

Решение PROMT по переводу для компании PayPal

Содержание

РЕШЕНИЕ PROMT ПО ПЕРЕВОДУ ДЛЯ КОМПАНИИ PAYPAL	1
СОДЕРЖАНИЕ	1
КРАТКИЙ ОБЗОР	1
ПРЕДПОСЫЛКИ ДЛЯ ИСПОЛЬЗОВАНИЯ МАШИННОГО ПЕРЕВОДА ПРИ ЛОКАЛИЗАЦИИ ВЕБ-САЙТА PAYPAL	1
ИСХОДНЫЕ ТРЕБОВАНИЯ PAYPAL К ПОСТАВЩИКАМ МАШИННОГО ПЕРЕВОДА	2
Бизнес-требования	2
Технические требования	2
РАЗВЕРТЫВАНИЕ РЕШЕНИЯ PROMT	2
PROMT Translation Server	2
Соответствие решения PROMT бизнес-требованиям PayPal	3
Соответствие решения PROMT технических требованиям PayPal	3
УСОВЕРШЕНСТВОВАНИЯ И ОБНОВЛЕНИЯ	4
Система с функциями нормализации	4
Интеграция со сторонними решениями	4
Создание словарей для перевода на английский	4
PROMT DeepHybrid	5
ЗАКЛЮЧЕНИЕ	5

Краткий обзор

Данное тематическое исследование описывает опыт и преимущества использования системы машинного перевода (МП) PROMT в процессе локализации веб-сайта PayPal.

Предпосылки для использования машинного перевода при локализации веб-сайта PayPal

Компания PayPal поддерживает локализованные веб-сайты на 25 языках, и синхронизация этих веб-сайтов является критически важной задачей. Компания регулярно выполняет перевод больших объемов документов различного содержания, таких как пользовательский интерфейс, интерактивная справка, ответы на часто задаваемые вопросы, сообщения об ошибках, шаблоны службы поддержки пользователей, руководства пользователя и т. д. Технологический процесс локализации веб-сайта PayPal организован на основе SDL WorldServer. Внешние бюро переводов выполняют перевод и редактирование содержимого, а штатные лингвисты PayPal контролируют результаты работы контрагентов. Компания PayPal приняла решение об использовании машинного перевода в процессе локализации по двум основным причинам.

1. **Сокращение времени выхода на рынок.** Объем содержимого веб-сайтов PayPal составляет приблизительно 1,5 миллиона слов для каждого языка; каждые две недели переводится приблизительно 15 000 слов. Время, выделенное для перевода, составляет незначительную часть всего жизненного цикла.
2. **Экономическая выгода.** Надлежащим образом настроенные системы МП позволяют создавать переводы, требующие лишь минимального постредактирования для достижения качества перевода, выполняемого профессиональными переводчиками. Как правило, бюро переводов предлагают значительные скидки на постредактирование по сравнению со стоимостью перевода «с нуля».

Исходные требования PayPal к поставщикам машинного перевода

Бизнес-требования

1. **Сокращение времени выхода на рынок.** Цикл перевода на основе МП с учетом постредактирования должен занимать значительно меньше времени, чем перевод «с нуля».
2. **Экономическая выгода => Качество машинного перевода должно обеспечивать возможность постредактирования.** Система МП должна обеспечивать предварительный перевод такого качества, которое позволяло бы поставщикам лингвистических услуг предлагать скидки на постредактирование до 25 %.

Технические требования

1. **Интеграция с SDL WorldServer.** Решение для МП должно быть частью текущего процесса локализации. WorldServer должен использовать систему МП без каких-либо дополнительных трудозатрат.
2. **Устойчивое и масштабируемое решение на основе сервера.** Корпоративное решение должно обеспечивать возможность выполнения значительных объемов перевода в непрерывном круглосуточном режиме без каких-либо сбоев.
3. **Сохранение специфических особенностей содержимого веб-сайтов PayPal.** Все теги, переменные и другую техническую информацию из первоисточника должны быть сохранены в тексте машинного перевода в своем исходном виде.
4. **Инструменты для дальнейшей внутренней настройки.** Лингвисты PayPal должны иметь возможность дополнительной настройки системы МП в целях сокращения времени на постредактирование.

