this feedback by refining the setup. For instance, we became more aware of some of the speed bottlenecks when delivering the workbench across the web over large distances (e.g., server in Scotland, user in Luxembourg) and addressed these.

Going forward, we expect that the increased awareness of the CASMACAT workbench among potential commercial users and the availability of the workbench as open source software will lead to broad adoption of the workbench. The efforts of our sister project MateCat, which is quite eager in pushing the technology to market and with whom we share the same open source infrastructure, will also broaden the appeal of the technology developed by the CASMACAT project.

5.4 Societal Impact

We directly engaged with professional translators in various ways, such as invited talks at conferences and meetings of professional translators associations (Imperial College London, TriKonf, ITI ScotNet, etc.). There was great interest in the technology that we developed.

It is worth noting that the attitudes of some translators towards machine translation is quite skeptical because of how the technology is presented to them. They are often confronted with machine translation as a cost saving effort, which quite directly means a reduction of their word rates, which they do not necessarily see justified by the actual benefit that machine translation provides to their productivity.

By establishing the community translation platform, we are directly inserting the CASMACAT workbench into the work of professional translators. The volunteer participants in these platforms are typically younger and more open to novel technology. The direct engagement with the end users of the technology will be very helpful for both sides.

In addition, the development of the CASMACAT workbench as open source software brings us researchers and developers in direct contact with translators who are able to judge the software on its merits outside the typical commercial constraints. We are optimistic that going forward the CASMACAT Home Edition can play a similar role as the open source OmegaT translation memory tool that has already gained significant market share among translators. Merging these projects, or at least integrating them may be a productive path going forward. We have foreseen such a possibility within the CASMACAT project and its modular design enables this. For instance, it would be relatively straightforward for OmegaT to directly connect to the API of the CAT server and MT server of the CASMACAT workbench.

The end effect of the direct connection with translators is that it opens channels between researchers and users. Some technically skilled users may become directly involved in the development of the software, thus driving the development by actual user needs, not only novel technical capabilities that are expected to meet user needs. We expect that such an environment based on open source software leads to a more rapid development of a better workbench and better tools than the current situation which is characterized by the bottlenecks of commercial developers and language service companies who have their own interest which may not always be to the benefit of the translator.

6 Use and Dissemination of Foreground

6.1 Template A1: List of Scientific Publications

No.	Title	Main author	Title of the periodical or the series	Number, date or frequency	Publisher	Place of publication	Year	Pages
1	Minimum Bayes' Risk Subsequence Combination for Machine Translation	González-Rubio	Pattern Analysis and Applications	in press			2015	
2	Building a Better Mousetrap: Compressing Mouse Cursor Activity for Web Analytics.	Leiva	Information Processing and Management	in press			2015	
3	Context-aware Gestures for Mixed-Initiative Text Editing UIs	Leiva	Interacting with Computers	in press			2015	
4	Translating without In-domain Corpus: Machine Translation Post-Editing with Online Learning Techniques	Lagarda	Computer Speech & Language	in press			2015	
5	Post-editing machine translation: Efficiency, strategies, and revision processes in professional translation settings	Carl	Psycholinguistic and Cognitive Inquiries into Translation and Interpreting	115	BTL		2015	
6	Interactive Translation Prediction vs. Conventional Post-editing in Practice: A Study with the CasMaCat Workbench	Sanchis-Trilles	Machine Translation	in press	Springer		2014	
7	On-line and Active Learning for Machine Translation and Computer-Assisted Translation	González-Rubio	Proceedings of the Workshop on Interactive and Adaptive Machine Translation		AMTA		2014	
8	Integrating Online and Active Learning in a Computer-Assisted Translation Workbench	Alabau	Proceedings of the Workshop on Interactive and Adaptive Machine Translation		AMTA		2014	
9	On String Prioritization in Web-based User Interface Localization	Leiva	Proceedings of the 15th International Conference on Web Information System Engineering (WISE)				2014	460-473
10	Training of On-line Handwriting Text Recognizers with Synthetic Text Generated Using the Kinematic Theory of Rapid Human Movements	Martn-Albo	Proceedings of the 14th International Conference on Frontiers in Handwriting Recognition				2014	543-548
11	A Systematic Comparison of 3 Phrase Sampling Methods for Text Entry Experiments in 10 Languages	Sanchis-Trilles	Proceedings of the 16th international conference on Human-computer interaction with mobile devices and services (MobileHCI)				2014	537-542
12	CFT13: A Resource for Research into the Post-editing Process	Carl	Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC)		ELRA	Paris	2014	1757-1764
13	SEECAT: Speech & Eye-tracking Enabled Computer Assisted Translation	Martnez	Proceedings of the 17th Annual Conference of the European Association for Machine Translation		EAMT	Basel	2014	81-88

