PROMT® Translation Solution for *PayPal*™ — Case Study

Contents

Summary

This case study describes the experience and benefits of using the PROMT machine translation (MT) engine in PayPal's localization process.

Background for Using MT at PayPal

PayPal maintains localized websites in 25 languages and it is critical that all these sites are synchronized. The company regularly translates large volumes consisting of various types of content: User Interface, Online help, FAQs, Error messages, Customer Support templates, User manuals, etc.

The PayPal localization workflow is based on SDL WorldServer. External language service providers (LSPs) translate and edit the content and PayPal's in-house linguists review the final output from the LSPs.

PayPal had two basic reasons to use MT in its localization process.

1. **Decreasing Time-to-Market.** PayPal's content size is about 1.5 million words for every language; around 15,000 words are translated on a bi-weekly basis. Time allocated to translation is a small portion of entire lifecycle.

2. **Cost Saving.** Properly customized MT engines produce translations that require minimal post-editing to achieve human translation quality. Typically, language service providers offer sizeable post-editing discounts compared to the cost of translating from scratch.

PayPal's Initial Requirements for MT vendor

Business Requirements

1. **Decrease Time-to-Market.** The translation cycle with MT+Post-Editing should take considerably less time than translating from scratch.

2. **Cost Saving => Translation Output Should Be Good Enough for Post-Editing.** The MT engine should produce raw translation of such quality that LSPs would be willing to offer post-editing discounts of up to 25%.

Technical Requirements

1. **Integration with SDL WorldServer.** The MT solution should be a component of the current localization process. WorldServer should invoke the MT engine without any additional engineering efforts.

2. **Stable and Scalable Server-Based Solution.** The Enterprise solution should deal with large translation volumes accessible twenty-four hours a day, seven days a week without any downtime.

3. **Retaining Specifics of PayPal's Content.** All the tags, variables, placeholders and other engineering metadata from the source should be preserved in the MT target.

4. **Tools for Further In-House Customization.** PayPal linguists should be able to customize the engine to decrease post-editing time.

PROMT Deployment

PROMT Translation Server

PROMT Translation Server (PTS) became the core of the PROMT Translation Solution for PayPal. PTS can successfully translate production volumes of up to 20 million words per day per translation server. At the same time, it has an API based on ASP.NET web-services, which allows integration of the MT engine into any workflow.

Within two months of the project kickoff in 2009, five engines had been trained:

- English > French
- English > German
- English > Italian
- English > Russian
- English > Spanish

How PROMT Addressed PayPal's Business Requirements

Training the MT engine is key to achieving the objectives of a shorter translation cycle and lower translation costs. Fast and flexible engine customization is what distinguishes PROMT engines, which are currently deployed in many Fortune 500 companies.

Of the utmost importance, all PayPal's existing linguistic assets were incorporated into the trained engine:

- PayPal Translation Memories (including 1.5 million words for each language) were attached to the MT engine.
- PayPal's terminology database was fully integrated into the PROMT user dictionaries (UDs).

- PayPal's syntactic and stylistic preferences were implemented into the PROMT Virtual Style Guide.

- Engineering specifics and all kinds of tagging were analyzed in PayPal's content and corresponding engineering rules were incorporated into the PROMT feature called XLIFF Rules.

How PROMT Addressed PayPal's Technical Requirements

1. **Integration with SDL WorldServer:** PROMT has developed an advanced MT adaptor for the WorldServer (XLIFF Connector), which sends to MT an XLIFF string with all formatting information:

```
This is a <ph id="1" x="&lt;b&gt;">{1}</ph>test<ph id="1" x="&lt;/b&gt;">{2}</ph>
```

… so the MT engine automatically puts placeholders into the correct position in the translation:

```
Das ist ein <ph id="1" x="&lt;b&gt;">{1}</ph>test<ph id="1" x="&lt;/b&gt;">{2}</ph>
```

2. **Robust and Scalable Server-Based Solution**. PROMT Translation Server is an advanced, scalable and robust solution, which is capable of handling huge production volumes. PTS powers PROMT free online websites with 10+ million words translated daily.