Развертывание решения PROMT

PROMT Translation Server

PROMT Translation Server (PTS) является ядром решения PROMT по переводу для компании PayPal. PTS позволяет переводить до 20 миллионов слов в день на каждом сервере перевода. Кроме того, PTS использует API, основанный на веб-сервисах ASP.NET, что позволяет интегрировать систему МП в любой технологический процесс.

В течение двух месяцев после начала проекта в 2009 г. решение для PayPal было настроено на перевод в пяти направлениях:

- английский язык > французский язык;
- английский язык > немецкий язык;
- английский язык > итальянский язык;
- английский язык > русский язык;
- английский язык > испанский язык.

Соответствие решения PROMT бизнес-требованиям PayPal

Настройка системы МП имеет ключевое значение для сокращения цикла перевода и снижения затрат на перевод. Возможность быстрой и гибкой настройки – отличительная особенность систем машинного перевода PROMT, которые в настоящее время развертываются во многих компаниях Fortune 500.

Очень важно, что в процесс настройки системы МП были включены все существующие лингвистические данные PayPal:

- Памяти переводов(Translation Memory) PayPal (включая 1,5 миллиона слов для каждого языка) были интегрированы с системой МП.
- Терминологическая база данных PayPal была полностью включена в пользовательские словари PROMT.
- Синтаксические и стилистические требования PayPal были точно отражены в МП путем установки соответствующих правил перевода, предусмотренных в системе PROMT.
- Были проанализированы все технические особенности и форматирование содержимого веб-сайтов PayPal, и соответствующие технические правила были реализованы в виде правил XLIFF.

Соответствие решения PROMT технических требованиям PayPal

1. **Интеграция с SDL WorldServer:** компанией PROMT разработан усовершенствованный адаптер МП для WorldServer (XLIFF Connector), который посылает в МП строку XLIFF, содержащую всю информацию о форматировании:

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

... таким образом, система МП автоматически размещает техническую информацию в переводе надлежащим образом:

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

2. **Надежное и масштабируемое серверное решение:** PROMT Translation Server является современным надежным масштабируемым решением, позволяющим обрабатывать большие объемы информации. PTS обеспечивает поддержку веб-сайтов PROMT для бесплатного онлайн-перевода в объеме свыше 10 миллионов слов ежедневно.
3. **Сохранение** технической информации **веб-сайтов PayPal (признаков, переменных и т.д.):** благодаря применению правил XLIFF для признаков и специальной обработки переменных, системы PROMT справляются с самой сложной технической информацией. Нижеприведенный пример иллюстрирует обработку специфичного для PayPal содержимого с помощью решения PROMT:

Your {1} Bank Account Is Confirmed

В этом случае при последующей обработке, выполняемой PayPal после МП и постредактирования, вместо {1} вставляется соответствующее прилагательное, определяющее язык, например «французский», «немецкий», «испанский» и т. д., и система МП автоматически размещает его в тексте

согласно порядку слов требуемого языка. Например, в переводе на французский язык (где прилагательное следует за существительным) фраза формируется следующим образом:

Votre compte bancaire {1} est confirmé

- 4. Инструменты для дальнейшей внутренней настройки:** PROMT предоставляет PayPal инструментальные средства для ведения терминологии и повышения качества перевода. Кроме того, компания PROMT также провела обучение специалистов PayPal по машинному переводу.

Усовершенствования и обновления

После успешного развертывания базового решения в 2010 г. компания PROMT осуществила следующие усовершенствования продукта в 2010-2012 гг.:

Система с функциями нормализации

Поскольку PayPal имеет различные локализации для англоязычных стран (США, Великобритании и Австралии), компания PROMT разработала механизмы лингвистической нормализации американского варианта английского языка для британского и австралийского веб-сайтов.

Интеграция со сторонними решениями

Для адаптации к потребностям PayPal в дополнительных языках компания PROMT создала платформу, объединяющую несколько систем МП сторонних поставщиков. Такая интегрированная платформа функционирует в рамках исходного интерфейса PTS и позволяет PayPal не менять технологический процесс перевода.

После проведения внутреннего анализа компания PayPal выбрала соответствующие системы машинного перевода с английского языка на китайский (упрощенный и традиционный), а также для скандинавских языков (датского, норвежского и шведского). Однако компании PayPal не хотелось изменять существующий технологический процесс МП в связи с использованием решения PROMT. Кроме того, эти сторонние системы не имели коннектора для WorldServer – критически важного элемента процесса локализации PayPal.

