Automating translation of high volumes for discovery processes

By meticulously training and fine-tuning MT engines tailored to each domain and language pair, Alpha CRC has developed a method that not only ensures the production of high-quality, consistent, and efficient translation outputs but also significantly reduces the time and costs associated with traditional translation methods.

Metrics

8m words
in 4 days

Tools

MT, TM

Outcomes

- **Improved** translation quality that meets industry-specific standards
- **Increased** efficiency with the ability to translate large volumes of text in a shorter time frame
- **Cost savings** due to the reduced need for extensive post-editing by human translators
- **Enhanced** scalability, allowing the corporation to easily adjust to varying volumes of translation needs without compromising on quality.

Use Case description

Alpha CRC rises to meet the challenge of translating vast volumes of documents for a multinational corporation by training and fine-tuning Machine Translation (MT) engines specific to each domain and language pair.

We curate high-quality, domain-specific corpora for iterative training cycles, ensuring the MT engines produce high-quality, consistent, and efficient translation outputs.

The outcome is a robust, fully automated translation workflow that delivers translations with improved accuracy and contextual relevance, leading to significant time and cost savings for the corporation. This innovative approach enhances the corporation's scalability and competitive edge in global markets, while also contributing to the advancement of MT technology and its application in business.

Higher volumes in shorter times

Challenge

A multinational corporation required the translation of vast volumes of documents across multiple domains (legal, technical, etc.) and language pairs in tight turnaround times.

In this case, when content is not highly visible, traditional human translation may be time-consuming and cost-prohibitive. However, untrained generic Machine Translation (MT) engines often produce suboptimal results, particularly with domain-specific jargon and nuances. The challenge is to ensure high-quality, consistent, and efficient translation outputs that are domain-appropriate and culturally sensitive.

Solution

To address this challenge, Alpha CRC's MT specialists train and fine-tune the most suitable MT engines for each domain and language combination. The process involves:

1. Selecting a range of MT engines with potential for the required tasks.

2. Curating domain-specific corpora to train the MT engines, ensuring that the training data is of high quality and representative of the actual content to be translated.

3. Conducting iterative training cycles where the MT engines learn from the domain-specific data.

4. Comparing the performance of the trained MT engines using a set of predefined metrics such as BLEU scores, TER, and domain-specific accuracy assessments by Alpha CRC's linguistic experts.

5. Selecting the best-performing MT engine(s) for each domain and language pair based on the comparative analysis.

6. Integrating the trained MT systems into a fully automated translation workflow, as depicted in the provided infographic.