No.	Title	Main author	Title of the periodical or the series	Number, date or frequency	Publisher	Place of publication	Year	Pages
14	Speech-enabled Computer-aided Translation: A Satisfaction Survey with Post-editor Trainees	Mesa-Lao	Proceedings of the Workshop on Humans and Computer-assisted Translation (HaCaT)		ACL		2014	99-103
15	Gaze Behaviour on Source Texts: An Exploratory Study Comparing Translation and Post-editing	Mesa-Lao	Post-editing of Machine Translation: Processes and Applications		Cambridge Scholars Publishing	Newcastle upon Tyne	2014	219-245
16	Investigating User Behaviour in Post-editing and Translation Using the CASMACAT Workbench	Elming	Post-editing of Machine Translation: Processes and Applications		Cambridge Scholars Publishing	Newcastle upon Tyne	2014	147-170
17	Collaborative Web UI Localization, or How to Build Feature-rich Multilingual Datasets	Alabau	17th Annual Conference of the European Association for Machine Translation		EAMT		2014	151-154
18	Efficient Wordgraph Pruning for Interactive Translation Prediction	Sanchis-Trilles	17th Annual Conference of the European Association for Machine Translation		EAMT		2014	27-34
19	CASMACAT: Cognitive Analysis and Statistical Methods for Advanced Computer Aided Translation	Koehn	17th Annual Conference of the European Association for Machine Translation		EAMT		2014	57
20	FBK-UPV-UEdin participation in the WMT14 Quality Estimation shared-task	Souza	9th Workshop on Statistical Machine Translation		ACL		2014	322-328
21	Online Optimisation of Log-linear Weights in Interactive Machine Translation	Chinea-Rios	9th Language Resources and Evaluation Conference		ELRA		2014	3556-3559
22	Evaluating the Effects of Interactivity in a Post-Editing Workbench	Underwood	9th Language Resources and Evaluation Conference		ELRA		2014	553-559
23	CASMACAT: A Computer-assisted Translation Workbench	Alabau	14th Conference of the European Chapter of the Association for Computational Linguistics		ACL		2014	25-28
24	Proofreading Human Translations with an E-pen	Alabau	Workshop on Humans and Computer-assisted Translation		ACL		2014	10-15
25	Representatively Memorable: Sampling the Right Phrase Set to Get the Text Entry Experiment Right	Leiva	SIGCHI conference on Human Factors in Computing Systems		SIGCHI		2014	1709-1712
26	The Impact of Visual Contextualization on UI Localization	Leiva	SIGCHI conference on Human Factors in Computing Systems		SIGCHI		2014	3739-3742
27	The New Thot Toolkit for Fully-Automatic and Interactive Statistical Machine Translation	Ortiz-Martínez	14th Conference of the European Chapter of the Association for Computational Linguistics		ACL		2014	45-48
28	Inference of Phrase-Based Translation Models via Minimum Description Length	González-Rubio	14th Conference of the European Chapter of the Association for Computational Linguistics		ACL		2014	90-94

