

# PROMT

ПРОМТ

Внедрение решений  
по машинному переводу на примере  
Adobe и ФГУ ФИПС

**PROMT**<sup>®</sup>

# PROMT

- 20 лет успешной работы в области разработки, развития и внедрения решений для автоматического перевода текстов.
- 27 направлений перевода.
- Более 10 000 компаний-клиентов в сферах IT, промышленности, финансов, торговли, науки, образования, в переводческих агентствах и государственных структурах.
- Международная компания – более 100 сотрудников в России, Германии и США.
- Единственный Machine Translation провайдер в России. Успешно конкурирует с другими провайдерами на международных рынках.
- Интеграция в международное сообщество разработчиков систем машинного перевода: ведущая роль в международных отраслевых ассоциациях, участие в проектах для Евросоюза.
- Стратегическое партнерство с SDL, возможность интеграции программ-переводчиков PROMT с любыми системами Translation Memory (базы переводческой памяти).
- Есть свое бюро переводов, использующее технологии машинного перевода.

# Основные типы использования машинного перевода

- Получение максимально качественного машинного перевода для постредактирования человеком.
- Gisting, или перевод для понимания общего смысла текста.

# Сценарий внедрения

- Создание пилотного проекта.
- Развертывание готового решения на базе пилотного проекта.
- Обучение специалистов компании или аутсорсинг лингвистических услуг.
- Включение решений по машинному переводу в документооборот компании.
- Глубокая настройка системы с использованием всех лингвистических данных клиента (гlossариев, параллельных текстов) в процессе реальной эксплуатации решения.

# Обработка лингвистических данных заказчика и настройка системы

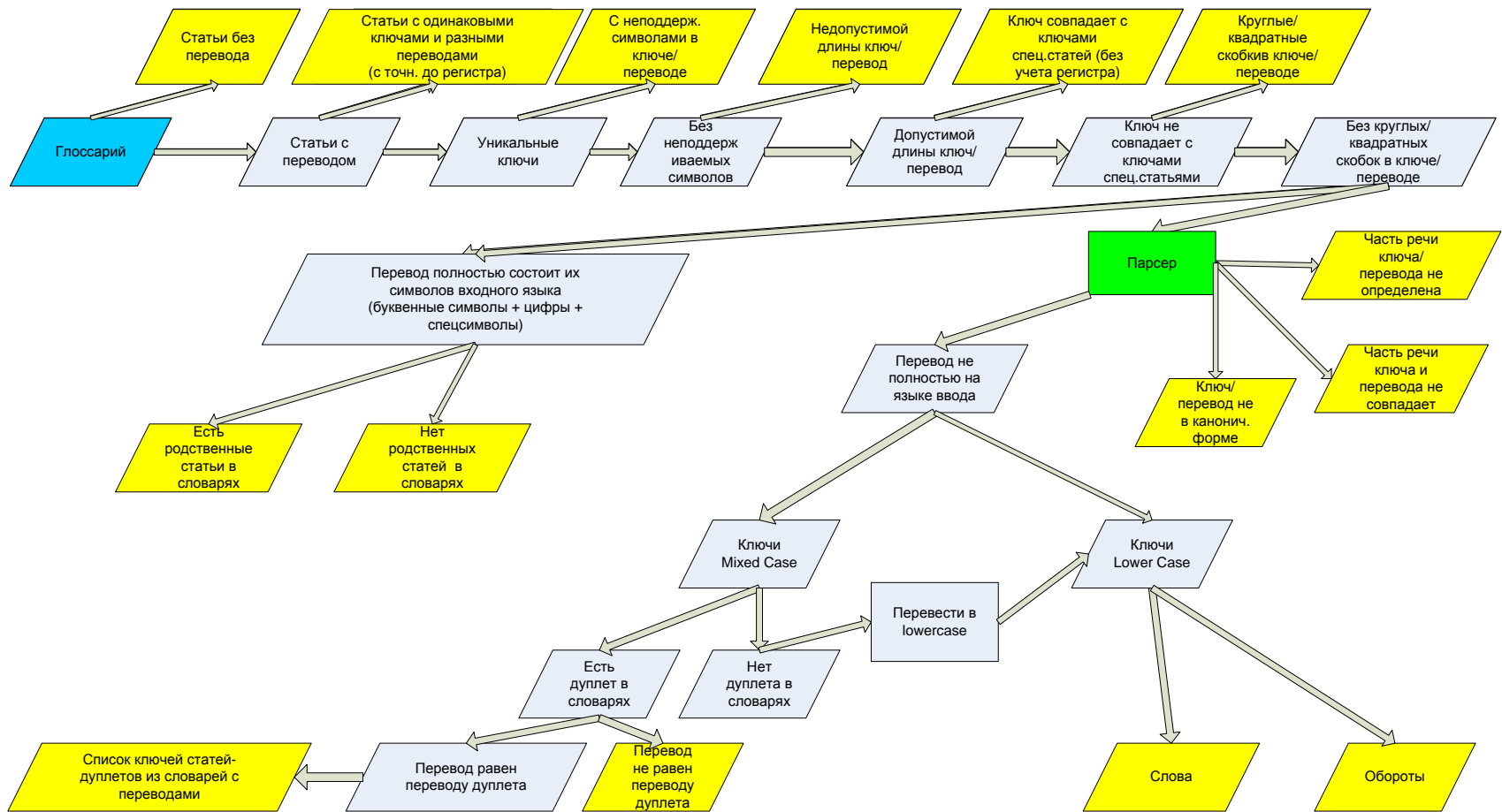
- Создание словаря машинного перевода на базе глоссариев.
- Подбор специализированных словарей PROMT.
- Выделение терминологии через PROMT Terminology Manager.
- Подбор правил перевода (т. е. способов синтеза тех или иных конструкций) с учетом типа текста.
- Тестирование результатов и коррекция настроек.

