Using PROMT for Localization at PTC, Inc. – Case Study

Contents

Synopsis

This case study highlights the benefits of using Machine Translation (MT) in the software localization process, based on the example of PROMT Translation Server implementation at PTC, Inc.

About PTC

PTC (Nasdaq: PTC) enables manufacturers to achieve sustained product and service advantage. PTC's technology solutions help customers transform the way they create, operate, and service products for a smart, connected world. Founded in 1985, PTC employs approximately 6,000 professionals serving more than 28,000 businesses in rapidly-evolving, globally-distributed manufacturing industries worldwide. Get more information at http://www.ptc.com.

Challenge

PTC localizes its core product lines from English into 10-14 languages. The localization process includes translating both the technical documentation and the user interface. This large volume of technical content needs to be translated accurately as well as quickly in order to meet project schedules and deadlines.

PTC's localization workflow for translating documentation is managed in SDL WorldServer. All translatable English content is extracted from the source files and segmented. Previously translated segments are leveraged against the Translation Memory (TM) database, while new segments are marked as such. The translation kit is then sent to Language Servers Providers (LSPs) for translation.

The primary objective of implementing Machine Translation in PTC's localization process was to reduce overall translation expense while keeping translation quality and localization schedules intact.

To achieve the goal the process needed to include pre-translation of new segments with a Machine Translation engine, so that a translator would post-edit machine-translated segments instead of translating these from scratch.

At the same time, the MT output needs to be of a high enough quality to enable translator's productivity to increase compared to the traditional translation metrics and to allow PTC to get significant post-editing discounts from LSPs.

Introducing PROMT into PTC's Localization Process

Pilot Project Overview

PROMT offered PTC a free MT pilot in Russian to demonstrate the advantages of Machine Translation. The English-to-Russian language pair was selected for the Machine Translation pilot for the following reasons:

- Because Russian is a highly-inflected language, the English-Russian language pair is challenging for MT, which makes it a good proof-of-concept selection.
- PTC had a reliable vendor for Russian, who was willing to participate in the pilot and provide feedback on the MT output quality.

PTC provided the following data for customization:

- PTC terminology database with ~12,000 English terms along with the Russian translations
- Translation Memory database with PTC Windchill segments that were translated in earlier localization projects (~70,000 segments or ~700,000 words)

After PROMT accomplished the initial customization and made the trained engine available to PTC, the PTC localization team translated a pilot project for PTC Windchill. After the Translation Memory was leveraged, new segments were pre-translated by the PROMT Translation Server.

The translation kit was delivered to PTC's vendor. The vendor performed post-editing tasks and provided valuable feedback on the MT quality, which they communicated directly to ProMT. After the completion of the pilot the vendor determined that post-editing took less time than translating from scratch and they offered PTC a post-editing discount for projects translated with MT.

Deployment into Production Workflow

After PROMT proved its efficiency during the pilot project, PTC purchased a perpetual license of PROMT Translation Server (for an in-house installation) together with full engine customization for all PTC products.

The full engine customization included the following:

- Creating a PTC dictionary from the PTC terminology database (~12,000 entries)
- Setting up and testing the PROMT Virtual Style Guide to correspond to PTC's Style Guide (~50 style guidelines were implemented)
- Building a Statistical Post-editing Model on PTC's Translation Memories for all products (~400,000 segments or ~4,000,000 words)
- Creating XLIFF rules for translating inline tags

In the scope of the initial deployment, PROMT successfully accomplished the following tasks:

- PROMT customized the PTC engine for all of PTC's products (all customization was added to the same engine, so all the PTC localization projects could be run with the same settings
- PROMT delivered an installation package of PROMT Translation Server with PROMT DeepHybrid Technology and an MT adaptor for WorldServer and helped PTC to install it on their server
- PROMT delivered 30 hours of on-site training to the PTC Localization team

Further Improvements

Automatic Action for UI elements

PTC technical documentation contains UI references that must appear in the documentation exactly the same way they appear in the running software. One possibility to ensure this consistency was for PTC to deliver software TMs to PROMT and for PROMT to create a dictionary from this database. However, this approach meant that such a dictionary would need updating on a regular basis. So PTC proposed to have the MT engine leave the UI elements in English and instead implemented an Automatic Action in WorldServer to insert translations for the UI Options. This Automatic Action captures the text of the English UI Option (since it is marked up with the 'uicontrol' tag) and retrieves the translation for this entry from the Software Terminology Database in WorldServer (which is regularly created from a SW TM and is therefore up-to-date).

For example, the following English segment…

```
When you move your cursor over text displayed in the <uicontrol>Multiple Rules</uicontrol> column, a tool-tip pops up, listing the applicable rules for the object in that row.
```

… is sent to PROMT, which translates the sentence as a whole, but intentionally leaves the content between the uicontrol tag intact, as it will be handled by the Automatic Action:

```
При перемещении курсора через текст, показанный в столбце <uicontrol>Multiple Rules</uicontrol> всплывающая подсказка раскрывается, внося в список применимые правила для объекта в той строке.
```

The machine-translation Automatic Action looks for the English string between the uicontrol tags in the Software TD in WorldServer and replaces it with the corresponding translation:

```
При перемещении курсора через текст, показанный в столбце <uicontrol>Различные правила (Multiple Rules)</uicontrol> всплывающая подсказка раскрывается, внося в список применимые правила для объекта в той строке.
```

In case the entry is not found in the SW TD, or multiple translations are found, WorldServer inserts the English string and outputs a comment to explain why the translation was not placed. The Automatic Action also inserts the English translation in parenthesis if the style guide for the product has such a requirement (`<uicontrol>Различные правила (Multiple Rules)</uicontrol>`).

The file with pre-translated segments is then sent to a vendor, who does not need to check for the correct translation for the UI element unless it has been left in English.

Adding a German Engine

After the deployment and positive results on the Russian engine, PTC evaluated an English-to-German PROMT customized engine and subsequently purchased it. Both the Russian and German engines have been successfully used in production since 2011/2012 respectively.

Upgrade to PROMT DeepHybrid v2.0

In late 2013, PROMT delivered a PROMT DeepHybrid 2.0 engine, which produced improved translation quality, and which was implemented on a more reliable architecture platform, including PTC-specific improvements (e.g., additional dictionaries and algorithms).

Conclusion

With over 23 years of progressive MT experience and technological advancement and through its comprehensive and detailed approach, PROMT has produced a solution for PTC which allows the company to cut translation costs while maintaining quality. The PROMT engine has assisted the PTC Localization Team in streamlining their production process by reducing tasks related to terminology research and tag placement on the MT content.

At PTC we knew we had to deploy machine translation technology in order to save on localization spend. But compromising on the quality is not an option for us: for the most part our documentation is highly technical and complex so usability is our prime concern. We require our vendors to deliver human-level translation quality on any projects, regardless of whether MT is used. Our vendors offer post-editing discounts based on the productivity assumptions, and therefore the quality of the raw MT output is extremely important to us.
I want to personally thank Alexey Borisov for providing a superb level of technical support, addressing any issues in a timely fashion, performing training sessions and advising on required customizations.
An integration of Machine Translation process into our WorldServer workflow allowed us to streamline the loc kit creation process and gain productivity at PTC.

– Natalia Levitina, Localization Director at PTC, Inc.

The efficiency of Machine Translation considerably increases if the system is set up for the type and subject-domain of the customer's documentation. The implementation of the PROMT solution in PTC fully confirms this.

– Julia Epiphantseva, Head of Business Development at PROMT