No.	Title	Main author	Title of the periodical or the series	Number, date or frequency	Publisher	Place of publication	Year	Pages
29	Improving on-line handwritten recognition in interactive machine translation	Alabau	Pattern Recognition	Vol. 47 (3)			2014	1217-1228
30	Cost-Sensitive Active Learning for Computer-Assisted Translation	González-Rubio	Pattern Recognition Letters	Vol. 37			2014	124-134
31	The Impact of Machine Translation Quality on Human Post-Editing	Koehn	Workshop on Humans and Computer-assisted Translation		EACL		2014	
32	Refinements to Interactive Translation Prediction Based on Search Graphs	Koehn	Annual Meeting of the Association for Computational Linguistics		ACL		2014	
33	Dynamic Phrase Tables for Machine Translation in an Interactive Post-editing Scenario	Germann	Workshop on Interactive and Adaptive Machine Translation		AMTA	Vancouver	2014	20-31
34	Error detection in native and non-native speakers provides evidence for a Noisy Channel Model of sentence processing	Hill	Annual Conference on Human Sentence Processing		CUNY	Columbus	2014	
35	Human detection of translation errors in text: unwrapping the dynamic process through eye-tracking	Hill	Translation in Transition: between cognition, computing and technology			Copenhagen	2014	
36	CASMACAT: A live demonstration, Architectures and Mechanisms in Language Processing	Hill	AMLaP XX			Edinburgh	2014	
37	CASMACAT: A Computer-Assisted Translation Workbench	Hill	DEMOfest			Edinburgh	2014	
38	Interactive Off-line Handwritten Text Transcription Using On-line Handwritten Text as Feedback	Martn-Albo	12th International Conference on Document Analysis and Recognition				2013	1280-1284
39	Interactive-predictive machine translation: towards the next generation of workbenches for translators	Sanchis-Trilles	Workshop on Multimedia Information Processing		DGA	Paris	2013	
40	Empirical Study of a Two-Step Approach to Estimate Translation Quality	González-Rubio	International Workshop on Spoken Language Translation				2013	227-234
41	Improving the Minimum Bayes Risk Combination of Machine Translation Systems	González-Rubio	International Workshop on Spoken Language Translation				2013	219-226
42	Multimodal post-editing	Alabau	Speech & Eye-Tracking Enabled CAT			Copenhagen	2013	
43	Interactive machine translation	Alabau	Speech & Eye-Tracking Enabled CAT			Copenhagen	2013	
44	Web Technologies in CASMACAT	Alabau	Speech & Eye-Tracking Enabled CAT			Copenhagen	2013	
45	Error-proof, High-performance, and Context-aware Gestures for Interactive Text Edition	Leiva	.ACM Annual Conference Extended Abstracts on Human Factors in Computing Systems		ACM		2013	1227-1232
46	Warped K-Means: An algorithm to cluster sequentially-distributed data	Leiva	Information Sciences	Vol 237:10			2013	196-210

No.	Title	Main author	Title of the periodical or the series	Number, date or frequency	Publisher	Place of publication	Year	Pages
47	Dimensionality reduction methods for machine translation quality estimation	González-Rubio	Machine Translation	Vol. 27 (3-4)	Springer		2013	281-301
48	Partial Least Squares for Word Confidence Estimation in Machine Translation	González-Rubio	Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA)	LNCS 7887	Springer		2013	500-508
49	User Evaluation of Advanced Interaction Features for a Computer-Assisted Translation Workbench	Alabau	Machine Translation Summit				2013	361-368
50	Advanced Computer Aided Translation with a Web-Based Workbench	Alabau	Workshop on Post-Editing Technologies and Practice				2013	53-62
51	CASMACAT: An Open Source Workbench for Advanced Computer Aided Translation	Alabau	The Prague Bulletin of Mathematical Linguistics	Vol. 100			2013	101-112
52	CASMACAT: Cognitive Analysis and Statistical Methods for Advanced Computer Aided Translation	Koehn	Machine Translation Summit				2013	411
53	Eye-tracking Post-editing Behaviour in an Interactive Translation Prediction Environment	Mesa-Lao	Journal of Eye Movement Research,	No. 3, Vol. 6		Lund, Sweden	2013	541
54	Interactive Machine Translation using Hierarchical Translation Models	González-Rubio	Conference on Empirical Methods in Natural Language Processing (EMNLP)				2013	244-254
55	Web Browsing Behavior Analysis and Interactive Hypervideo	Leiva	ACM Transactions on the Web	Vol. 7 (4)			2013	1-28
56	Findings of the 2013 Workshop on Statistical Machine Translation	Bojar	Workshop on Statistical Machine Translation		ACL		2013	
57	Paraphrasing in Context of Computer Aided Translation	Rustamli	MSc thesis	U. of Edinburgh			2013	
58	Dynamic Programming for Re-Mapping Noisy Fixations in Translation Tasks	Carl.	Journal of Eye Movement Research	Vol. 6, No. 2			2013	1-11
59	Towards a Classification of Translation Styles based on Eye-tracking and Keylogging Data	Dragsted	Journal of the Writing Research	Vol. 5, No. 1			2013	133-158
60	Scaling up Translation Process Research: Analysis of the Translation Process Using the CRITT TPR Database	Balling	Translation Studies: Centers and Peripheries		Eu.Soc. for Transl. Stud.		2013	210-211
61	The Casmacat Post-editing Workbench Prototype-II: A Research Tool to Investigate Human Translation Processes for Advanced Computer Aided Translation	Carl	Translation Studies: Centers and Peripheries		Eu.Soc. for Transl. Stud.		2013	230-231
62	Production Time Across Languages and Tasks: A Large-scale Analysis using the CRITT Translation Process Database	Winther	The development of translation competence: Theories and methodologies from psycholinguistics and cognitive science				2014	