3. **Retaining Metadata of PayPal's Content (Tags, Variables, etc.)**: Due to the XLIFF Rules for tags and special processing for variables, PROMT can handle advanced engineering metadata. The example below shows how PROMT can manage PayPal-specific items:

Your {1} Bank Account Is Confirmed

In this instance, in PayPal's post-processing occurring after MT and post-editing, {1} would be filled in with the appropriate language-specific adjective such as "French", "German", "Spanish", etc. and the MT engine automatically places it in the correct order for the language. For instance, in the French translation, it would appear as follows (an adjective follows a noun in French):

Votre compte bancaire {1} est confirmé

4. **Tools for Further In-House Customization:** PROMT provided PayPal with a workbench for maintaining terminology and improving translation quality. PROMT also trained PayPal's MT specialists.

Enhancements and Upgrades

After the successful deployment of a baseline solution in 2010, PROMT introduced the following improvements in 2010 - 2012:

Normalization Engines

Since PayPal has different localizations for English-speaking countries (the US, the UK, and Australia), PROMT has developed custom normalization engines that produce the corresponding output from US source segments.

Integration with 3rd Party Engines

To accommodate PayPal's need for extra languages, PROMT built a platform that integrates multiple 3rd party MT engines. Such an integrated platform works from within the same PTS interface and allows PayPal to continue to utilize the same translation workflow.

After some internal evaluations, PayPal identified suitable machine translation engines for English-to-Chinese (Simplified and Traditional) and Nordics (Danish, Norwegian and Swedish). However, PayPal didn't want to change the existing MT workflow with PROMT. In addition, these 3rd party engines did not have a connector for WorldServer, a vital prerequisite for the PayPal localization process.

In response, PROMT created a wrapper of the Chinese and Nordic engines: PayPal's WorldServer sends all segments to the PROMT Translation Server, which works as a hub that redirects the translation requests for Chinese and Nordics to the respective 3rd party engines.

Creating Dictionaries for Reversed Pairs

PayPal has Business Units and Content Managers for French, Italian and Spanish (among other languages), so PayPal linguists are often requested to translate content stemming from an office in another country but that needs to be available in English as well.

To help PayPal with this task, PROMT has reversed existing English-to-other-languages dictionaries for translating from French, German, Italian, Russian and Spanish to English.

These reversed dictionaries were especially useful in urgent and high volume tasks.

PROMT DeepHybrid

In 2011, PayPal became one of the first PROMT clients in the US to take advantage of PROMT's new MT technology, DeepHybrid (DH). PROMT DeepHybrid combines the flexibility and predictability of rule-based MT with the fluency and lexical acumen of statistical MT.

After creating DeepHybrid profile on PayPal data, PROMT reported about the following results, which were received on 1% random segments from PayPal TM that had been considered as test part and had not been used in training process:

- **Scores:** BLEU increased about 15-20%, Editing Distance decreased on 10-15%.

- **Human evaluation:** in 80% cases, the DeepHybrid output is more fluent, more accurate and requires less post-editing.

PROMT DeepHybrid dramatically increased the translation quality for all of the PROMT language pairs at PayPal.

Conclusion

The deployment of PROMT's overall solution, including all of its various components, has helped PayPal translate faster and save on translation costs. Machine translation reduced mechanical work such as typing, terminology lookup and tag placement and allowed post-editors to focus more on style and overall fluency. Additionally, the in-house customization of the PROMT engine allows for the full utilization of PayPal's linguistic assets and further improves translation quality.

"Using PROMT technology in PayPal localization process can be considered as MT industry standard for software localization. PayPal maintains its website on more than 20 languages and all its releases should be translated very frequently. Trained PROMT engine translates the updated content well enough, so MT saves time and costs dramatically. This is a perfect background for using MT," says Julia Epiphantseva, Head of Business Development at PROMT.

Olga Loskutova, PayPal MT specialist, summarizes the PayPal experience with PROMT as quite successful:

"PayPal localization team has successfully integrated PROMT MT solution in the translation environment and has been using it in production for almost 2 years for a number of languages. With our tight schedule

and a growing need for quick turn-around time without compromising the quality of translation PROMT engines have proven to be dedicated and tireless robotic translators we cannot do without. Being able to customize the engines in-house also has become our greatest advantage as our trained linguists can prepare MT systems for human post-editing by coding new terminology into the MT user dictionaries per each release. The productivity growth varies across language pairs due to morphological and syntactic complexities of some languages. There is no doubt that with regular customization and upgrades benefits of using MT steadily become more and more noticeable. MT saves cost, time and helps improve the quality of translation."