



Нейросетевой машинный перевод в бизнесе

Юлия Епифанцева, PROMT

ПЕРЕВОД КАК СТРАТЕГИЧЕСКАЯ ЗАДАЧА

- За 1950-2000гг мировая торговля выросла в 20 раз, производство в 6 раз. Объем информации, произведенной человечеством, удваивается каждые 15-16 месяцев.
- Для охвата 90% аудитории необходимо
 - 2001 – 13 языков
 - 2020 – 48 языков
- Евросоюз. 28 языков. Информация должна быть представлена на всех языках ЕС.
- Программы ЕС по развитию науки и технологий Horizon 2020 (бюджет на 7 лет 80 миллиардов евро).
- Выручка каждой из ста крупнейших переводческих компаний в мире составляет от 10 до 700 миллионов долларов.
- По данным Global Market Insights, Inc. к 2021 году объем мирового рынка лингвистических услуг составит 56 миллиардов долларов.

ИИ: ЗАДАЧИ, РЕШАЕМЫЕ ЧЕЛОВЕКОМ – МОЖЕТ РЕШАТЬ МАШИНА

Лингвистические задачи в области ИИ

- Распознавание речи
- Классификация текстовых сообщений
- Идентификация людей по голосу
- Перевод текста

Машинный перевод- одна из самых
успешных реализаций ИИ

Предпосылки для развития МП на основе нейронных сетей

- Рост вычислительных мощностей
- Развитие нейросетевых технологий
- Рост цифрового контента

Данные в цифровой экономике
– это новая нефть

ПОЧЕМУ МАШИННАЯ ОБРАБОТКА ЕСТЕСТВЕННОГО ЯЗЫКА СЛОЖНАЯ ЗАДАЧА



МНОГООБРАЗИЕ КОНТЕНТА

Заголовки и большие тексты, распознанная устная речь, терминология из разных отраслей науки и бизнеса, деловой, научный стиль и сленг представляют собой «мини-языки» внутри большого языка



АСИММЕТРИЯ ЯЗЫКОВ

Разные языки по-разному описывают мир и используют разные подходы в морфологии, грамматике, синтаксисе, семантике



ДАННЫЕ ДЛЯ ОБУЧЕНИЯ

Сложности в поиске и подборе данных для обучения и настройки систем на перевод под конкретные бизнес-задачи

**FOLHFY
BVDETJ
ESZXYD**

СРЕДНЯЯ ДЛИНА СЛОВ В РАЗНЫХ ЯЗЫКАХ

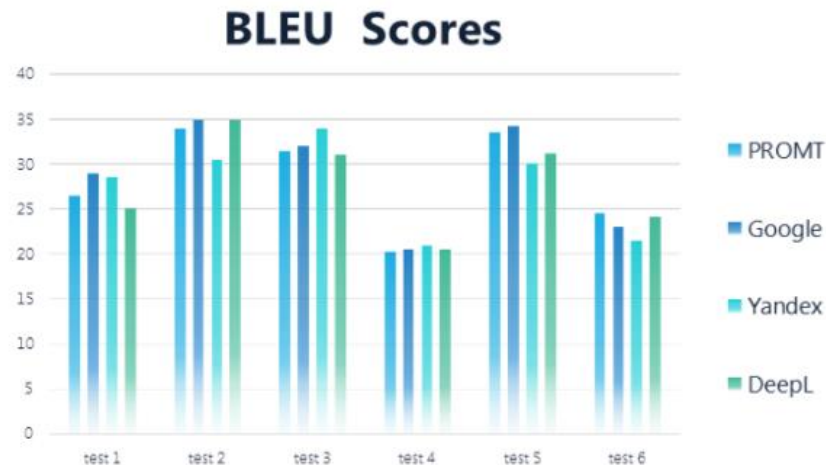
Увеличение или уменьшение количества символов в тексте перевода по сравнению с оригиналом влияет на сохранение форматирования документа при переводе

PROMT Neural, Google, Yandex, DeepL

ТЕСТИРОВАНИЯ НА ТЕКСТАХ
ОБЩЕЙ ТЕМАТИКИ
АНГЛО-РУССКАЯ ПАРА

Экспертная оценка:

