

Pilot Project Proposal

Prepared for: Apple Inc.

Prepared by: Daniel Birriel, Justin Flaherty, Rebecca Guttentag, Juan Ortiz

February 20, 2019



PILOT PROJECT OBJECTIVES

Objective

For this project, our group aims to create an estimate of the time, money, and effort that will be needed in order to train a statistical machine translation (SMT) engine. The specific engine we are looking to create will automatically translate future updates to the Software License Agreements created for new Apple iOS devices and updates, specifically those in the iPad, iPhone, and iPod touch line of products.

Goals

In order to achieve the above objective, we aim to build our engine on the expectation that all post-edited machine translations (PEMT) will meet three major criterion:

- Automated Efficiency: the automated engine will allow for an expedited process when performing translation on product updates
- Cost Benefit: the engine will cut costs significantly compared to standard human translation
- Quality Assurance: the engine will perform in line with the QA standards dictated by the table below

Quality Assurance

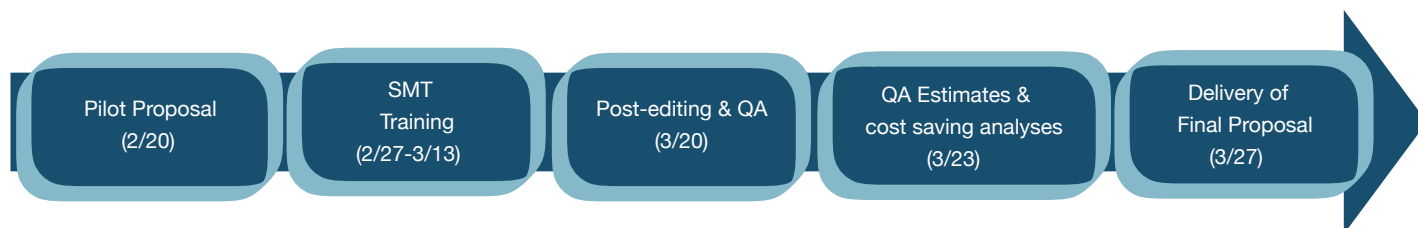
Description	Minor	Major	Critical
Mistranslation	3	7	10
Omission/Addition	2	4	7
Inconsistent Terminology	2	4	6
Grammar / Punctuation	2	3	5
Improper Link Cross-Reference	1	2	3
Awkward	1	2	3
Inconsistent Style	1	2	4

The threshold of quality that we will hold all machine translations to is a maximum score of 30 points. Translations that exceed this limit will be considered unacceptable and will be rerun after further improvement of the engine.

PROCESS

For the testing data, we have selected the most recent Terms and Conditions agreements found in the three most recent iPad, iPhone, and iPod touch iOS updates: iOS 12, iOS 11.2, and iOS 11. Future tuning will be done with older Terms and Conditions agreements found within the same line, while training will be done using the Terms and Conditions from other lines of Apple products, as well as other legal documentation that can be publicly accessed from Apple's website.

PROJECT TIMELINE



BUDGET

Estimated project costs include the following:

Task	Estimated Hours	Rate	Subtotal
MT Engine Setup	1	\$ 40	\$ 40
Alignment	50	\$ 40	\$ 2,000
MT Training Rounds	10	\$ 76	\$ 760
Post Editing	10	\$ 25	\$ 250
Quality Assurance (QA)	4	\$ 40	\$ 160
Total			\$ 3,210

DELIVERABLES

At the end of the project, the client will receive the following documents:

- The final, updated project proposal
 - Status reports and metrics on the data used, including training, tuning, and testing data sets
 - A compiled list of each round of the MT Training, along with the associated BLEU scores
 - Estimates of the time and costs needed to achieve the client's previously stated goals
 - An overall verdict on the project & any suggestions for adjustments or changes that will need to be made to increase feasibility
-