

Pilot Project Proposal

FOR:

Wizards of the Coast LLC

P.O. Box 707

Renton, WA 98057



BY:

DnDorks (MIIS)

Dean, Christopher

Guedez, Marianna

Shin, Sook

Soza, Gary

Sublet, Eric

Pilot Project Objectives

Following the news that ***Wizards of the Coast LLC*** partnered with *Gale Force Nine* in localizing the ***Dungeons & Dragons*** fifth edition tabletop RPG content into multiple languages, the **DnDorks team** is proposing the following pilot project. This pilot project aims to determine the effort required in the training of a Neural Machine Translation (NMT) engine intended for use in the translating of content involved in the role-playing game (RPG), ***Dungeons & Dragons***, with a particular focus on both official and fan-base campaign materials (English to Spanish). This project's goal is for post-edited machine translations (PEMT) made with this engine to meet the following target criteria to achieve efficiency, cost, and quality.

- Efficiency: PEMT approximately **30%** faster than human translation
- Cost: PEMT approximately **40% savings** over human translation
- Quality: PEMT with an acceptable score based on the following model:

HUMAN EVALUATION SCORECARD
ADAPTED FROM DQF METRIC

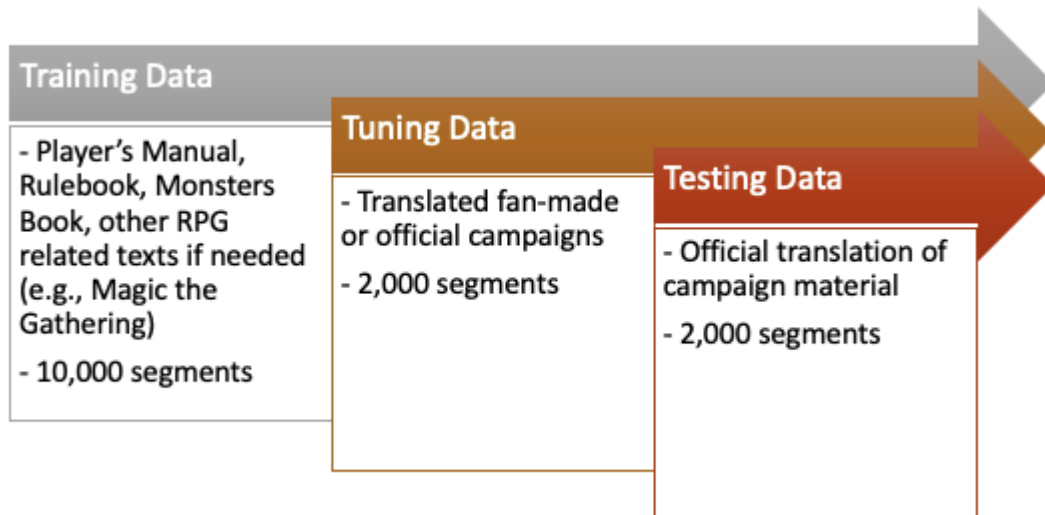
ACCURACY	ERROR TYPES (ACCURACY) <input type="checkbox"/> Mistranslation <input type="checkbox"/> Misspelling <input type="checkbox"/> Omission <input type="checkbox"/> Punctuation error <input type="checkbox"/> Grammatical error <input type="checkbox"/> Formatting (dates, etc)
FLUENCY	
VERITY	
HIT POINTS Threshold: 25 hit points /500 words	

An evaluation of the MT translation will be carried out by a human reviewer using a set of criteria (adapted from the DQF model) for error types and severity, performed towards the beginning and end of the engine’s training. (The format has been adapted to mimic scorecards typically used in Dungeons & Dragons to motivate and retain human evaluators, who likely have experience with the game.) The engine will lose *hit-points (HP)* for each error based on error type and severity. If the HP of the engine falls below 0, the round will be classified as **fail**.

Note: Verity error types will usually represent a major or critical error. While accuracy error types will represent minor errors unless the error is so big that changes the complete meaning of the segment.

