

# О тенденциях и технологиях

# ПРОМТ в мировом пространстве

- ПРОМТ Россия
- ProMT GmbH
- ProMT Americas

# Рынок услуг по переводу и рынок машинного перевода (МП)

- Рынок услуг по переводу текстовой информации огромен: он составил более 10 миллиардов долларов в 2010 году (Common Sense Advisory)
- Рынок МП и услуг, связанных с МП, в 2010 составил около 2 миллиардов долларов (WinterGreen Research)

# Для каких целей используется машинный перевод (МП)

- Для общения
- Для получения информации из текстов на иностранном языке
- Для получения перевода текста с целью публикации

# Ядро системы – искусственный интеллект

- Для разных целей нужны совершенно разные продукты, с разным интерфейсом и разными возможностями настройки
- Общая часть всех продуктов, online-сервисов, плагинов – модуль перевода

# Команда разработчиков

- В ПРОМТ R&D работает около 70 человек
- Из них одна группа занята лингвистическими исследованиями и развитием ядра системы – модулей перевода на основе технологии rule-based, другие группы работают над гибридной технологией и статистическим переводчиком
- Выделенная группа занимается разработкой средств разработки

# Команда разработчиков

Средства разработки – это:

- специальные тестеры для поддержки эволюционирования системы
- генераторы объектов тестирования
- средства сравнения парсинга
- средства для поиска в текстовых корпусах любого размера
- статистические модели языка

# Как устроены модули перевода и что нового в PROMT 9.0

За последние годы мы:

- научили систему общаться со сложно устроенными семантическими моделями и заложили технологическую основу для развития семантического компонента системы
- эффективно применили в разработке корпусные и статистические методы для значительного расширения синтаксических и семантических правил
- создали гибридную технологию перевода (DeepHybrid) для английского языка и внедрили ее у первых клиентов
- создали собственную технологию статистического перевода

# Что делает R&D сегодня

- Фокус исследований – продолжение работы над гибридной моделью модуля перевода, которая объединяет преимущества rule-based систем и статистических систем перевода
- Основы статистических моделей:
  - моноязычные текстовые корпуса в несколько миллиардов словоупотреблений и, в частности, текстовые корпуса наших online-сервисов и наших клиентов
  - параллельные двуязычные тематические корпуса, например, корпус Европарламента в несколько десятков миллионов словоупотреблений
  - С 2011 – модуля перевода для английского, работающий по гибриднему принципу

# Потребность в машинном переводе

Факторы, влияющие на рост интереса к машинному переводу:

- контент компаний практически полностью становится цифровым и существует в стандартных форматах
- сокращение времени появления локализованной информации о продуктах и услугах, сокращение расходов на эту деятельность – требование рынка
- люди абсолютно готовы воспринимать перевод, сгенерированный компьютером, когда используют эту услугу для целей общения

# Потребность в машинном переводе

- МП не заменяет человеческий перевод
- МП – это необходимый инструмент, который создает дополнительные возможности по сокращению времени и затрат на изготовление перевода
- МП нужен там, где человеку не справиться с потоком информации в реальном времени



Спасибо за внимание!

Светлана Соколова  
Генеральный директор  
компании ПРОМТ

---

**PROMT**<sup>®</sup>