

DATA DESCRIPTION

MT Quality Evaluation carried out by human evaluators based on the MQM framework

"GT" = Google Translate

"DL" = Deep

"SY" = Systran Translate

01-total errors, sd, av, weighted avg, anova includes:

- a pdf file with summary data about descriptive statistics and ANOVA analysis
- 3 csv files with results. for each MT engine (GT, DL, SY) for each text, of all the error types and severity attributed by the evaluators.

02-error severity, categories, types includes, post-editing time, includes:

- a pdf file with summary data about errors from the perspective of severity.
- a pdf file with summary data about the 4 error categories analysed by the manual evaluation, namely Accuracy, Fluency, Terminology and Other of the MQM framework. Some error categories group together more than one error type. This is the case for Accuracy and Fluency.
Accuracy is composed by Mistranslation, Addition/Omission, Untranslated/DNT.
Fluency is composed by Grammar/Syntax, Register, Inconsistency, Spelling.
- a pdf file with summary data about all the error types
- a pdf file with summary data about the average time required for post-editing the translations performed by the three MT engines

03-statistical analysis on text length and readability, includes:

- a pdf file with the results of the statistical analysis aimed to investigate possible correlations between MT errors with text length and readability, and differences among the MT engines.
- csv file with the text length, readability, MQM errors and automated Bert Score error.