To determine the work involved in bringing the MT to this benchmark, we will train a preliminary engine over the course of two to three weeks. We will estimate the requirements of a completed MT training project based on the costs and work involved in training the machine engine and by assessing the rate of improvement and the final quality.

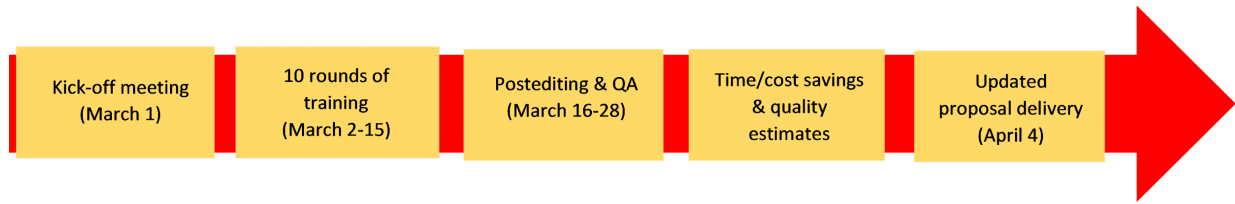
Data Sets



Process

- **File preparation:** Conversion of data sets to DOC/TEXT format
 - Relevant PDFs will be downloaded online and then converted into Docx using Adobe Acrobat Pro DC.
 - Files will be aligned into TMX using SDL Trados and then cleaned in Olifant to create acceptable segments for training.
 - Identify terms to replace with a unique token ID.
- **Training rounds:** MT training with Microsoft Translator Hub
 - DnDorks will evaluate the BLEU score after each round of training.
 - Different strategies will be implemented and adjusted based on each testing round's findings to improve the BLEU score (if there is a decrease in BLEU score, each subsequent round will be based on the most recent successful model). Including, but not limited to:
 - Addition of new training documents
 - Addition/modification of termbases
 - More detailed cleanup of training/tuning segments
- **Assessment:** Estimation of PEMT time, cost, and quality based on results of training rounds
 - Machine-translated documents will undergo at least two rounds of timed post-editing by human evaluators.
 - Post editors aim for grammatically, syntactically, and semantically correct translation. They won't be asked to make stylistic or preferential changes.
 - DnDorks will calculate the time required for post-editing time in the training process.
 - Cost savings will be approximated based on post-editing time, with a standard of \$30.00/hr rate.
 - DnDorks will use a customized form of the DQF quality metrics to determine if the translations are of acceptable quality.
- **Proposal update:** Based on the assessment of PEMT, we will reexamine the proposal estimates.

Timeline



Costs

The following table details the estimated time and costs for this pilot project.

Task	Estimated Hours	Rate	Subtotal
Document alignment	1	\$15.00	\$15.00
Document acquisition/cleanup	20	\$35.00	\$700.00
Glossary Creation	2	\$25.00	\$50.00
Post-editing (per 2,000 words)	3	\$30.00	\$90.00
QA	2	\$20.00	\$40.00
Microsoft Translator Cost			\$249.86
Post Mortem	4	15.00	\$60.00
PM	N/A	15% of the total	\$120.48
		Total	\$1,325.34

Deliverables

The following will be delivered to the client upon completion of the pilot project.

- Updated proposal for training an MT engine to translate Dungeons and Dragons campaigns
- Details of the data used for pilot project training (training, tuning, and testing data sets)
- Records of iterations made for training the engine and BLEU score for each round
- Estimates of the feasibility, time, and cost required to achieve the stated goals
- Suggestions for ways to continue training the MT engine

Terms & Conditions

A "Scope Change" shall mean any change to the work as outlined in a work order. Either party may propose a Scope Change. Upon being informed of any proposed Scope Change, **DnDorks** shall promptly go into a breakdown of the increased or decreased cost, if any, created by the change and any other impact such as the impact on **DnD Project Pilot's schedule** and impact on any agreed-to Milestone Dates. A Scope Change shall not be effective until a written change order, signed by both parties, has been executed.