# Создание словаря из глоссариев

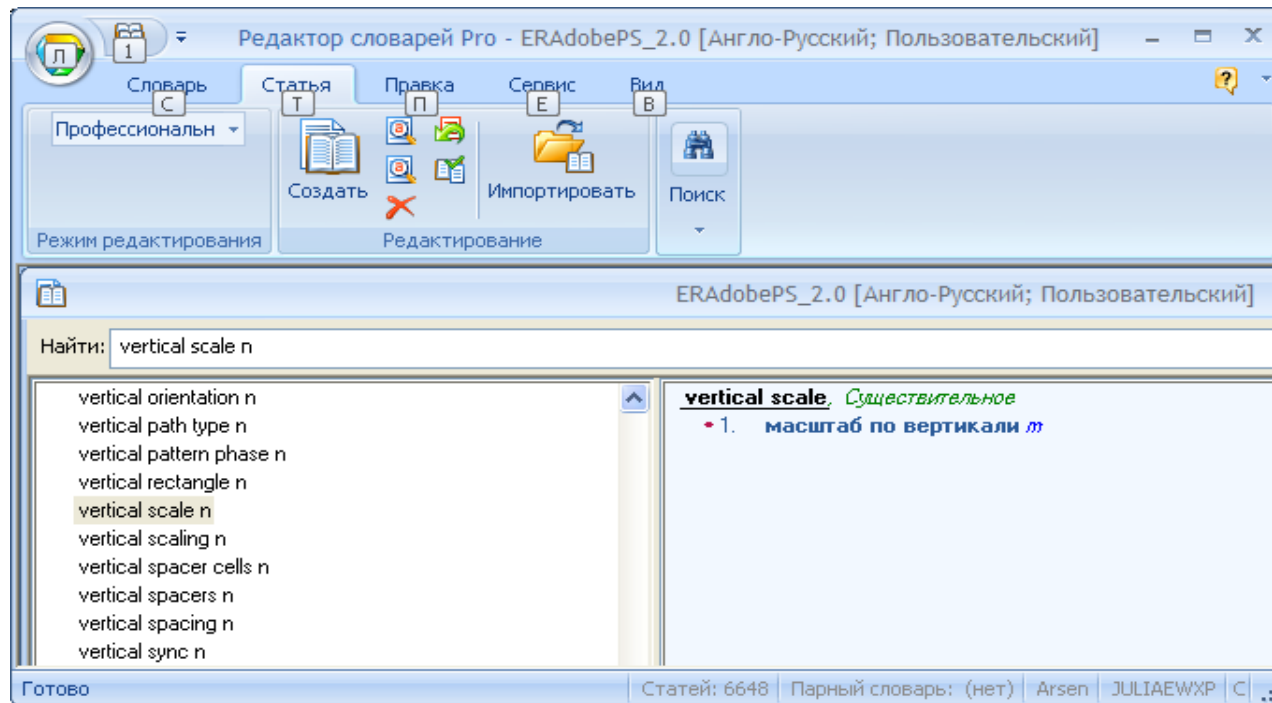
Глоссарии могут иметь некорректные с точки зрения технологии машинного перевода PROMT записи

- PROMT SplitGlossary Tool – выделение «чистой» части исходного глоссария и сортировка по группам.
- Автоматическое создание словаря в PROMT Dictionary Editor.
- Проверка словарных статей, имеющих маркировку unverified.