No.	Title	Main author	Title of the periodical or the series	Number, date or frequency	Publisher	Place of publication	Year	Pages
63	Shared Representations and the Translation Process	Schaeffer	Describing Cognitive Processes in Translation: Acts and events	TIS 8:2			2013	
64	Incremental Learning for Statistical Machine Translation	Ortiz-Martínez	Workshop on Future Directions in Translation Research			Nara (Japan)	2012	
65	User Evaluation of Interactive Machine Translation Systems	Alabau	Annual Conference of the European Association for Machine Translation				2012	20-23
66	Transcribing Handwritten Text Images with a Word Soup Game	Alabau	Annual Conference on Human Factors		ACM		2012	2273-2278
67	Study of Electronic Pen Commands for Interactive-Predictive Machine Translation	Alabau	International Workshop on Expertise in Translation and Post-editing Research and Application				2012	4-7
68	PRHLT Submission to the WMT12 Quality Estimation Task	González-Rubio	Workshop on Statistical Machine Translation		NAACL		2012	104-108
69	Online adaptation strategies for statistical machine translation in post-editing scenarios	Martínez-Gómez	Pattern Recognition	45:9			2012	3193-3203
70	The next generation translator's workbench: post-editing in CASMACAT v.1.0	Mesa-Lao	Translating and the Computer Conference		ASLIB		2012	
71	On the Optimal Decision Rule for Sequential Interactive Structured Prediction	Alabau	Pattern Recognition Letters	33:6			2012	2226-2231
72	Building task-oriented machine translation systems	Sanchis-Trilles	PhD Thesis		UPVLC		2012	
73	Active learning for interactive machine translation	González-Rubio	Conference of the European Chapter of the Association for Computational Linguistics				2012	245-254
74	A finite-state approach to phrase-based statistical machine translation	González	International Workshop on Finite State Methods				2012	95-103
75	The CASMACAT workbench: a tool for investigating the integration of technology in translation	Elming	International Workshop on Expertise in Translation and Post-editing - Research and Application			Copenhagen	2012	
76	Translating vs Post-editing: A pilot study on eye movement behaviour across source texts	Mesa-Lao	International Workshop on Expertise in Translation and Post-editing - Research and Application			Copenhagen	2012	