- Перевод легко читается
- Корректная структура предложений
- Нет ошибок согласования
- Перевод требует незначительного или вообще не требует редактирования



СЦЕНАРИИ ПРИМЕНЕНИЯ МАШИННОГО ПЕРЕВОДА

СЛУЖБЫ ПОДДЕРЖКИ

для оперативного ответа на запросы

E-COMMERCE

локализация контента для покупателей

КОРПОРАТИВНАЯ ПЕРЕПИСКА

общение внутри глобальных компаний

ДОКУМЕНТАЦИЯ К ПРОДУКЦИИ

документация к ПО, техническая документация

ЛОКАЛИЗАЦИЯ САЙТОВ

перевод контента сайтов на лету

ПОИСК ИНФОРМАЦИИ

в мультязычных БД и сервисах

ИНТЕГРАЦИЯ В ПРИЛОЖЕНИЯ

как дополнение к основным функциям

ОБРАЗОВАТЕЛЬНЫЕ СЕРВИСЫ

для изучения языков

Взаимодействие пользователей в онлайн-пространстве

Крупнейший в мире туристический
сайт (переведен на 21 язык)

Основной контент – миллионы отзывов
пользователей со всего мира
(более 385 миллионов отзывов)

Машинный перевод с английского на русский
60 000+ запросов и 450 000+ слов в день



Обеспечение работы служб поддержки и общения с клиентами

Одна из крупнейших в мире глобальных систем онлайн-дистрибуции билетов

Условия бронирования формируются в режиме реального времени
на английском языке

Более **5 000** пользователей
(туристических агентств в России)

Более **6 000** запросов на перевод или
более **4 млн** слов в неделю

amadeus



Международное сотрудничество

Перевод внутренней документации с английского, немецкого, испанского и др.

Пользователи в филиалах по всему миру (Австралия, Северная и Южная Америка, ЮАР, Финляндия, Болгария и др.)

Быстрый перевод для понимания деловой переписки, юридических и бухгалтерских документов

Экономия времени и ресурсов - перевод у бизнес-пользователей в режиме реального времени



Оптимизация затрат на профессиональный перевод

Обновление документации
на сайтах и в новых версиях
продуктов

Перевод в короткие сроки
с английского на 30 языков

С машинным переводом рост
производительности
переводчиков на 15+%



ЧМ 2018 в России

Мобильный переводчик,
работающий без Интернета и
интегрированный в рабочий
инструмент проводников

Голосовой и текстовый перевод

Тематические разговорники

Режим «Диалог» для общения в
режиме реального мнения с
иностранцами



COVID. Сервисы для врачей

Российский сервис для врачей, где интегрированы технологии перевода PROMT для перевода статей и другой медицинской и фармакологической информации из зарубежных источников.

В марте и апреле 2020 общее количество запросов на перевод увеличилось в 160 раз, а объемы перевода – в 55 раз!



Сколько книг переводится на PROMT.One в день?

Всего переведено за

1 год = 69 млрд слов

В среднем за

1 день = 190 млн слов

Для перевода 190 млн слов нужно:

31 717 дней = 87 лет

работы без перерыва 1 переводчика

76 120 переводчиков,
чтобы перевести все за **1 день**



WWW.TRANSLATE.RU

2016

5 ПРИЧИН ЗАДУМАТЬ О ПРИМЕНЕНИИ МАШИНОГО ПЕРЕВОДА СЕГОДНЯ

- Необходимость перевода с и на иностранные языки никогда не исчезнет
- Технологии будут все больше и больше использоваться для дополнения и расширения работы профессиональных переводчиков
- Нейронный машинный перевод — это инновация и революция в области ИИ
- Актуальный вызов – это не только перевод текста, но и перевод документа
- Вопросы безопасности в ИТ не утратят своей значимости



ARTIFICIAL
INTELLIGENCE
JOURNEY

Спасибо за внимание!

Julia.Epiphantseva@prompt.ru