

The logo for PROMT, featuring the word "PROMT" in white, bold, uppercase letters on a blue rectangular background. A registered trademark symbol (®) is located at the top right of the word.

PROMT[®]

For Americas: 330 Townsend St., Suite 117, San Francisco, CA 94107 Tel: (415) 913-7586 Fax: (415) 913-7589 promptamericas@prompt.com

For other regions: 16A Dobrolubova av. ("Arena Hall" business center, building 2), 197198, St. Petersburg, Russia Tel. +7 (812) 611-0050
Alexey.lyapunov@prompt.ru

www.prompt.com

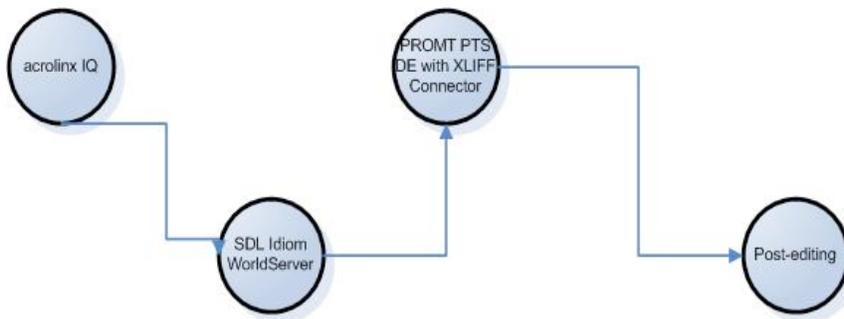
PROMT-Adobe Case Study:

Best Practices in Integrating Authoring, Globalization Management and Machine Translation

SUMMARY

Adobe is one of the acknowledged world leaders in developing cutting-edge software with more than half of the company's revenue coming from outside the United States. It is not surprising that Adobe uses the same stringent standards in building its localization process. Most machine translation use cases involve reduction of localization costs and time-to-market delta. The PROMT-Adobe case study is of particular interest since it also showcases a set of best practices for the entire content authoring and localization lifecycle:

- from acrolinx IQ for checking the authoring style to
- SDL Idiom WorldServer for globalization management to
- PROMT Translation Server (PTS) for machine translation to
- human post-editing of machine translation



PROMT has helped Adobe gain significant translation efficiencies, especially for structured content where translation volumes are high and content changes frequently. We're beginning to realize the promise of machine translation for the enterprise.

Christine Duran, Senior Translation Technology Manager at Adobe Systems

AT A GLANCE

Client Challenges

High quality machine translation solution to reduce localization costs and time-to-market delta

Interaction with SDL WorldServer

Metadata preservation requirement

Interaction with acrolinx IQ

Integration of best-of-class MT engines into primary MT engine

PROMT Products

PTS Developer Edition Server
Idiom WorldServer-PTS XLIFF Connector

XML Rules

Dictionary Editor

Term Manager

Content Maximization Suite

Virtual Style Guide

Results

Improved translation productivity by 20-30%

DEPLOYMENT OVERVIEW

Three language pairs and four locales in deployment:

English (US)

Portuguese (Brazil)

Russian

Simplified Chinese

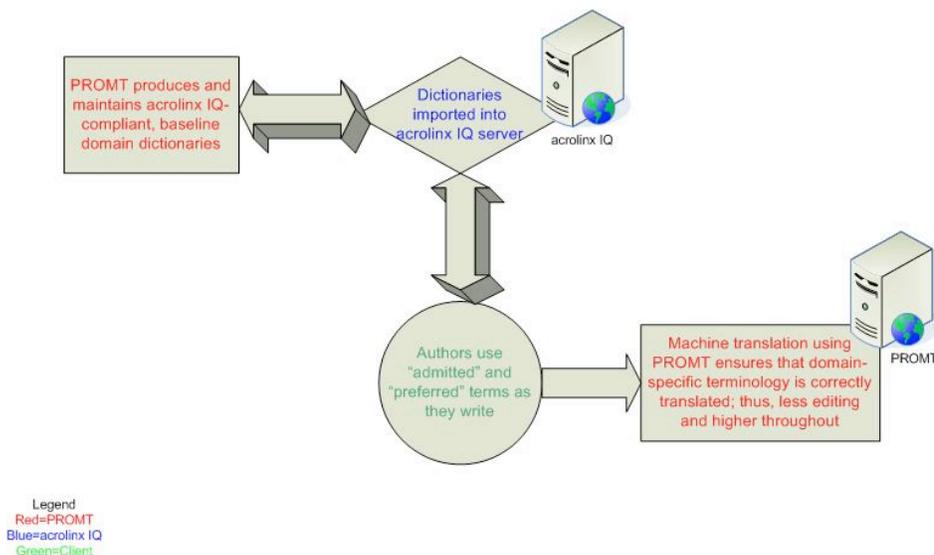
Content types: UI, Online help, FAQ, User Guide, Error Messages, Marketing, Field offices feedback, User feedback