No.	Title	Main author	Title of the periodical or the series	Number, date or frequency	Publisher	Place of publication	Year	Pages
77	Post-editing integration in a translation agency workflow	Silva	International Workshop on Expertise in Translation and Post-editing - Research and Application			Copenhagen	2012	
78	Inside the Monitor Model: Processes of Default and Challenged Translation Production Translation	Carl	Computation, Corpora, Cognition	Vol 2, No 1			2012	
79	The CRITT TPR-DB 1.0: A Database For Empirical Human Translation Process Research	Carl	Workshop on Post-editing Technology and Practice		AMTA		2012	
80	Translog-II: a Program for Recording User Activity Data for Empirical Translation Process Research	Carl	International Conference on Language Resources and Evaluation			Istanbul	2012	
81	Integrating Treebank Annotation and User Activity in TranslationResearch	Carl	Linked Data in Linguistics: Representing and Connecting Language Data and Language Metadata,		Springer		2012	77-85
82	Black Box Features for the WMT 2012 Quality Estimation Shared Task	Buck	Workshop on Statistical Machine Translation		ACL	Montréal	2012	91-95
83	Sentence Completion Prediction Based on Hierarchical Models	Assavanives	Honours project report		U. of Edinburgh		2012	
84	The CASMACAT Project: The Next Generation Translator's Workbench	Ortiz-Martínez	Jornadas en Tecnologa del Habla and the 3rd Iberian SLTech Workshop (IberSPEECH)				2012	326-334
85	Online learning of log-linear weights in interactive machine translation	Lpez-Salcedo	Jornadas en Tecnologa del Habla and the 3rd Iberian SLTech Workshop (IberSPEECH)				2012	277-286
86	Building task-oriented machine translation systems	Sanchis-Trilles	Jornadas en Tecnologa del Habla and the 3rd Iberian SLTech Workshop (IberSPEECH)				2012	480-489
87	Selecting translations to be post-edited by Sentence-Level Automatic Quality Evaluation	Garca-Martnez	Master's thesis		UPVLC		2012	
88	Aprendizaje online de los pesos del modelo log-lineal en traduccin	Salcedo	Master's thesis		UPVLC		2012	

All publications are available open access on the CASMACAT web site.

6.2 Template A2: List of Dissemination Activities

No.	Type of activities	Main leader	Title	Date	Place	Type of audience
1	symposium	CBS	Symposium on Eye tracking in Translation and Post-editing Studies	2012	Tobii conference Eye Tracking Behaviour	researchers
2	workshop	CBS	Workshop Eye tracking and Natural Language Processing	2012	COLING	researchers
3	workshop	CBS	Workshop on Speech and Gaze in Translation	2013		researchers
4	symposium	CBS	Special symposium on Empirical approaches to gaze data analysis in reading, writing and translation	2013	European Conference on Eye Movements	researchers
5	tutorial	CBS	Tutorial on Empirical Translation Process Research	2013	NLPCS	researchers, students
6	workshop	CBS	Workshop on Multimedia Information Processing	2013	DGA	researchers
7	workshop	UEDIN	Workshop on Humans and Computer-assisted Translation	2014	EACL	researchers
8	workshop	CBS	Workshop on Translation, Bilingualism and Translation Technology	2014		researchers
9	workshop	UEDIN	Workshop on Interactive and Adaptive Machine Translation	2014	AMTA	researchers
10	workshop	CBS	1st Workshop on Future Directions in Translation Research	2012	Copenhagen	researchers
11	workshop	CBS	2nd Workshop on Future Directions in Translation Research	2014	Copenhagen	researchers
12	course	CBS	Third PhD Course in Translation Processes Research	2013	Copenhagen	researchers, students
13	course	CBS	Fourth PhD Course in Translation Processes Research	2014	Copenhagen	researchers, students
14	project	CBS	Summer project on Translation Data Analytics	2014	Copenhagen	researchers
15	workshop	UEDIN	7th Workshop on Statistical Machine Translation	2012		researchers
16	workshop	UEDIN	8th Workshop on Statistical Machine Translation	2013	Sofia	researchers
17	workshop	UEDIN	9th Workshop on Statistical Machine Translation	2014	Baltimore	researchers
18	event	UEDIN	7th Machine Translation Marathon	2012	Montréal	students
19	event	UEDIN	8th Machine Translation Marathon	2013	Prague	students
20	event	UEDIN	9th Machine Translation Marathon	2014	Trento	students

6.3 Section B1: Patents, Trademarks, Registered Designs

The project did not file for patents, trademarks, or registered designs. All significant project outcomes were released open source.

