

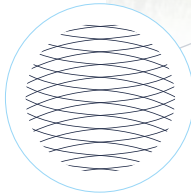
# Data-Driven Machine Translation Evaluation for Multilingual Excellence

## Case Study

## Objective

This client is engaged in the development of machine translation models across multiple language combinations, utilizing the 23 official languages of the European Union. Their objective is to gain a clear understanding of the quality achieved through customization of these machine translation models. They aim to involve linguists in the evaluation process to obtain a more comprehensive overview jurisdictions.

Translation | Interpreting | Intellectual Property



UNITED  
BY WORDS

## Challenge

### Framework

- Machine translation quality is primarily intended for informational purposes (referred to as 'gisting').
- Models are developed for various language combinations among the 23 official EU languages.
- Informative content can encompass a wide range of materials, including general corporate communications, technical documentation, intellectual property (IP) decisions, or other IP-related documents, website content, and more, all tailored to the client's needs.

### Technical specifications

- Evaluation is to be performed at the segment level, using bilingual files delivered in .tmx or .xlx formats.
- Statistical information on results should be customized to meet the client's specific requirements.
- No modifications are required in the output translation.

### Evaluation specifications

- Machine translation outputs should be assessed on a segment-by-segment basis within the bilingual document, utilizing a rating scale from one (1) to four (4) as follows:



Ideal



Possibly acceptable



Acceptable



Unacceptable



## Our client

A European Union official Agency, a world-class organization with global reach.

### Linguists' profile requests

- Each document should be independently evaluated by two (2) to three (3) linguists who possess an impeccable command of both the source and target language pair of the given bilingual file (C2 is the minimum required level).

### Environment for evaluations

- The client stipulated the use of a dedicated and appropriate environment that provides linguists with the necessary tools and ensures real-time availability of results.

### Estimated volumes

- The average document size can vary from 1,100 to 11,000 words, with an average of approximately 5,000 words per document.
- Each document should be evaluated independently between two and three times., converting the estimated volume into 15,000 words per analysis.
- While exact figures vary, there is an average of 6 jobs per month making the scope of the project exceed 5.32 million words over a 4-year project.

