



**Machine Translation Enhanced
Computer Assisted Translation**

D4.3 – Third version of MateCat tool

Author(s): Alessandro Cattelan.

Dissemination Level: Public

Date: October 2nd, 2014

Third version of MateCat tool

Grant agreement no.	287688
Project acronym	MateCat
Project full title	Machine Translation Enhanced Computer Assisted Translation
Funding scheme	Collaborative project
Coordinator	Marcello Federico (FBK)
Start date, duration	November 1 st 2011, 36 months
Dissemination level	Public
Contractual date of delivery	September 30 th , 2014
Actual date of delivery	October 2 nd , 2014
Deliverable number	4.3
Deliverable title	Third version of MateCat tool
Type	Report
Status and version	Final
Number of pages	7
Contributing partners	Translated S.r.l.
WP leader	Translated S.r.l.
Task leader	Translated S.r.l.
Authors	Alessandro Cattelan
Reviewers	Marcello Federico
EC project officer	Alexandra Wesolowska
The partners in MateCat are:	Fondazione Bruno Kessler (FBK), Italy
	Université Le Mans (LE MANS), France
	The University of Edinburgh (UEDIN)
	Translated S.r.l. (TRANSLATED)

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

FBK

MateCat

Marcello Federico

Povo - Via Sommarive 18

I-38123 Trento, Italy

federico@fbk.eu

Phone: +39 0461 314 552

Fax: +39 0461 314 591

Copies of reports and other material can also be accessed via <http://www.matecat.com>

© 2014, Alessandro Cattelan

No part of this document may be reproduced or transmitted in any form, or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission from the copyright owner.

Executive Summary

In the third year of the project, the MateCat tool has reached a high level of maturity and solidity, which makes it suitable for professional use. Translated S.r.l., the commercial partner of the consortium, has switched its entire production to MateCat, translating over 35 million words with it. In the last few months, over 3,000 translators have started using the MateCat tool for professional work.

This document provides an overview of key features of the 0.4.1 release of MateCat published on September 23rd, 2014.

Table of Contents

Executive Summary	3
1. MateCat tool features	5
1.1 Workflow automation (API)	5
1.2 Project management and language resources	5
1.3 Improvement of translation experience	6
2. Documentation	7
3. MateCat tool URL and GIT repository	7

1. MateCat tool features

In the third year of the project, the development of the MateCat tool has focused on creating a more solid and advanced translation software for professional translators and language service providers. The main advances are in three main areas: workflow automation, project management and language resources, improvement of translation experience.

1.1 Workflow automation (API)

MateCat provides a set of Rest API to let users integrate the CAT tool in their translation management systems or in other applications. The API allows to create projects and check their status from outside the MateCat tool.

Specific documentation for the API can be found at the following URL:
<http://pro.matecat.com/api/docs>

Translated S.r.l. tested the API in its workflow by connecting the MateCat tool to its translation management system to automatically create projects without human intervention.

1.2 Project management and language resources

The management panel has been extended with the functionality to archive and delete projects. Users can decide to archive projects when they finish their translations and they may also choose to remove the project from the MateCat tool by deleting them. For increased security, we also implemented an auto-archiving feature whereby all projects that have not been edited for more than 30 days are archived automatically and are therefore only accessible via the management panel. Archived projects can still be de-archived from the management panel.

The latest versions of the MateCat tool let users create and use private translation memories along with the collaborative translation memory provided by default. Users can create private TMs by generating a private key or by uploading TMX files to their project both from the project creation page and from the translation editor. Users can also create project-specific glossaries from the *Glossary* tab in the translation editor. Allowed fields are source term, target term and comments.

Third version of MateCat tool

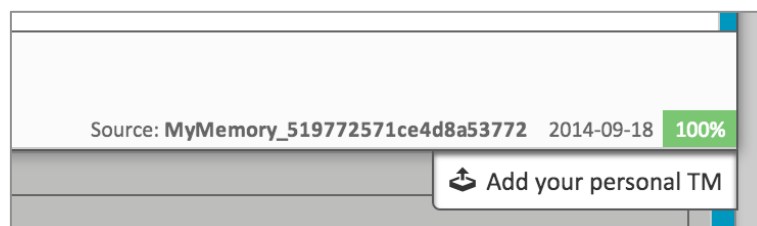


Figure 1. Adding personal translation memories.



Figure 2. Adding glossary terms.

1.3 Improvement of translation experience

The translation editor has been enhanced with advanced editing functionalities such as the capitalization of selected words by selecting them and switching the case with three dedicated buttons, and with the possibility to enter non-breaking spaces, tabulations and soft returns in the target segment.

A powerful search and replace functionality has been added to the MateCat tool. Clicking on the magnifying lens on the top right corner of the tool or using the shortcut CTRL+F (or CMD+F) brings up the search and replace panel. With this functionality, you can search for a specific pattern in the source or target segments and also for occurrences of pairs of words in the source and target simultaneously. From the same panel, you can also replace words in the target segment.

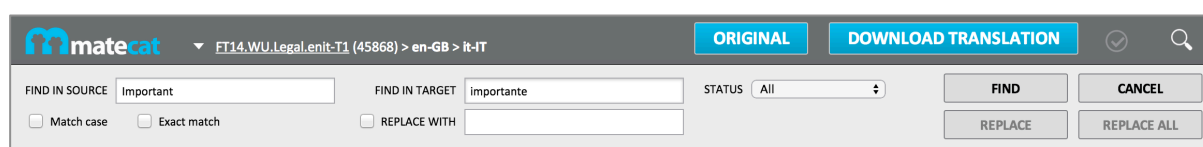


Figure 3. Search and replace panel.

We improved the management of tags by adding shortcuts to add tags in the target segment and by implementing automated controls for potential issues such as extra whitespaces, inconsistencies among tag IDs and missing tags.

Third version of MateCat tool

In terms of quality assurance, we developed a functionality which automatically propagate translation for a segment that is repeated throughout the document. If the user enters two or more different translations for identical source segments, the MateCat tool prompts a warning and displays the differences between the two versions, thus avoiding potential inconsistencies.

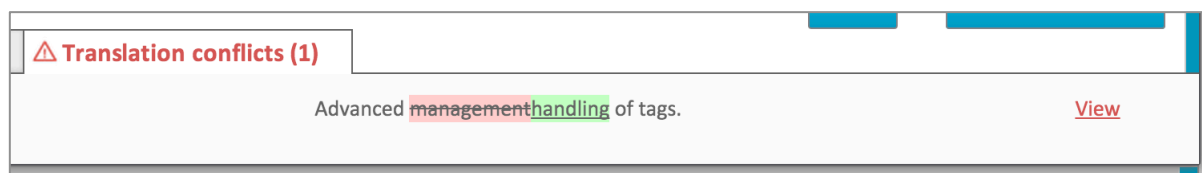


Figure 4. Inconsistencies in repeated segments.

Great effort has also gone into small improvements and minor bug fixes which resulted in over 300 small improvements as documented in our project management system (Asana). This resulted in a solid and easy to use translation software for professional use.

2. Documentation

The documentation for the MateCat tool is available online at the following URL:

- <http://docs.matecat.com>

Specific instructions on how to install the open source version are available at the following URL:

- <http://docs.matecat.com/installation-guide>

3. MateCat tool URL and GIT repository

The third release of the MateCat tool can be found at the following URL:

- <http://pro.matecat.com>

The source code is published regularly in the following GIT repository:

- https://www.assembla.com/code/matecat_source/git/nodes