6.4 Section B2: Exploitable Foreground

Type of Exploitable Foreground	Description of exploitable foreground	Exploitable product(s) or measure(s)	Timetable, commercial or any other use	Owner & Other Beneficiary (s) involved
Resource	Translation Process Database	Database	2014	Translation agencies, Localization companies
Software	Web-based advanced CAT tool	Workbench	2014	Translation agencies, Localization companies
Software	Home Edition	Workbench	2014	Translation agencies, Localization companies
Commercial exploitation of R&D results	Automatic and manual correction of gaze data	Workbench	2014	Translation agencies, Localization companies
Commercial exploitation of R&D results	Interactive-predictive machine translation	Workbench	2016	Translation agencies, Localization companies
Commercial exploitation of R&D results	On-line learning for machine translation	Workbench	2016	Translation agencies, Localization companies
Commercial exploitation of R&D results	Quality estimation of machine translation	Workbench	2015	Translation agencies, Localization companies
Commercial exploitation of R&D results	Active interaction for computer-assisted translation	Workbench	2015	Translation agencies, Localization companies
Commercial exploitation of R&D results	Active learning for computer-assisted translation	Workbench	2016	Translation agencies, Localization companies
Commercial exploitation of R&D results	E-pen for computer-assisted translation	Workbench	2016	Translation agencies, Localization companies
General advancement of knowledge	Bayesian adaptation for computer-assisted translation	Workbench	2017	Translation agencies, Localization companies

None of the exploitable foreground is confidential, embargoed. For all exploitable foreground, the sector is ranslation and interpretation activities.

Brief description of the exploitable foregrounds

- Translation Process Database

  A unique collection of translation process data generated by the project was made available for research purposes, creating the basis for analysis on a much larger scale than previously possible. The establishment of data standards enables interoperability.

- Web-based advanced CAT tool

  Based on the core MATECAT CAT tool, an advanced workbench for professional translators with many unique and novel feature was developed and released open source. It is ready to use, free of charge, by commercial users and researchers to carry out experiments and to build on.

- Home Edition

  Not only the workbench, but a full machine translation training environment has been developed for use by individual translators on their home computers. This is the first time that non-technical translators are able to build their own machine translation system on their own machines without need to resort to cloud services.

- Automatic and manual correction of gaze data

  To assist the analysis of eye tracking data, both automatic and manual gaze-to-eye mapping methods have been developed and integrated into the workbench. This is an essential tool for fine-grained analysis of eye tracking data and creates the basis for in-depth studies of translator behaviour.

- Interactive-predictive machine translation (IPMT)

  This foreground aims to increase the overall (machine translation + human) productivity by incorporating human correction activities within the translation process itself. It can be used in the translation workflow of the companies. Current state-of-the-art of IPMT is left-to-right, a more natural extension is to fix all the correct parts of the translations. Another extension is the use of syntax-based models for IPMT. Both extensions requiere more research.

- On-line learning for machine translation

  The text obtained by an IPMT or a post-editing (PE) system together with the corresponding aligned source segments can generally be converted into new, fresh training data, useful for adapting the system to changing environments. The current state of this foreground can be used in any statistical IPMT or PE system.

- Quality estimation of machine translation

  Enriching the output of the translation systems by providing additional information from the model allows to estimate the unreliable parts of the automatic translation. This foreground is the basis of the active interaction and active learning. Further research is necessary to get real time state-of-the-art quality estimation techniques.

- Active interaction for computer-assisted translation

  The user effort can be reduced by only editing those parts which the confidence measure classifies as incorrect. This foreground can be part of any machine translation system.

- Active learning for computer-assisted translation

  The combination of quality estimation and online learning allows benefits of the active interaction plus an improvement of the quality of the models used for translation: the human expert is not required to exhaustively supervise all translations and the subset of reviewed/corrected sentences can be used to improve the quality of the translation model.

- E-pen for computer-assisted translation

  The e-pen modality is more natural than others as the mouse or the keyboard and it is suitable in those scenarios where only few translation errors should be corrected. More research is necesary for devices such as touch screens.

