

**ПРЕМЬЕРА ГОДА**

# КАСТОМИЗИЦИЯ МАШИННОГО ПЕРЕВОДА:

как получить индивидуальную СМП  
для своей отрасли, бизнеса, проекта

Юлия Епифанцева, PROMT  
Маргарита Меняйлова, EGOTECH

**16 / 03 / 2022**

## О PROMT

- Отечественный разработчик ПО в области ИИ
- Лучшие эксперты в области МП России с 1991г.
- Один из ведущих вендоров МП в мире
- Уникальная команда разработчиков – более 40 программистов и лингвистов
- Все продукты компании внесены в Единый реестр отечественного ПО

## Участие в WMT - главной конференции по МП

С 2013г. PROMT участвует в Workshop on Machine Translation (WMT) (<http://www.statmt.org/>) в направлении по переводу новостных текстов.

В 2021 году PROMT участвовал в Workshop on Machine Translation со своей технологией управления терминологией в нейросетевом машинном переводе.  
Лучшие результаты по настройке англо-русской языковой пары.

# О МНОГОЦЕЛЕВОМ ЛИНГВИСТИЧЕСКОМ ПРОВАЙДЕРЕ ЭГО ТРАНСЛЕЙТИНГ

Группа компаний ЭГО Транслейтинг (осн. 1990 г.) – многоцелевой лингвистический провайдер, лидер цифровых технологий в лингвистике, один из флагманов российского и восточноевропейского рынков лингвистических услуг.



**ТОП-5 на рынке Восточной Европы**

(по данным международного агентства  
Common Sense Advisory)



**ТОП-5 по комплексу услуг и №1  
по устным переводам в России**

(по версии [translationrating.com](http://translationrating.com))



**Действующий член Европейской  
ассоциации переводческих компаний ELIA**  
(European Language Industry Association)

Технологический суббренд  
ГК ЭГО Транслейтинг

Эксперты по комплексной поддержке  
внедрения, обучения, сопровождения и  
оценке качества машинного перевода



# МАШИННЫЙ ПЕРЕВОД В РОССИИ И МИРЕ

Растет кол-во компаний, использующих технологии МП в своей деятельности



**СИБУР**



**SpanishDict**

**TADVISER**  
Государство. Бизнес. ИТ



Революционный скачок в качестве перевода в связи с разработками в области нейронных сетей и накопления качественных данных для обучения и тренировки систем



Бурный рост количества запросов на формирование глоссариев, тематических словарей и иных задач по структурированию лингвистических данных



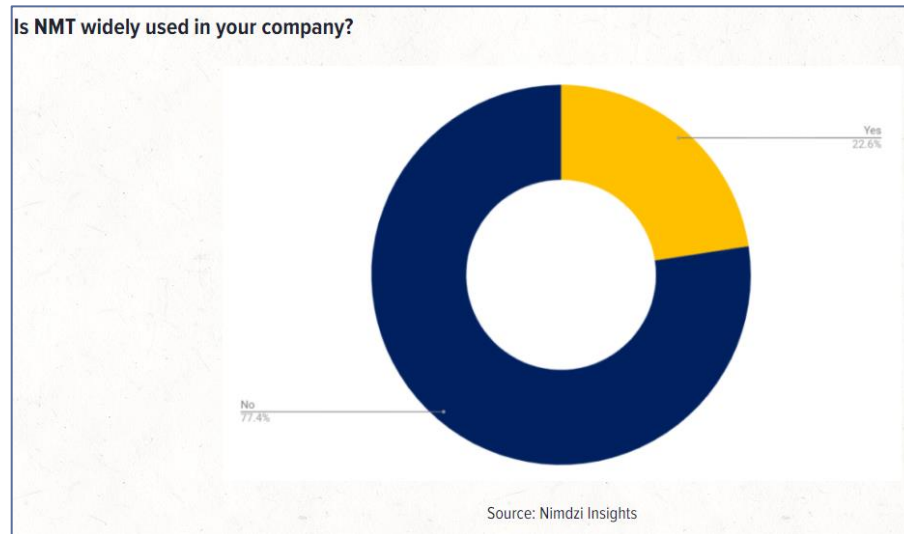
Всплеск интереса к метрикам качества оценки МП

