



Proceedings of the
**21st Annual Conference of
the European Association
for Machine Translation**

28–30 May 2018
Universitat d'Alacant
Alacant, Spain

Edited by

Juan Antonio Pérez-Ortiz
Felipe Sánchez-Martínez
Miquel Esplà-Gomis
Maja Popović
Celia Rico
André Martins
Joachim Van den Bogaert
Mikel L. Forcada

Organised by



Universitat d'Alacant
Universidad de Alicante

transducens
research group



The papers published in this proceedings are —unless indicated otherwise— covered by the Creative Commons Attribution-NonCommercial-NoDerivatives 3.0 International (CC-BY-ND 3.0). You may copy, distribute, and transmit the work, provided that you attribute it (authorship, proceedings, publisher) in the manner specified by the author(s) or licensor(s), and that you do not use it for commercial purposes. The full text of the licence may be found at <https://creativecommons.org/licenses/by-nc-nd/3.0/deed.en>.

© 2018 The authors

ISBN: 978-84-09-01901-4

Terminology validation for MT output

Giorgio Bernardinello

STAR Group

Wiesholz 35, 8262 Ramsen

Switzerland

giorgio.berardinello@star-group.net

Abstract

WebTerm Connector is a plugin for STAR MT Translate which combines machine translation with validated terminology information. The aim is to provide “understandable” information in the target language using corporate language and terminology.

1 WebTerm

WebTerm is STAR’s web-based terminology management system that can be seamlessly integrated into STAR MT Translate in order to search for terms and display terminology information.

2 STAR MT Translate

STAR MT Translate is STAR’s web-based application for MT systems that can also be integrated into Microsoft Office products. STAR MT engines are based on customer-specific translations and terminology. As additional feature, the web-based application can access the company’s Transit translation memory.

3 WebTerm Connector

The Web Term Connector plugin allows users to search for translations of terms in the company’s dictionaries within the STAR MT Translate environment. Furthermore, it is possible to display additional terminology information for the source and the target language terms, if this is available in the dictionaries.

3.1 Searching for terms

On the assumption that the company dictionaries contain validated terminology, the translation is searched for in the dictionaries first. If no translation is found, especially for longer strings and sentences, the TM is then checked for perfect matches. If none are found, the MT engine will provide the translation.

3.2 Information about terms

If more information on parts of the source or the target sentence is needed, the user can simply highlight the text to search for in the company-specific dictionaries. If the word is found, its translation and the available terminology information are displayed at the bottom of the same web page. The customer can define what information from the dictionaries should be shown.

The major advantage of this solution is that users can quickly access company terminology without having to switch between browser tabs.

3.3 Translation of terminology information

Sometimes the dictionaries contain definitions or context descriptions in just in one language, e.g. the target language. But the user might need them in the source language in order to understand them, which is why MT technology is also integrated into the terminology area. This makes it possible to obtain an automatic translation of text parts of the dictionary simply by clicking the “Translate” button there.

3.4 Improved term searching functionality

All strings that were highlighted for terminology searching but for which no match was found in the dictionaries are listed in a log file. The customer can then use this information to improve not only their WebTerm dictionaries but also their MT engines.

© 2018 The authors. This article is licensed under a Creative Commons 3.0 licence, no derivative works, attribution, CC-BY-ND.