- Bayesian adaptation for computer-assisted translation

  This foreground allows to avoid over-training towards development set by not forgetting the generality provided by the training set. Furthermore, re-estimating the systems parameters from scratch may imply a computational overhead which may not be acceptable in certain environments, such as translation systems configured for post-editing or IPMT, in which the final human user is waiting for the translations to be produced.

Report on societal implications

A **General Information** *(completed automatically when **Grant Agreement number** is entered.*	
Grant Agreement Number:	ICT 287576
Title of Project:	CASMACAT
Name and Title of Coordinator:	Professor Philipp Koehn

B **Ethics**	
1. Did your project undergo an Ethics Review (and/or Screening)? • If Yes: have you described the progress of compliance with the relevant Ethics Review/Screening Requirements in the frame of the periodic/final project reports? Special Reminder: the progress of compliance with the Ethics Review/Screening Requirements should be described in the Period/Final Project Reports under the Section 3.2.2 *'Work Progress and Achievements'*	***No***
2. Please indicate whether your project involved any of the following issues (tick box) :	
RESEARCH ON HUMANS	
• Did the project involve children?	No
• Did the project involve patients?	No
• Did the project involve persons not able to give consent?	No
• Did the project involve adult healthy volunteers?	No
• Did the project involve Human genetic material?	No
• Did the project involve Human biological samples?	No
• Did the project involve Human data collection?	No
RESEARCH ON HUMAN EMBRYO/FOETUS	
• Did the project involve Human Embryos?	No
• Did the project involve Human Foetal Tissue / Cells?	No
• Did the project involve Human Embryonic Stem Cells (hESCs)?	No
• Did the project on human Embryonic Stem Cells involve cells in culture?	No
• Did the project on human Embryonic Stem Cells involve the derivation of cells from Embryos?	No
PRIVACY	
• Did the project involve processing of genetic information or personal data (eg. health, sexual lifestyle, ethnicity, political opinion, religious or philosophical conviction)?	No
• Did the project involve tracking the location or observation of people?	No
RESEARCH ON ANIMALS	
• Did the project involve research on animals?	No
• Were those animals transgenic small laboratory animals?	No
• Were those animals transgenic farm animals?	No
• Were those animals cloned farm animals?	No
• Were those animals non-human primates?	No
RESEARCH INVOLVING DEVELOPING COUNTRIES	
• Did the project involve the use of local resources (genetic, animal, plant etc)?	No
• Was the project of benefit to local community (capacity building, access to healthcare, education etc)?	No
DUAL USE	
• Research having direct military use	No

• Research having the potential for terrorist abuse	No

C Workforce Statistics

3. Workforce statistics for the project: Please indicate in the table below the number of people who worked on the project (on a headcount basis).

Type of Position	Number of Women	Number of Men
Scientific Coordinator	0	1
Work package leaders	1	3
Experienced researchers (i.e. PhD holders)	1	7
PhD Students		7
Other	14	8

4. How many additional researchers (in companies and universities) were recruited specifically for this project?	**12**
Of which, indicate the number of men:	10

D Gender Aspects

5. Did you carry out specific Gender Equality Actions under the project?	O X	Yes No

6. Which of the following actions did you carry out and how effective were they?

	Not at all effective				Very effective
❑ Design and implement an equal opportunity policy	O	O	O	O	O
❑ Set targets to achieve a gender balance in the workforce	O	O	O	O	O
❑ Organise conferences and workshops on gender	O	O	O	O	O
❑ Actions to improve work-life balance	O	O	O	O	O
O Other:					

7. Was there a gender dimension associated with the research content – i.e. wherever people were the focus of the research as, for example, consumers, users, patients or in trials, was the issue of gender considered and addressed?

O Yes- please specify

X No

E Synergies with Science Education

8. Did your project involve working with students and/or school pupils (e.g. open days, participation in science festivals and events, prizes/competitions or joint projects)?