# Алгоритм обработки глоссария в PROMT SplitGlossary Tool



# Автоматическое создание словаря

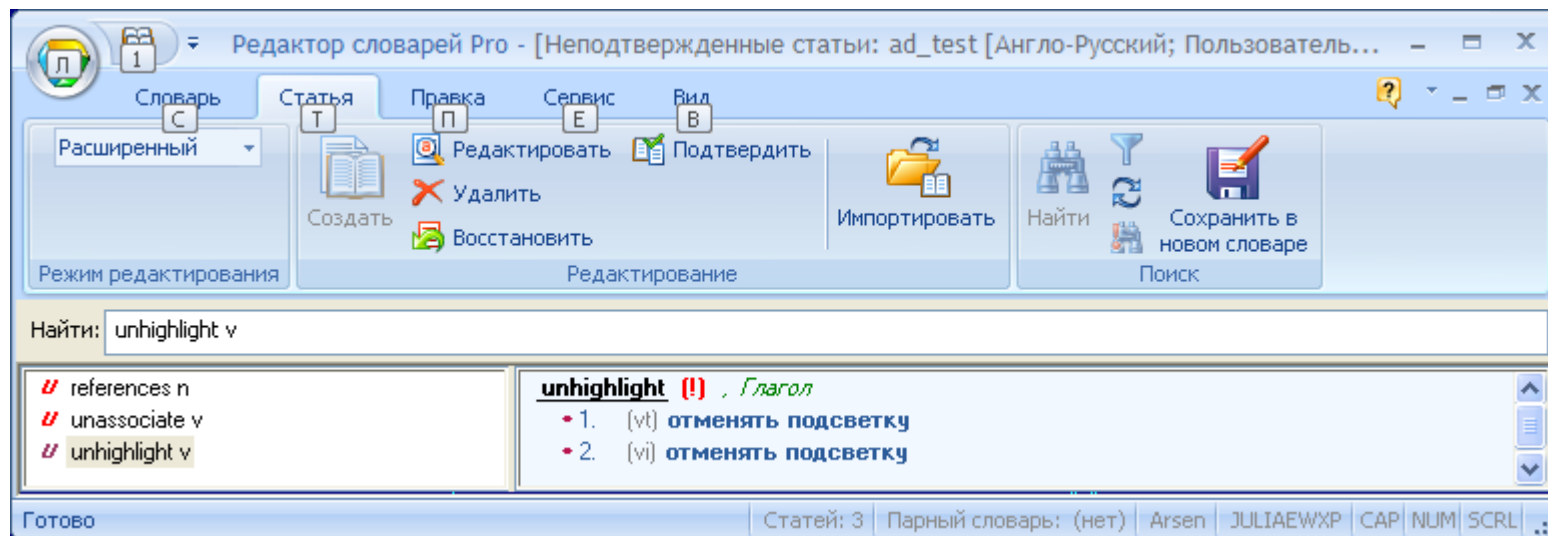


	В	С
	али	
	ная ориентация	
	ный текст по контуру	
	ная фаза узора	
	ный прямоугольник	
	по вертикали	
	по вертикали	
8	Vertical Spacer Cells	Вертикальные ячейки-разделители
9	Vertical Spacers	Вертикальный разделитель
10	Vertical Spacing	Вертикальная разрядка
11	Vertical Sync	Синхронизация по вертикали
12	vertical type	вертикальный текст
13	vertical type mask	вертикальный текст-маска
14	Vertical Type Mask Tool	Инструмент Вертикальный текст-маска
15	Vertical Type on a Path Tool	Инструмент Вертикальный текст по контуру
16	Vertical Type Tool	Инструмент Вертикальный текст
17	Vertical Type Tool	Инструмент Вертикальный текст

# Проверка отдельных словарных статей

Специальная маркировка для словарных статей, содержащих информацию, требующую экспертной проверки:

- Ключ
- Перевод(ы)



# Подбор специализированных словарей PROMT

Технология подбора включает в себя:

- экспертный отбор потенциальных словарей-кандидатов,
- сопоставление лексического состава словарей и тестовых текстов через статистический анализ,
- построение пробного перевода и его экспертный анализ.

# Специализированные словари PROMT

Компания PROMT выпускает более 170 словарей по различным областям науки и бизнеса для разных направлений перевода. В коллекцию словарей по бизнесу, например, входят словари следующих тематических областей:

- Финансы
- Менеджмент
- Экономика
- Логистика
- Бизнес-документация
- IT для бизнеса
- Коммерция
- Политика
- Юридический



# Сопоставление лексического состава словарей и текстов

**FrequencyListTool**

Список файлов для обработки:  
 \\alexeybw2k3\shared\pre-sales\АГиП.txt

Файл с результатом:  
 E:\EX\General\MARKETING\Пилоты\АГиП\frqlist1

**Действия**

Построение частотного списка

Обработка частотного списка

Прогресс обработки:

**Параметры частотного списка**

2 - минимальная длина n-граммы

4 - максимальная длина n-граммы

2 - минимальная частотность

Не удалять не алфавитные символы

- регистро-независимый частотный список

- выводить абсолютную частотность

- выводить относительную частотность

- выводить ранг (только при сортировке по частоте)

**Сортировка**

по алфавиту

Отменить

acute disseminated myositis	дерматомиозит	n	m
acute ear	острый катаральный средний отит	n	m
acute exposure	кратковременное облучение	n	n
acute hospital	больница скорой медицинской помощи	n	f
acute isolated myocarditis	идиопатический миокардит	n	m
acute management	неотложная помощь	n	f
acute MI	ОИМ	n	m
acute necrotizing pancreatitis	острый панкреонекроз	n	m
acute nonsuppurative thyroiditis	подострый тиреоидит	n	m
acute phase protein	белок острой фазы	n	m
acute splenic tumor	острое увеличение селезенки	n	n
acute stage plasma proteins	белок острой фазы	n	m
acute vascular purpura	геморрагический васкулит	n	m
acutely injured patient	шоковый пострадавший	n	m
acutenaculum	иглодержатель	n	m
acutorsion	акуторзия	n	f
acutrauma	акустическая травма	n	f

# Пробный перевод и оценка

## Difference # 4

S:

K:\Texts\@Parallel Texts\

On-Surface Coordinate System

T1:

Система координат на поверхности

T2:

Поверхностная система координат

T\_mst:

Система координат "на поверхности"

## Difference # 5

S:

K:\Texts\@Parallel Texts\

The coordinate system is created with the origin on the surface and the first axis normal to the selected surface.

