

Информационная безопасность и интеллектуальный анализ неструктурированных текстовых данных для бизнеса и государства



НИКИТА ШАБЛЫКОВ
коммерческий директор PROMT

Неструктурированные текстовые данные. Вызовы

Сохранность значимой информации

- ▶ Необходимость обработки информации в закрытой от интернета среде
- ▶ Многоязычность и нехватка лингвистической компетенции у специалистов

Поиск и извлечение значимой информации

- ▶ Невозможность оперативного отслеживания потенциальных угроз в потоке Больших Данных
- ▶ Разные форматы (текст, аудио, видео, графика)

Критическая информация для бизнеса и государства

Сохранность значимой информации

- ▶ Перевод конфиденциальной документации в корпоративных сетях, закрытых контурах
- ▶ Интеграция функций перевода в специальные программные комплексы и решения

Поиск и извлечение значимой информации из данных

- ▶ Извлечения фактов и сущностей из многоязычных потоков
- ▶ Интеграция технологий анализа текстов на естественных языках в поисковые системы, ИА и BI системы

PROMT Translation Server

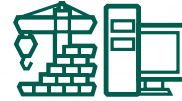
Система перевода текстов,
документов различных форматов,
текстовых Больших Данных (Big Data)
без обращения в Интернет



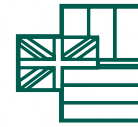
Модули для перевода речи



Перевод текстов, сайтов и
документов с сохранением
форматирования



Настройка перевода на
предметную область



20+ языков

80+ языковых комбинаций



Лучшие оценки экспертов
за перевод с английского
на русский в 2013-2016гг

Почему не онлайн-сервис?

Онлайн-сервисы – **один из каналов утечки:**

Барак Туровски,
руководитель проекта
[Google Translate](#)



*«... я передам вам наши логи,
и вы сможете посмотреть
на все запросы, которые
мы обрабатываем»*

7 марта 2017

PTS установлен внутри организации
пользователя

PROMT 
Translation Server 18

**исключает
несанкционированный
доступ к данным**

PROMT Analyzer SDK

Система выделения смысла из любых неструктурированных данных



Выделение связей между сущностями



Точное извлечение сущностей (персоналий, организаций, событий) и фактов из текстов – до 95%



Определение тональности высказываний



Выделение характеристик найденных сущностей



Финалист конкурса Информационно-аналитических инструментов 2017

Европейский проект от Еврокомиссии



Задача:

- ▶ Сбор информации из интернета (Surface web, Deep Web, Dark Web)
- ▶ Анализ информации на разных языках, извлечение сущностей, связанных с терроризмом (персоналий, организаций, событий)
- ▶ Фильтрация, кластеризация, семантическая классификация, резюмирование информации

Решение:

Выявление и анализ онлайн-контента, связанного с финансированием террористической деятельности и вербовкой

Пользователи:

Сотрудники правоохранительных органов

Мобильное приложение для ФПК



Федеральная
пассажирская
компания

Задача:

Создание доступной и открытой международной коммуникативной среды

Пользователи:

Сотрудники компании, предоставляющей услуги по перевозке пассажиров и грузобагажа железнодорожным транспортом

Решение:

Приложение PROMT для перевода текста и голоса

- ▶ В режиме оффлайн на устройствах контроля легитимности электронных билетов (УКЭБ)
- ▶ На сервере штабного вагона ЕИИПП АО «ФПК»
- ▶ Высокая скорость перевода специализированной лексики (ж/д транспорт, чрезвычайные ситуации и т.д.)

Преимущества

- ▶ **Качественный перевод конфиденциальной документации в закрытой от интернета среде**
- ▶ **Работа с большими объемами текстовых данных в режиме реального времени**
- ▶ **Многоязычность: 20+ языков и 88+ языковых пар**
- ▶ **Гибкая настройка под конкретные задачи**

Клиенты PROMT



УПРАВЛЕНИЕ ДЕЛАМИ
ПРЕЗИДЕНТА РФ





**Бесплатный доступ к демо-
версии PROMT Translation Server
и PROMT Analyzer SDK**

для всех участников конференции
«Critical Communications Russia 2018»

НИКИТА ШАБЛЫКОВ

Коммерческий директор

Nikita.Shablykov@promt.ru

+7 (495) 580-4848

+7 911 836-8704

Critical Communications Russia 2018