Для решения этой задачи в компании PROMT была создана оболочка для систем машинного перевода на китайский и скандинавские языки, функционирующая следующим образом: WorldServer, установленный в PayPal, посылает все сегменты в PROMT Translation Server, который является центральным узлом, переадресовывающим запросы перевода на китайский и скандинавские языки в соответствующие сторонние системы.

Создание словарей для перевода на английский

Компания PayPal имеет подразделения, использующие французский, итальянский, испанский и другие языки, а также контент-менеджеров, работающих с содержимым на этих языках. Таким образом, от лингвистов PayPal часто требуется перевод содержимого, поступающего из зарубежных офисов, которое, в то же время, должно быть доступно и на английском языке. Для решения этой задачи компания PROMT создала словари для перевода, позволяющие осуществлять перевод с французского, немецкого, итальянского, русского и испанского языка на английский язык. Основой новых словарей стали ранее разработанные словари для перевода с английского, и они оказались особенно востребованными при выполнении срочных и крупномасштабных задач.

PROMT DeepHybrid

В 2011 г. компания PayPal стала одним из первых клиентов PROMT в США, который воспользовался преимуществами новой гибридной технологии машинного перевода PROMT – DeepHybrid (DH). PROMT DeepHybrid сочетает в себе гибкость и предсказуемость машинного перевода на основе правил с гладкостью статистического МП.

После создания гибридного профиля для данных PayPal компания PROMT сообщила о следующих результатах, полученных на основе 1 % произвольно выбранных сегментов машинного перевода PayPal, которые рассматривались в качестве материала для тестирования и не использовались в процессе настройки:

- **Количественные показатели:** показатели оценки качества перевода с использованием методологии BLEU увеличились приблизительно на 15-20 %; редакционное расстояние (количество необходимых правок) уменьшилось приблизительно на 10-15 %.
- **Оценка с точки зрения профессионального переводчика:** в 80 % случаев DeepHybrid обеспечивает более гладкий, более точный МП, требующий меньшего времени на постредактирование.

Технология PROMT DeepHybrid позволила значительно повысить качество перевода для всех языковых пар PROMT в рамках проекта PayPal.

Заключение

Развертывание решения PROMT в полном объеме, включая все существующие компоненты, позволило PayPal увеличить скорость перевода и существенно снизить затраты на перевод. Машинный перевод уменьшил объем рутинных работ, таких как набор текста, поиск терминологии и размещение тегов, а также позволил специалистам по постредактированию уделять большее внимание стилю и общему качеству перевода. Кроме того, внутренняя настройка системы PROMT предполагает использование всех лингвистических данных PayPal и обеспечивает дополнительное повышение качества перевода.

«Группа локализации PayPal осуществила успешную интеграцию решения для МП PROMT в среду перевода и использует его в производственном процессе в течение почти двух лет для перевода на целый ряд языков. В условиях напряженного графика и все возрастающей потребности в сокращении сроков выполнения работ без ущерба для качества перевода система PROMT доказала свою незаменимость в качестве эффективного инструмента автоматизированного перевода. Кроме того, возможность внутренней настройки механизмов является наиболее значимым преимуществом, поскольку наши высококвалифицированные лингвисты могут подготовить систему МП к последующей обработке профессиональными переводчиками благодаря включению новой терминологии в пользовательские словари МП для каждого конкретного проекта. Рост производительности варьируется для различных языковых пар в силу морфологических и синтаксических особенностей некоторых языков. Нет сомнения, что благодаря возможности надлежащей настройки и внесения соответствующих обновлений преимущества от использования МП становятся все более очевидными. Машинный перевод позволяет увеличивать экономическую эффективность, снижать трудозатраты и повышать качество перевода».

Ольга Лоскутова, специалист по машинному переводу в PayPal

«Использование технологии PROMT в процессе локализации PayPal можно считать отраслевым стандартом использования МП при локализации программного обеспечения. Веб-сайт PayPal доступен более чем на 20 языках, и его обновления требуют постоянного перевода. Обучаемый механизм PROMT предоставляет достаточно качественный перевод обновляемого содержимого, обеспечивая таким образом существенную экономию времени и снижение затрат. Это весомая причина для использования машинного перевода».

Юлия Елифанцева, директор по развитию бизнеса компании
PROMT