T1:

Система координат создается с началом на поверхности и перпендикуляре первой оси по выбранной поверхности.

T2:

Система координат создана с началом на поверхности и перпендикуляре первой оси к выбранной поверхности.

T\_mst:

Система координат создается с началом на поверхности и первой осью по нормали к выбранной поверхности.

## Difference # 6

S:

K:\Texts\@Parallel Texts\

Click Insert > Model Datum > Coordinate System.

T1:

Выберите пункты "Вставить" (Insert) > Опорный элемент модели (Model Datum)> Система координат (Coordinate System).

T2:

Нажмите на Вкладыш (Insert)> Опорный элемент модели (Model Datum)> Система координат (Coordinate System).

T\_mst:

Выберите пункты "Вставить" (Insert) > "Опорный элемент модели" (Model Datum) > "Система координат" (Coordinate System).

# PROMT Terminology Manager (1)

**Словари, в которых термин встречается**

**Термин из исходного текста**

**Как будет выглядеть перевод этого термина в текстах**

**Примеры предложений с этим термином**

Гарантирует наличие правильного перевода существенных терминов и выражений из данного текста. Освобождает от ручного сопоставления глоссариев заказчика с текстом.

# PROMT Terminology Manager (1)

The screenshot displays the PROMT Terminology Manager interface. The top window shows the 'Expanded mode' (Расширенный режим) with a detailed view of a dictionary entry for 'extreme programming'. A red callout box points to this mode, stating: 'Расширенный режим с возможностью установки лингвистических параметров каждой статьи' (Expanded mode with the ability to set linguistic parameters for each article). The bottom window shows the 'Simple mode' (Простой режим) with a table view of the same dictionary entries. A red callout box points to this mode, stating: 'Простой режим для быстрого ввода и просмотра содержимого словаря' (Simple mode for quick input and viewing of the dictionary content). To the right, a Microsoft Word window shows the dictionary content after export, with a red callout box stating: 'Вид словаря после экспорта в Word' (View of the dictionary after export to Word).

Входное слово	Перевод	Часть речи	Комментарий	Активный перевод
Agile n	Agile	Существительное		✓
be [ 6 ] well suited v	подходить	Глагол		✓
company environment n	корпоративная среда	Существительное		✓
critical study n	основополагающее исследование	Существительное		✓
express in terms v	выражать в терминах	Глагол		✓
extreme programming n	экстрим-программирование	Существительное	to be confirmed	✓
extreme project management n	экстрим-менеджмент	Существительное		✓
in several cases d	в ряде случаев			✓
maneuverability n	способность			✓
multi-project a	мультипроект			✓

- Формирование списка переводов для извлеченных терминов.
- Разрешение противоречий в терминологии, в т. ч. поступающей от заказчика. Выбор конкретного варианта перевода, например «объем» или «объём», «опция» или «параметр».

# Правила перевода

Правила перевода – это набор настроек, которые позволяют влиять на синтез перевода, когда пользователь через специальный интерфейс выбирает один из предложенных вариантов синтеза.

**Примеры правил перевода для англо-русского направления:**

- Перевод местоимения you на русский язык (Вы, вы, ты).
- Перевод модальных глаголов личной или безличной формой (you can – Вы можете/можно, you should – Вы должны/нужно (следует)).
- Перевод императива императивом или инфинитивом (Activate the account – активируйте счет/активировать счет).
- Перевод названий компаний или сохранение оригинального написания (John Benjamins Publishing Company – John Benjamins Publishing Company/ Издательство Джона Бенджеминса).
- Выбор совершенного или несовершенного вида в пассиве (is built - построен/строится).

# Подбор правил перевода (1)

**PROMT™**

Найдено: **47 579** вхождений, **5** документов

Размер подкорпуса: **100%** от общего объема корпуса Adobe(en)

Английский: **Adobe(en)**

Русский : **Adobe(ru)**

**Make** changes as necessary to the following options, and then click OK:

Внесите необходимые изменения в следующие параметры, а затем нажмите кнопку «OK».

**Drag** the FLVPlayback and FLVPlaybackCaptioning components into the Library panel.

Перетащите компоненты FLVPlayback и FLVPlaybackCaptioning на панель «Библиотека».

**Use** this behavior to create puzzles, slider controls, and other movable interface elements.

Используйте это поведение, чтобы создать пазлы (игру), элемент управления «ползунок» и другие перемещаемые элементы интерфейса.

**Specifies** a data file for the ActiveX control to load.

Указывает файл данных для загрузки элемента ActiveX.

**Use** the Size control to adjust the grid spacing and the Color control to change the color of the grid.

Размер ячеек сетки настраивается с помощью элемента управления «Размер», а цвет линий — с помощью элемента управления «Цвет».