## Основные сценарии применения машинного перевода:

- ✓ Международное сотрудничество;
- ✓ Исследовательская деятельность;
- ✓ Локализация описания товаров и услуг, сайтов и сервисов, мобильных приложений и т.д.;
- ✓ Обеспечение работы служб поддержки и взаимодействия с клиентом;
- ✓ Взаимодействие пользователей в онлайн-пространстве;
- ✓ **Интеграция в работу бюро переводов;**
- ✓ **Интеграция в работу отделов перевода компаний, служб корпоративных коммуникаций, служб информационного обеспечения.**

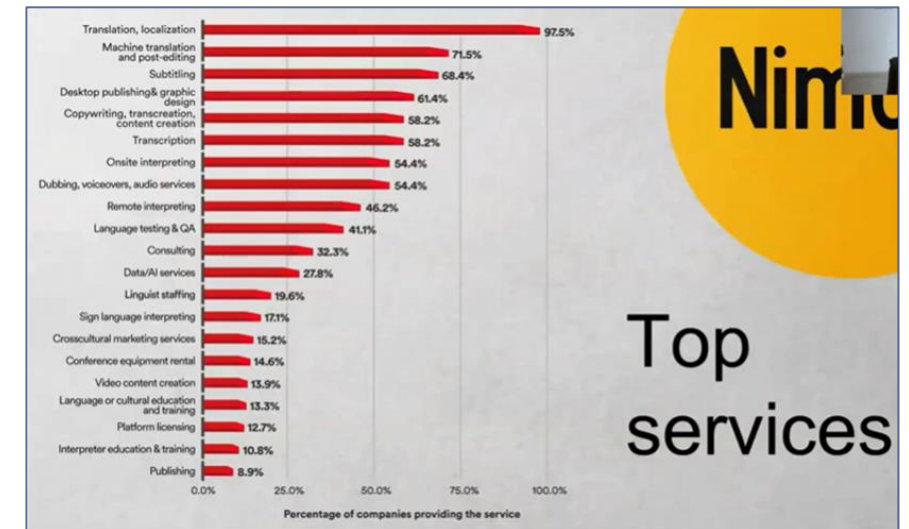
# МАШИННЫЙ ПЕРЕВОД – БЕЗОГОВОРОЧНЫЙ МИРОВОЙ ТРЕНД ИНДУСТРИИ ПЕРЕВОДОВ

¼ компаний, располагающих собственными отделами по переводу, используют МП



Nimdzi, 2020 г.

71,5% переводческих компаний предлагают клиентам машинный перевод



Nimdzi, 2021 г.

Машинный перевод и постредактирование – несомненные тренды международного рынка переводческих компаний.

¼ профзаказчиков используют МП в своей повседневной работе.

**Следующий этап эволюции** – кастомизированный машинный перевод с учетом специфики контента заказчика.

## Цели, стоящие перед отраслью перевода:

- ✓ **Ускорить** сроки подготовки проекта
- ✓ **Сократить** производственные издержки
- ✓ **Повысить** конкурентоспособность

