

# Машинный перевод для профессионалов

**Юлия Епифанцева,**  
*Директор по развитию PROMT*

# PROMT

С 1991 одна из ведущих в мире лингвистических IT-компаний

- **20+ языков**, 80+ языковых пар
- **Технологии:** PROMT Rule-based MT, SMT, Hybrid MT
- **Клиенты:** Лаборатория Касперского, TripAdvisor, PayPal, PTC, Cisco, Amadeus, Федеральный институт промышленной собственности РФ, LSA, Multilizer и др.
- **Лучшее качество перевода с английского на русский** по мнению экспертов на [ежегодном конкурсе](#) автоматических переводчиков под эгидой ACL с 2013г.

# Плюсы и минусы машинного перевода

## Значительно упрощает процесс перевода

- ✓ Высокая скорость
- ✓ Экономия затрат на перевод
- ✓ Конфиденциальность
- ✓ Предсказуемость результата  
*(нет человеческого фактора)*

## но требует участия пользователя

- Необходимо контролировать качество перевода
- Необходимо овладеть инструментарием для настройки
- Необходимо иметь навыки постредактирования

# Быть или не быть МП в проекте

- **Большой объем документации** с короткими сроками на выполнение перевода
- **Однотипная документация:** техническая документация, мануалы, отчеты, пользовательский контент
- **Четкие требования к результатам** перевода
- **Наличие материалов для адаптации** (настройки) решения по МП (*базы ТМ, словари, глоссарии*)

# Online-сервисы или PROMT for SmartCAT

PROMT

Online-service

Настройка перевода



Гарантия конфиденциальности



Перевод без интернета



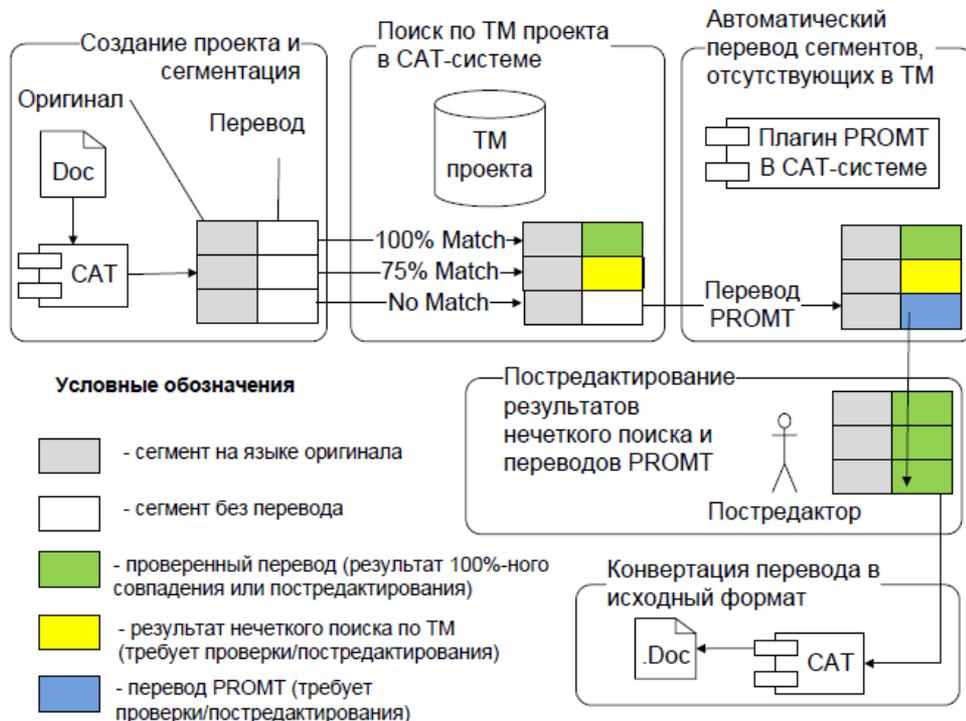
# Решение для SmartCAT

Платформа

PROMT Translation Server 12

- Клиент-серверное решение с поддержкой любых технологий перевода: Rule-based MT, SMT, Hybrid MT
- Возможность обучения системы на материалах заказчика
- Результаты настройки в виде профилей перевода (без ограничения их количества для одной языковой пары)
- Онлайн и оффлайн перевод

# Схема интеграции



# Технологии PROMT

## RBMT

Rule-based machine translation

---

### Компоненты:

- ▶ двуязычные словари
- ▶ морфологические таблицы
- ▶ списки префиксов
- ▶ базы имен

## SMT

Statistical machine translation

---

### Компоненты:

- ▶ фразовая таблица
- ▶ языковая модель

# Технологии PROMT

## RBMT

Rule-based machine translation

---

### Преимущества:

- ▶ синтаксическая и морфологическая точность
- ▶ предсказуемость результата
- ▶ возможность настройки на предметную область

## SMT

Statistical machine translation

---

### Преимущества:

- ▶ быстрая настройка (параллельные тексты)
- ▶ гладкость перевода

# Технологии PROMT

## RBMT

Rule-based machine translation

---

### Недостатки:

- ▶ необходимость поддерживать и актуализировать лингвистические базы данных
- ▶ «машинный акцент» при переводе

## SMT

Statistical machine translation

---

### Недостатки:

- ▶ многочисленные грамматические ошибки
- ▶ нестабильность перевода

# Методы настройки PROMT

- Настройка силами специалистов PROMT через заказ в SmartCAT
- Самостоятельная настройка пользователем через дополнительные инструменты PROMT
- Настройка на имеющихся материалах заказчиков (базы ТМ, глоссарии, словари)
- Возможна настройка без предоставления материалов заказчика

# Средства настройки PROMT

- **Специализированные словари**  
с терминологией из определенных предметных областей
- **Пользовательские словари**  
с терминологией заказчика
- **Правила перевода**  
для стилистических и синтаксических предпочтений
- **Списки Do-Not-Translate Words**
- **Статистическая модель**  
на основе параллельных текстов от 100 тыс. сегментов

# Профили перевода PROMT

## RBMT

Rule-based machine translation

---

- ▶ Набор словарей
- ▶ Списки Do-Not-Translate Words
- ▶ Правила перевода

## SMT

Statistical machine translation

---

- ▶ Корпуса параллельных текстов, релевантных для данного проекта

# Постредактирование МП

## Исходный текст

*Distributor may not take legal action against any infringer without SDA's prior written consent.*

## Машинный перевод:

*Дистрибьютор не может подать в суд против никакого нарушителя без предварительного письменного согласия SDA.*

## Итоговый результат:

*Дистрибьютор не имеет права подавать в суд на какого-либо нарушителя без предварительного письменного согласия SDA.*

Объем постредактирования : **34%**

# Преимущества использования PROMT в SmartCAT

- **Ускорение переводческих процессов** - время на редактирование существенно меньше, чем время на перевод «с нуля»
- **Сокращение рутинных операций** для переводчика и упор на творческую часть процесса перевода
- Широкие **возможности обучения и настройки** переводчика
  - *Терминологические словари, в т.ч. и на базе глоссариев переводчиков*
  - *Управление стилем перевода через «правила перевода»*
- **Снижение влияния человеческого фактора**
- **Онлайн и офлайн перевод**
- **Техническая и лингвистическая поддержка**

# Спасибо за внимание!

**Юлия Епифанцева**

Директор по развитию PROMT

[julia.epiphantseva@prompt.ru](mailto:julia.epiphantseva@prompt.ru)