CASE STUDY

argos multilingual

Argos AI TM Cleanup Tool Improves Machine Translation Results for Ariel

The Challenge

As the largest manufacturer of gas compressors in the world, Ariel Corporation wanted to take their localization program to the next level by employing more machine translation (MT) to streamline their translation processes across multiple languages. Because Ariel operates globally, the company requires translations in several languages to effectively meet market demands and comply with local regulations.

Ariel's customary process — using MT for translation and then having human linguists post-edit (MTPE) to refine the translation quality — worked well. However, Ariel wanted to process some content exclusively with machine translation, but the quality of the MT output had to improve before they could shift to an "MT and publish" process. Additionally, the company also wanted to reduce post-editing costs and get to market faster.

Ariel understood the need for a long-term strategy to support the continuous improvement of the MT engine to keep delivering quality translations. This would require cleaning their translation memory (TM) to eliminate inconsistent translations and terminology, misalignments, and obsolete entries. Typically, this kind of TM clean-up is performed manually, making it both time- and cost-prohibitive.

To resolve this dilemma, Ariel contacted their long-term language services partner, Argos for help.

The Solution

Argos proposed a two-step strategy to Ariel: First, clean the TM to eliminate false positives, inconsistencies, and risky TM matches. Then, use the cleaned TM to train a customized MT engine, which would improve accuracy and output so that post-editing would no longer be needed to meet their MT quality goals.

Argos' TM clean-up service, powered by AI, makes the arduous and costly process of cleaning up TMs efficient and affordable. Argos uses multiple AI networks that work in harmony. Layer 1 addresses TM quality distribution by categorizing the TM elements. Layer 2 AI is trained to analyze the target and source to determine if a segment should be translated or not.

These steps are supported by several other checks:

- **Consistency:** Is something translated in multiple ways? If it is, is it intentional or context-dependent?
- **Numbers:** Are there numerical differences between source and text?
- **Glossary:** Does it adhere to terminology? Does it contain 100 terms or less?
- **Regular expressions:** Are there differences in units of measurement, spelled-out numbers, or symbols?

These checks are conducted over four phases:

1. **Phase 1: Scoping:**
Argos and Ariel worked together to define the scope of content for review and auxiliary services. The resulting QA report shows the overall health and weakness of the TMs.

2. **Phase 2: Quoting:**
Argos's AI-driven system processes the TMs and generates a TM distribution report. The quote is based on the total number of issues identified. This report was used to narrow down the scope and focus on areas that would bring the most benefit.

3. **Phase 3: Implementation:**
Argos linguists review the automated TM cleaning. The removal of duplicate TM segments follows consistency checks and a final spot-check.

4. **Phase 4: Finalization:**
Argos issues a final, easy-to-understand report comparing the TM before and after cleaning. The report details the number of errors fixed and TM distribution.

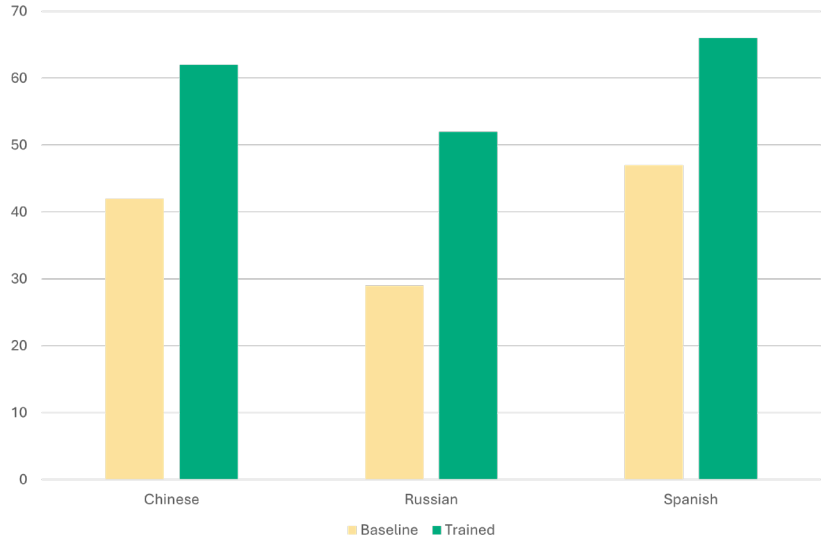
The next step was to use SYSTRAN Model Studio to build custom MT engines that were trained with the cleaned TMs. SYSTRAN Model Studio enabled a more precise customization of the MT engines, focusing specifically on the linguistic patterns and common errors identified in Ariel's previous translations. The cleaned TM allowed the machine translation to offer improved suggestions, supporting Ariel's quality goals.

The Result

Ariel was pleased with the results of the new approach, especially the high level of quality. After the TM cleanup, the translation error rate (TER) for 55% of the content was 0%, and another 24% of the content was between 1-5%.

The machine translation error rate also improved. For example, the percent of Spanish content with a 0% error rate increased from 39.8% to 60%. This demonstrates that using high-quality translation memory data to create a customized MT engine shows results in the translated content.

Although MT quality can vary by language, Ariel has achieved across-the-board improvement in Simplified Chinese, Russian, and Spanish from the AI TM cleanup and MT training strategy. The BLEU scores for Spanish and Simplified Chinese have gone from approximately 40 to over 60. Russian has likewise improved from a BLEU score of less than 30 to now over 50.



To ensure the engines operate at optimized levels, quality levels are closely monitored, and MT retraining is performed as needed.

Ariel has seen localization costs decrease as context matches and 100% matches no longer need to be manually reviewed, and translated content can now be published faster.

Although many organizations are unsure about the health of their TM, most understand that translation results are better if the TM is cleaned regularly. But because TM cleanup has historically been done manually, making it time-consuming and expensive, this important step has not been taken.

Today, there are more reasons than ever to perform a TM cleanup. Translation memory cleaning is now becoming the first step in a new and improved process, which involves using the TM data for machine learning and training. By incorporating AI, TM cleanup is now faster and more affordable, supporting ongoing improvements in translation accuracy and operational efficiency.

What They Say

"Working with Argos Multilingual has truly transformed our translation processes. We've been able to significantly reduce our post-editing cost and improve the accuracy of our translations. The two step strategy for cleaning the TM data with the Argos AI TM cleanup tool and building custom MT engines has been incredibly effective for our needs. We've learned that data quality is essential for getting the best results from AI and Argos has provided us the processes and expertise to efficiently and effectively clean our TM data and improve our MT performance. We're thrilled with the results and we look forward to our continuing partnership to improve our language services even further."
- Ron Egle, Ariel Corporation