Scope: over 1,000,000 words per language pair per product release

ADDRESSING CLIENT CHALLENGES

Integration with acrolinx IQ and the PROMT Virtual Style Guide

It is a well known and documented fact that good source leads to good translations and poor source leads to poor translations. The PROMT-acrolinx IQ integration goes one step beyond the source. PROMT produces and maintains acrolinx IQ-compliant dictionaries. In addition, some options in PROMT's *Virtual Style Guide*, a compendium of syntactic choices such as resolving part of speech homonymy, coincide with the acrolinx IQ set of rules. These two factors, coupled with verification of consistent use of preferred and admitted terms and adherence to company stylistic guidelines in English, make the English source highly amenable to getting a good machine translation.



Examples of PROMT/acrolinx IQ shared rules:

1. Use of the Imperative
2. Avoid Passive Voice
3. Use a Decimal Point

WIN: Adherence to company style and tone of voice. Excellent source leads to more elegant machine translation and faster, cheaper post-editing.

Metadata preservation through Idiom WorldServer integration

Adobe's localization process is based on WorldServer workflows. WorldServer sends a request to the machine translation engine beyond a certain fuzzy match threshold, usually 75%. However, unlike plain text connectors which are unable to handle placeholders and metadata information, the Idiom WorldServer-PROMT XLIFF Connector can process this metadata "noise" while retaining syntactic and semantic integrity. Consequently, instead of investigating where to place the tags, post-editors focus on linguistics. Lastly, the Idiom WorldServer-PROMT XLIFF Connector contains a self-correcting mechanism based on multiple pre- and post-processing algorithms and rules.

A small percentage of the time, the grammatical specificities of a language may create more tags than originally sent by WorldServer. If all these tags are sent to WorldServer, the segment will be rejected. The self-correcting mechanism strips out the extraneous tags and matches the number of tags originally sent. The intuitiveness of the connector guarantees that all valid XML/XLIFF strings are translated, an important measure if there is no post-editing for items such as knowledge bases.

Example of handling tags/metadata

Source: For more information, see the Kuler website at `<ph id="1" x="<xref href=""">{66}</ph>http://kuler.adobe.com/<ph id="2" x="" scope=""external" class=""- topic/xref ">">{67}</ph>http://kuler.adobe.com/<ph id="3" x="</xref>">{68}</ph>`.

Brazilian Portuguese Target: Para mais informações, ver o site web de Kuler `<ph id="1" x="<xref href=""">{66}</ph>em http://kuler.adobe.com/<ph id="2" x="" scope=""external" class=""- topic/xref ">">{67}</ph>http://kuler.adobe.com/<ph id="3" x="</xref>">{68}</ph>`.

Example of self-correcting mechanism

Source: The `{1}VideoPlayer{2}` and `{3}VideoElement{4}` classes can take a `StreamingVideoSource` instance as its `{5}source{6}` property.

Bad Russian target: `{1}Видеопроигрыватель{2}` и `{4}{3}` классы `VideoElement` могут взять образца `StreamingVideoSource` как свои `{6}{5}` свойства источника.