**Select** whether to save the channels in an interleaved or non-interleaved order.

Выберите, как должны сохраняться каналы: с чередованием или без чередования.

**Paste** and move collapsed code fragments

Вставка и перемещение свернутых свернутых фрагментов кода

Корпус: Adobe(en)  
Язык: english

Настроить корпус

словоформа    лемма

Омонимия:

До снятия

Часть речи

Синтаксис...  
Verb

Пунктуация:  
слева    справа

Позиция в предложении:  
в начале

Регистр:  
с большой буквы

Искать отрицание

Настройки поиска: дополнительно

Искать    Очистить

Настройки выдачи  
Поиск в новом окне

При наличии баз ТМ для подбора соответствующих моделей синтеза используется PROMT Corpus Search, который позволяет осуществлять поиск в проиндексированных параллельных корпусах.

**PROMT®**

# Подбор правил перевода (2)

Подбор правил синтеза в экспертном режиме с учетом типа и стиля текста.

Настройки перевода

Название	Значение
<b>Переводимые элементы</b>	По умолчанию
<b>Режим перевода</b>	«Умный» перевод
Перевод модальных глаголов	личный
Перевод причастия I после условных союзов	предлог "при" + существительное
Перевод герундия после предлога "by"	предлог "путем" + существительное
Перевод инфинитива цели	придаточное предложение
Порядок следования главных членов в нераспространенном предложении	переставлять
Порядок частей в сложных условных предложениях	переставлять
Перевод временных придаточных с подлежащим "you", вводимых союзами "after/b...	предлог "после/перед" + существительное
Перевод условных и временных придаточных с подлежащим "you"	придаточное предложение
Перевод объектных придаточных с подлежащим "you"	придаточное предложение
Перевод придаточных со сказуемым в сослагательном наклонении и подлежащим...	инфинитив
Перевод определительных придаточных	причастие
Перевод оборота с причастием "using"	предлог "с помощью"
<b>Стиль текста</b>	По умолчанию
<b>Разрешение грамматической неоднозначности</b>	По умолчанию

**Перевод причастия I после условных союзов**  
Английские временные, уступительные и условные конструкции с причастием I и союзами "when", "while", "if" могут переводиться разными способами - с помощью деепричастного оборота или придаточного предложения и существительного или инфинитива с предлогом. Тип синтеза «существительное/инфинитив с предлогом» позволяет придать переводу более формальный характер.

**ПРИМЕРЫ:**  
деепричастный оборот: Be careful while turning the device on. -> Будьте осторожны, включая устройство.  
предлог "при" + существительное: Be careful while turning the device on. -> Будьте осторожны при включении устройства.

# Тестирование результатов

- Сравнение результатов машинного перевода до и после настройки.
- При наличии параллельных текстов – сравнение результатов машинного перевода с переводами, выполненными человеком, через PROMT QALab.
- Корректировка параметров настройки и дополнительное обучение системы.

# PROMT Quality Assurance Lab

PROMT QALab показывает процент отличий в переводах и позволяет сопоставить результат машинного перевода с эталонным переводом.

The screenshot displays the PROMT QA Lab interface for a translation comparison. The window title is "PROMT QA Lab - Перевод 2". The menu bar includes "Файл", "Правка", "Действие", "Вид", and "Справка". The language is set to "Англо-Русский" and the device is "Аппаратное обеспечен".