O Yes- please specify

X No

9. Did the project generate any science education material (e.g. kits, websites, explanatory booklets, DVDs)?

X Yes – demonstration web site and integration into existing translation platforms

O No

F Interdisciplinarity

10. Which disciplines (see list below) are involved in your project?

X Main discipline[1]: computer science (1.1), cognitive science (5.1, 5.4), translation studies (6.2)

O Associated discipline[1]: | O Associated discipline[1]:

G Engaging with Civil society and policy makers

11a Did your project engage with societal actors beyond the research community? *(if 'No', go to Question 14)*	X O	Yes No

11b If yes, did you engage with citizens (citizens' panels / juries) or organised civil society (NGOs, patients' groups etc.)?

O No

O Yes- in determining what research should be performed

O Yes - in implementing the research

X Yes, in communicating /disseminating / using the results of the project

[1] Insert number from list below (Frascati Manual).

11c	**In doing so, did your project involve actors whose role is mainly to organise the dialogue with citizens and organised civil society (e.g. professional mediator; communication company, science museums)?**	○ X	Yes No

12. Did you engage with government / public bodies or policy makers (including international organisations)

- ○ No
- ○ Yes- in framing the research agenda
- ○ Yes - in implementing the research agenda
- X Yes, in communicating /disseminating / using the results of the project

13a Will the project generate outputs (expertise or scientific advice) which could be used by policy makers?

- X Yes – as a **primary** objective (please indicate areas below- multiple answers possible)
- ○ Yes – as a **secondary** objective (please indicate areas below - multiple answer possible)
- ○ No

13b If Yes, in which fields?

				Information Society Research and Innovation	

13c If Yes, at which level?

- ○ Local / regional levels
- ○ National level
- X European level
- ○ International level

H Use and dissemination

14. How many Articles were published/accepted for publication in peer-reviewed journals?	**17**
To how many of these is open access[2] provided?	**17**
How many of these are published in open access journals?	**5**
How many of these are published in open repositories?	**0**
To how many of these is open access not provided?	**0**
Please check all applicable reasons for not providing open access:	
❑ publisher's licensing agreement would not permit publishing in a repository ❑ no suitable repository available ❑ no suitable open access journal available ❑ no funds available to publish in an open access journal ❑ lack of time and resources ❑ lack of information on open access ❑ other[3]: ……………	

15. How many new patent applications ('priority filings') have been made? *("Technologically unique": multiple applications for the same invention in different jurisdictions should be counted as just one application of grant).*	**0**

16. Indicate how many of the following Intellectual Property Rights were applied for (give number in each box).	Trademark	**0**
	Registered design	**0**
	Other	**0**

17. How many spin-off companies were created / are planned as a direct result of the project?	**1**
Indicate the approximate number of additional jobs in these companies:	**5**

18. Please indicate whether your project has a potential impact on employment, in comparison with the situation before your project:

X	Increase in employment, or	X	In small & medium-sized enterprises
❑	Safeguard employment, or	X	In large companies
❑	Decrease in employment,	❑	None of the above / not relevant to the project
❑	Difficult to estimate / not possible to quantify		

19. For your project partnership please estimate the employment effect resulting directly from your participation in Full Time Equivalent (*FTE = one person working fulltime for a year*) jobs:	*Indicate figure:32.8*

[2] Open Access is defined as free of charge access for anyone via Internet.
[3] For instance: classification for security project.

Difficult to estimate / not possible to quantify	❑

I Media and Communication to the general public

20. As part of the project, were any of the beneficiaries professionals in communication or media relations?

○ Yes X No

21. As part of the project, have any beneficiaries received professional media / communication training / advice to improve communication with the general public?

○ Yes X No

22 Which of the following have been used to communicate information about your project to the general public, or have resulted from your project?

❑	Press Release	X	Coverage in specialist press
❑	Media briefing	X	Coverage in general (non-specialist) press
X	TV coverage / report	❑	Coverage in national press
❑	Radio coverage / report	❑	Coverage in international press
❑	Brochures /posters / flyers	X	Website for the general public / internet
❑	DVD /Film /Multimedia	❑	Event targeting general public (festival, conference, exhibition, science café)

23 In which languages are the information products for the general public produced?

❑	Language of the coordinator	X	English
❑	Other language(s)		