Case Study NO. 027

How domain-specific Neural Machine Translation solutions support greater output accuracy and cost efficiencies.

ULG created and trained domain-specific Neural Machine Translation (NMT) engines and tested their effectiveness across multiple domains. The results? A methodology that can be applied to industries and clients for an improved quality output – even with non-traditional languages.

Background

ULG provides Neural Machine Translation (NMT) in over 130 translation directions for global clients across multiple markets and domains.

Each engine is designed to provide translations within the scope of a generic subject area. In order to integrate NMT in the production workflow, domain engines needed to be created, so ULG built domain-specific engines pertaining to medical devices, pharmaceutical, heavy machinery, IFU-DFU and Healthcare.

The Challenge

Developing domain-specific NMT engines is a unique undertaking that poses several distinct challenges.

For one, the NMT team requires a much larger pool of data and content to build out the terminology and segments. Additionally, to prepare the NMT engines for the basic production domains, ULG had to ensure consistency in the cleaned and aligned data, train the NMT engines to provide higher output quality and then test the integration quality in a standalone and production workflow.

The Solution

ULG created custom domain-specific NMT engines that increased the quality scores of the translation output.

Translation quality is scored using multiple evaluation methods, including the standard BLEU quality score which is a computer-generated metric that analyzes the results by comparing the NMT file to a reference file. ULG also uses TER (automated edit distance scoring) and independent Distance Scoring performed by a team of linguists to determine and validate quality scoring of NMT.

When creating the new engines, ULG's NMT team followed the standard engine customization process and obtained significant BLEU score increases compared to the score of the in-market engine and the scores of the new domain engines.

The Results

After creating domain-specific engines, the actual scores in live production projects showed that:

- In all cases the lower score of a Live project was higher than the score of the original engine.
- Average scores of Live projects were between 13 and 26 points higher than those of the original engine.

Translation direction	Domains	Original BLEU	Live production BLEU (Range)
English-Spanish	Generic	35,53	57-73
English-Spanish	Medical Devices	53,34	66-89
English-Spanish	Pharma	48,21	69-78
English-Spanish	Heavy Machinery	53,09	66-78
English-Spanish	IFU-DFU	55,29	59-88
English-Spanish	Healthcare	57,36	71

These results indicate ULG's domain-specific NMT will align with client content requirements to drive higher quality outputs.

The same methodology the ULG team employed to train the domain-specific engines is also what is used for client-specific engines, ensuring that terminology, corpora and overall content is better and more accurate to bring faster turnarounds, lower costs and increased quality.

- **130** NMT language directions
- **162** customized, domain-specific NMT engines
- **26** average increase in quality score using Live production NMT