Исходный текст	Перевод 1	Перевод 2
<p>Введите имя файла</p> <p>In Internet Explorer, go to the web page. In Windows, you can switch to a different preset from within the Adobe PDF Settings dialog box. In the JDF Job Definitions dialog box, select a job definition, and click Edit. If the Preflight dialog box is not open, choose Advanced &gt; Preflight. In the Contacts section, click Add. If another panel is displayed in the Preflight dialog box, click the Profiles tab. In the JDF Job Definitions dialog box, click Contacts Manager. In the Preferences dialog box under Categories, select Documents. For the greatest control over how comments are printed, choose Comments &gt; Print With Comments Summary. Content reference to: Adobe PDF presets Type the appropriate information in the email message that opens after the conversion is complete. To remove the part from the model, right-click and choose Part Options &gt; Hide. Convert part of a web page to PDF web pages:adding to PDFs Drag the pointer to select text and images on a web page. Right-click on the selected content and choose one of the</p>	<p>Введите имя файла</p> <p>В Internet Explorer, <b>пойдите в web-страницу</b>. В Windows <b>можно перейти к различным установкам изнутри диалогового окна Adobe PDF Settings</b>. В диалоговом окне JDF Job Definitions, выберите <b>определение работы</b>, и нажмите <b>Edit</b>. Если диалоговое окно <b>Preflight</b> не открыто, выберите <b>Advanced</b>&gt;, <b>Выполняют предпочтательную проверку</b>. В разделе <b>Контакты</b>, нажмите <b>Add</b>. Если <b>другая панель отображается</b> в диалоговом окне <b>Preflight</b>, нажмите <b>табуляторы Профили</b>. В диалоговом окне JDF Job Definitions, нажмите <b>Contacts Manager</b>. В <b>Диалоговом окне настроек под Категориями</b>, выберите <b>Документы</b>. <b>Для самого большого контроля над тем, как комментарии печатаются, выберите Comments</b>&gt; <b>Печать Со Сводкой Комментариев</b>. Ссылка на раздел к: <b>Стили Adobe pdf</b> <b>Тип соответствующая информация в электронном письме, которое открывает</b> после преобразования, <b>полон</b>. <b>Чтобы удалить часть</b> из модели, щелкните правой кнопкой <b>мыши</b> и выберите <b>Part Options</b>&gt;, <b>Скрывают</b>. <b>Преобразуйте часть web-страницы к PDF</b> <b>сеть pages:adding к PDFs</b></p>	<p>Введите имя файла</p> <p>В <b>браузере Internet Explorer перейдите к web-странице</b>. В Windows в диалоговом окне «Параметры Adobe PDF» <b>можно переключиться на другой набор параметров</b>. В диалоговом окне «Описания заданий JDF» выберите <b>описание задания</b> и нажмите кнопку «<b>Редактировать</b>». Если диалоговое окно «Предпочетная проверка» не открыто, то выберите меню «<b>Дополнительно</b>» &gt; «<b>Предпочетная проверка</b>». В разделе «<b>Контактные лица</b>» нажмите кнопку «<b>Добавить</b>». Если в диалоговом окне «Предпочетная проверка» <b>отображается еще одна панель</b>, нажмите вкладку «<b>Профили</b>». В диалоговом окне «Описания заданий JDF» нажмите кнопку «<b>Диспетчер контактов</b>». В диалоговом окне «<b>Установки</b>» в разделе «<b>Категории</b>» выберите «<b>Документы</b>». <b>Чтобы получить больший контроль над процессом печати комментариев, выберите меню «Комментарии»</b> &gt; «<b>Печать со сводкой комментариев</b>». Ссылка на раздел: «<b>Стили Adobe PDF</b>» Введите <b>нужные данные</b> в сообщение электронной почты, которое <b>открывается</b> после завершения преобразования.</p>

Готово

Изменений в абзаце: 55.6818% Всего изменений: 65.9574%

# Adobe

- Один из крупнейших в мире производителей ПО для графического и веб-дизайна.
- Доход за 2010 г. – ок. 4 миллиардов долларов.
- Более половины этих доходов получено за пределами США.
- Важность локализации программных продуктов/перевода документации (более 100 продуктов/34 языка).

# Тип переводимых текстов

- элементы пользовательского интерфейса
- online-помощь
- FAQ
- руководство пользователя
- сообщения об ошибках

# История внедрения

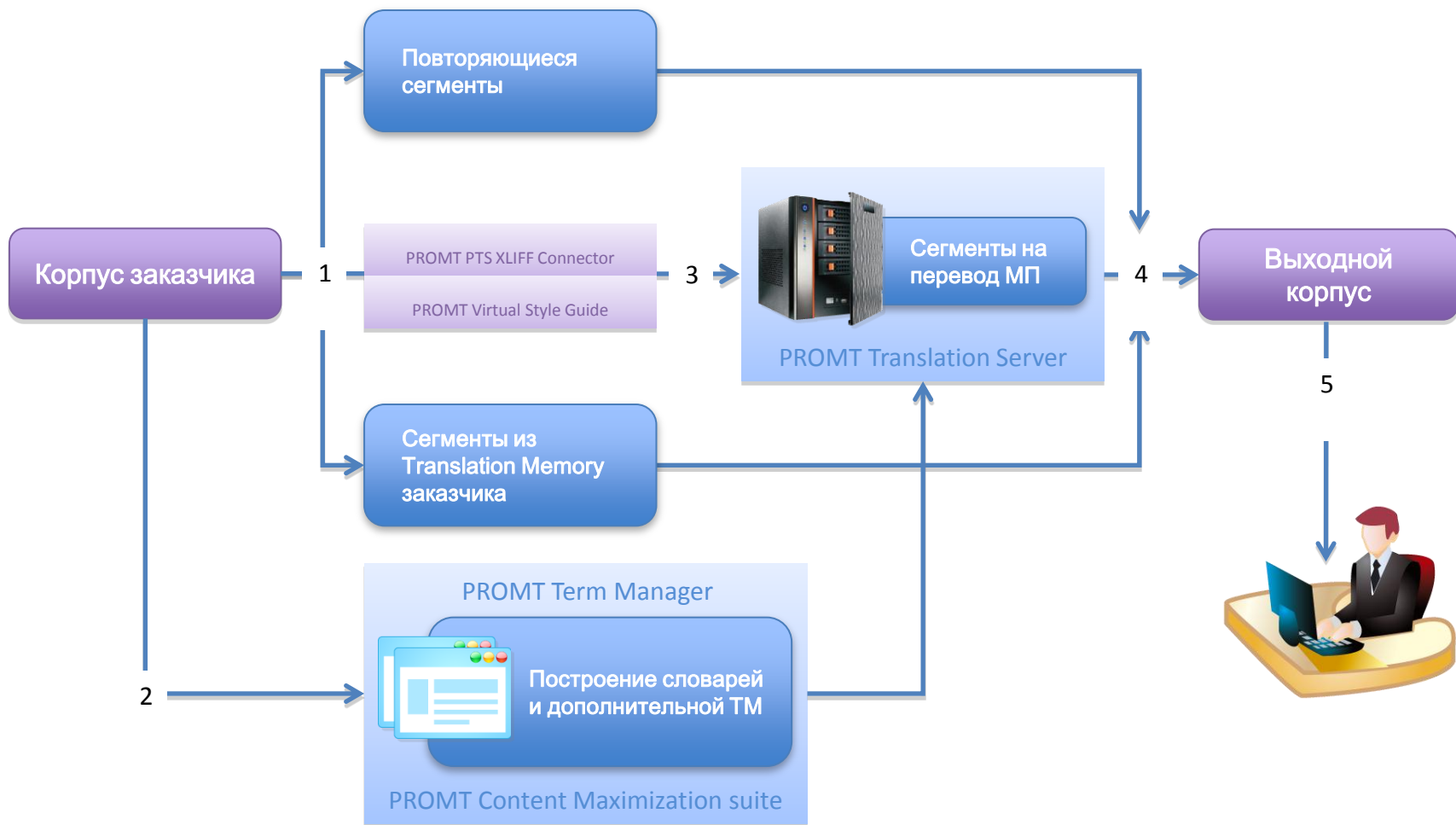
- Пилотный проект (2 месяца): демонстрация возможностей англо-русского перевода.
- Развертывание готового решения на базе пилотного проекта и внедрение в процесс документооборота.
- Глубокая настройка системы.
- Постоянная лингвистическая поддержка со стороны PROMT.

