

ADVANCED FUZZY MATCHING IN THE TRANSLATION OF EU TEXTS

Margita Šoštarić

Abstract

In the translation industry today, CAT tool environments are an indispensable part of the translator's workflow. Translation memory systems constitute one of their most important features, and the question arises how to best use them to make the translation process faster and more efficient. This research examines the methods used to retrieve translation suggestions, especially focusing on whether using more sophisticated algorithms and features leads to significant improvement in the quality of retrieved matches. To that end, a number of matching configurations and their combinations are applied to the syntactic parse trees produced from the DGT-TM. The matches are evaluated by means of automatic evaluation, based on correlations and mean scores, and human evaluation, based on correlations of the derived ranks and scores. The goal is to determine whether the implementation of the tested fuzzy matching metrics should be considered in the commercial CAT tools to improve the translation process.