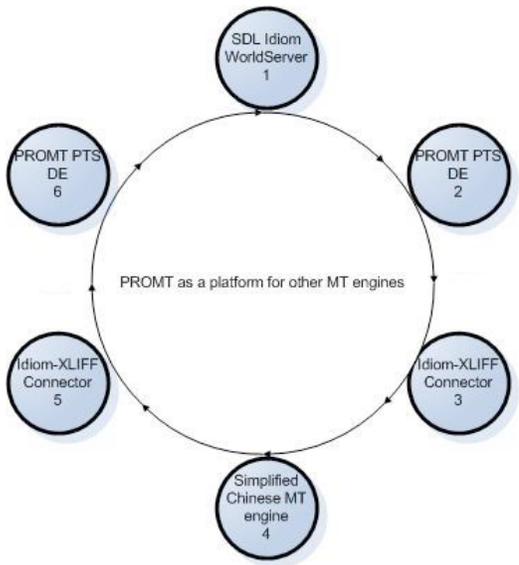
Self-corrected Russian target: `{1}Классы VideoPlayer и VideoElement могут взять случай StreamingVideoSource в качестве его исходной собственности{2}{3}{4}{5}{6}`.

WIN: MT post-editors can focus on their main task: post-editing the content, not rearranging the tags.

Integration of Simplified Chinese

Deployment of Simplified Chinese showcased PROMT' server's ability to serve as a platform, not just a stand-alone machine translation engine. After some internal evaluations, Adobe identified a suitable machine translation engine for Simplified Chinese. However, Adobe didn't want to maintain and train on an additional machine translation engine. In addition, the engine did not communicate with Idiom WorldServer, a vital prerequisite for Adobe's localization process.

Thanks to its modular architecture, PROMT created an integration of the Simplified Chinese engine; the engine was available from the regular PROMT web user interface and Idiom WorldServer was able to call the Chinese engine through the Idiom WorldServer-XLIFF Connector and PTS DE.



WIN: Since PROMT is used as a platform, the client is not required to buy, train and integrate multiple engines.

Terminology management and cleanup

Recognizing that brand erosion can have significant detrimental effects, Adobe vigilantly maintains an extensive multilingual product glossary. Porting Adobe's glossary into PROMT format addressed three issues simultaneously:

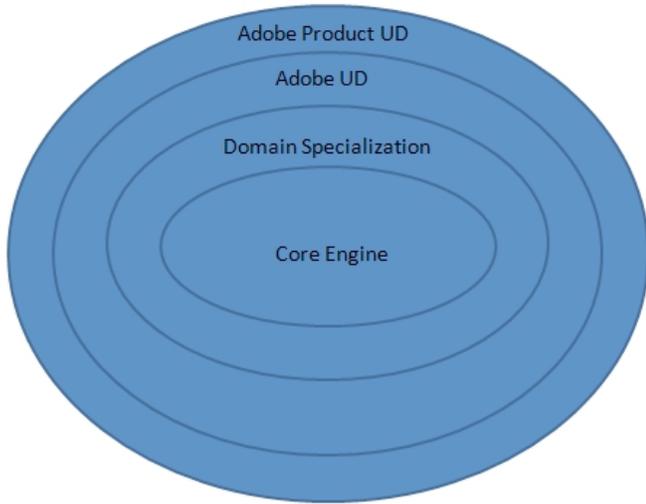
- identification and repair of inconsistencies between translation memory and glossary
- organization and taxonomization of glossary by product line
- automatic extraction of additional glossary candidates

PROMT used two best-of-breed terminology tools:

- the *Term Manager* module, which analyses an entire corpus, extracts a list of terminological candidates, and suggests a context-based translation
- *Content Maximization* suite, a toolkit containing a character counter for TMs and a statistical sub-sentential aligner with probability of lexical weight and translation for TMs

Both tools significantly augmented the number and quality of entries in the client glossary.

Ultimately, PROMT created a baseline user dictionary for Adobe terms, as well as individual product line user dictionaries. All dictionaries adhere to PROMT's "white-box" approach and are easily modified by client linguists. This multi-tiered approach (baseline plus specific product dictionaries) allowed Adobe to obtain relevant translations for the specific product line, especially in the cases of "collisions," a similar term, which has different translations depending on the product line.



Example of collision

Source1: title

Target1: титр

Source2: title

Target2: заголовок

WIN: relevant terminology and high terminology coverage by PROMT engine.