# Процесс перевода в Adobe (1)

- Для управления переводческими проектами используется SDL Worldserver (файлы в формате DITA XML).
- SDL Worldserver + MT connector (разработка PROMT).
- Сегменты с машинным переводом редактируются человеком (перевод, выполненный человеком).

Основная идея: получить человеческий перевод из машинного быстрее, чем переводить «с нуля», при условии предварительного обучения системы МП под конкретную задачу.

# Процесс перевода в Adobe (2)



# Предварительная обработка текстов

Исходные тексты проходят проверку на правописание и стилистическое единство.

Разделение на группы

- дублирующие друг друга сегменты,
- сегменты, найденные в Translation Memory, с совпадением более 75 %,
- уникальные сегменты → машинный перевод.

# Обработка терминологии и подбор правил синтеза

- Обработка глоссариев и создание словарей МП.
- Подбор спецсловарей PROMT.
- Обработка терминологии с помощью PROMT Terminology Manager и подбор переводов через PROMT Content Maximization suite.
  - набор словарей для конкретного корпуса и решение проблемы перевода терминологии разных продуктов внутри одной линейки.
- Подбор правил перевода (PROMT Virtual Style Guide).

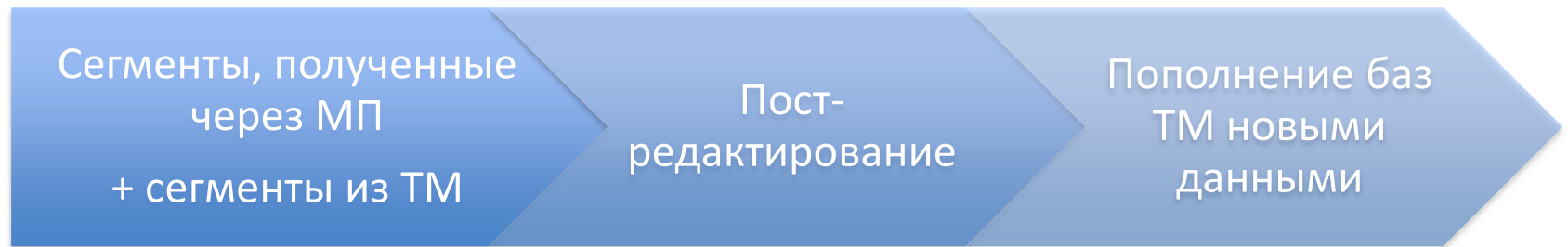
# Сохранение форматирования

PROMT Translation Server XLIFF Connector → уникальные сегменты переводятся в формат XLIFF для сохранения форматирования внутри строки.

*Check `NativeApplication.supportsSystemTrayIcon` to determine whether system tray icons are supported on the current system.*

*Проверить `NativeApplication.supportsSystemTrayIcon` чтобы определить, поддерживаются ли значки в системе в текущей системе.*

# Постредактирование



Технология PROMT обеспечивает накопление информации, что увеличивает экономию времени с каждым новым переводом.

# Пример перевода

When there are multiple databases connected in RoboSource Control, there could be an issue when the database is not found and cannot be connected to.

## Solution

---

This is caused by RoboSource Control storing the information for the last database which was logged into. If this does not match then the error message will be returned.

Workaround:

To fix this issue, open the RoboSource Control client and switch the database which is active. Then close RoboSource Control and open the project through RoboHTML.

---

Когда есть многократные базы данных, подключенные в RoboSource Control, могла быть проблема, когда база данных является не найденной и не может быть подключена к.


## Решение

---

Это вызвано RoboSource Control, хранящим информацию для последней базы данных, которая была зарегистрирована в. Если это не совпадет тогда, то сообщение об ошибке будет возвращено.

Обходной путь:

Чтобы исправить эту проблему, откройте клиента RoboSource Control и перейдите базу данных, которая является активной. Тогда закройте RoboSource Control и откройте проект через RoboHTML.



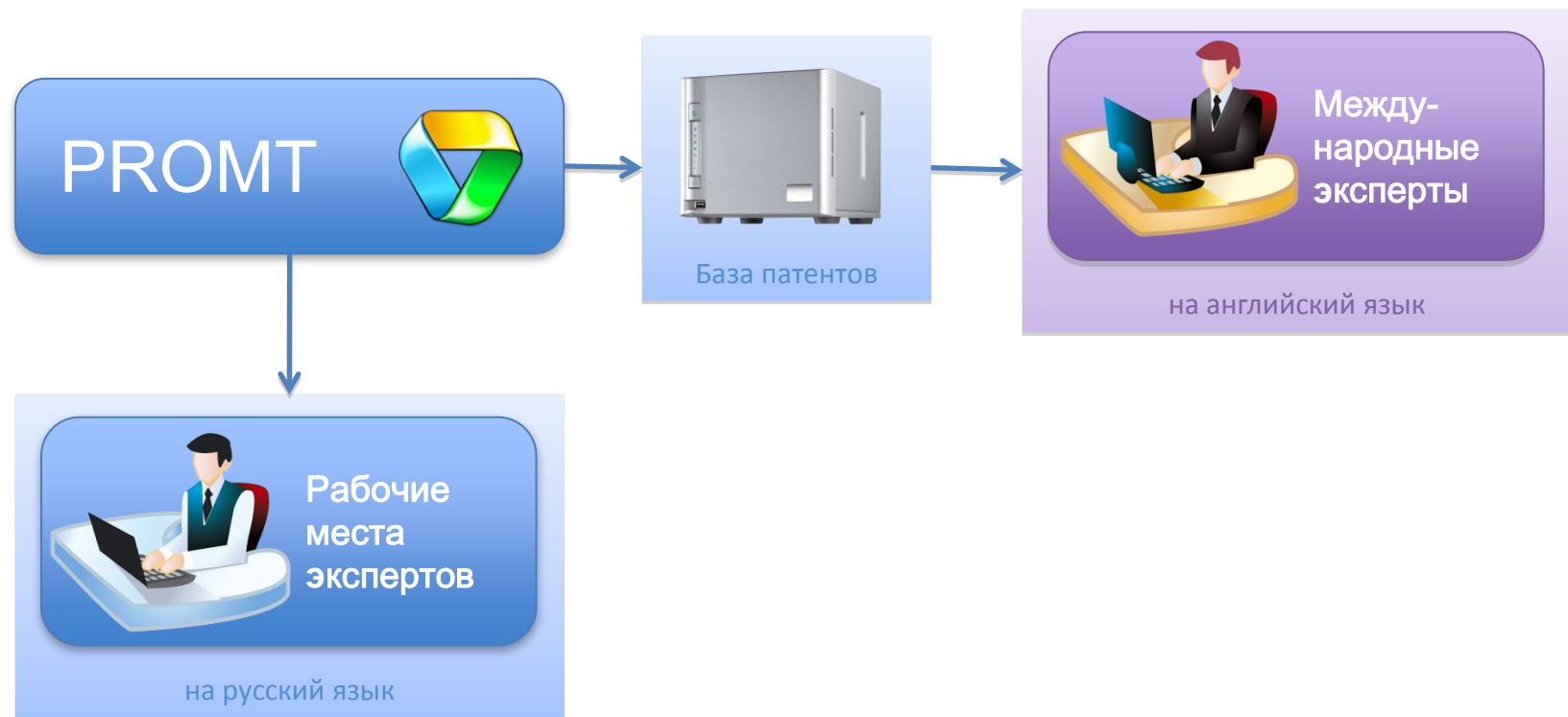
# Что дало решение PROMT в локализации продуктов Adobe

Скорость подготовки документации  
увеличилась на 20–30 %.

# ФГУ ФИПС

- Прием и экспертиза заявок на объекты патентного права, в т. ч. международных заявок.
- Обработка международных заявок на товарный знак.
- Выпуск официальных бюллетеней о зарегистрированных свидетельствах.
- Справочно-информационное обслуживание.

# Схема использования



# Перспективы ближайшего развития

- Создание пользовательских словарей экспертами в процессе актуальной эксплуатации.
- Аудит пользовательских словарей со стороны PROMT и внедрение их в общий документооборот.
- Создание словарей противоположного направления и внедрение их в общий документооборот.



Спасибо!

Юлия Епифанцева,  
директор по развитию бизнеса

PROMT

[www.promt.ru](http://www.promt.ru)

**PROMT**<sup>®</sup>