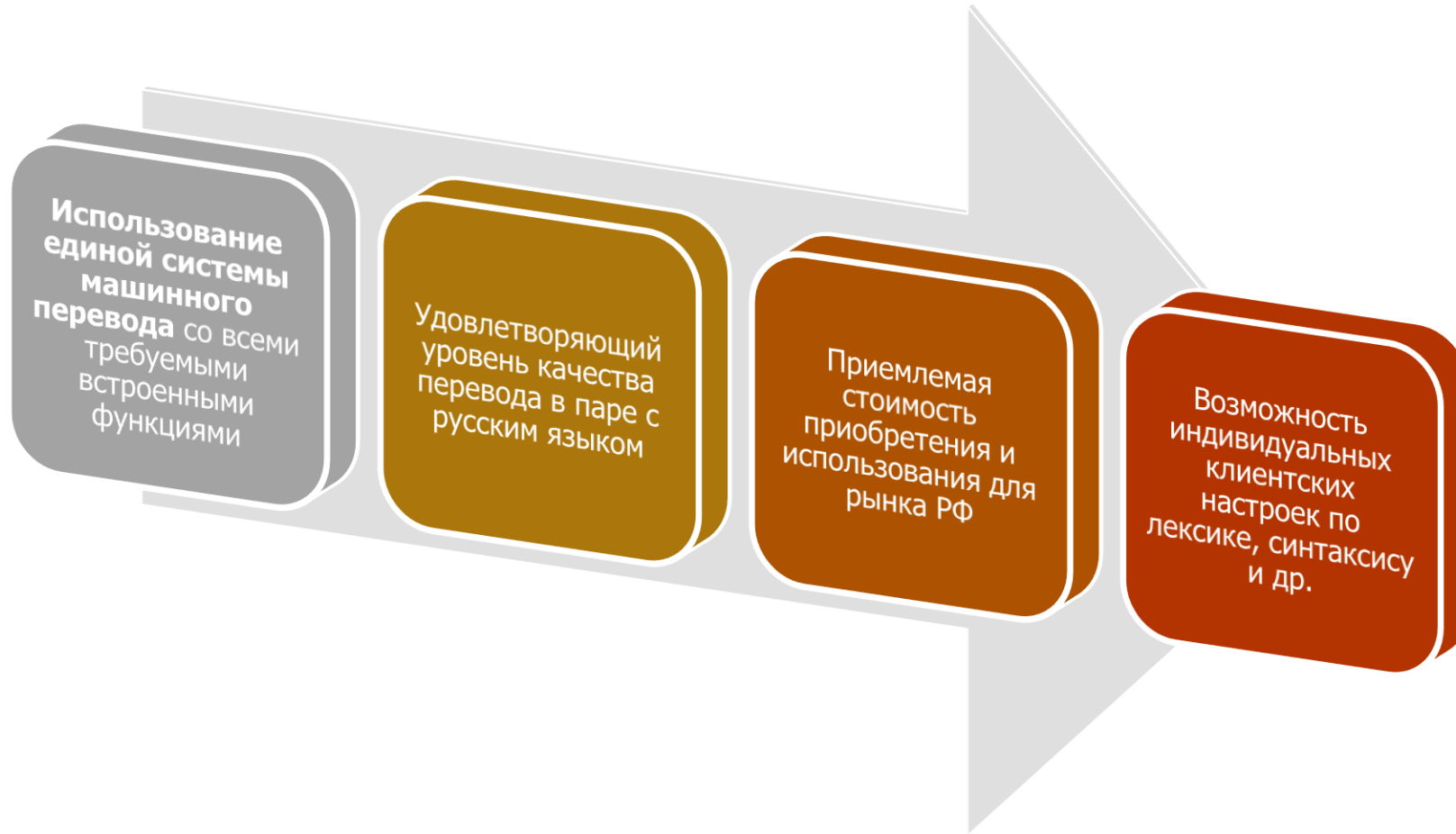
## Условия для реализации целей:

- ✓ Наличие объемных и продолжительных проектов с тематически однородными текстами
- ✓ Готовность обучать команду постредактированию
- ✓ Инфраструктура для внедрения МП
  - ✓ Для применения МП онлайн
  - ✓ Для развёртывания сервера или установки ПО на локальные компьютеры





# КЛЮЧЕВЫЕ ТРЕБОВАНИЯ К ХАРАКТЕРИСТИКАМ СИСТЕМ МП



**Решение:**  
симбиоз качественной технологии и индивидуальных клиентских данных и настроек  
в единой СМП  
по приемлемой для рынка РФ стоимости!

# ПРЕИМУЩЕСТВА ИСПОЛЬЗОВАНИЯ МАШИННОГО ПЕРЕВОДА И ПОСТРЕДАКЦИИ (МТ+ МТРЕ)



2019 г.



2021 г.




\*по сравнению с показателями до 2018 года

# PROMT И ЭГО ТРАНСЛЕЙТИНГ СМОТРЯТ В БУДУЩЕЕ ОТРАСЛИ ПЕРЕВОДА



Взаимная заинтересованность в синергии технологий и специалистов



Наши компании на практике осознали необходимость совмещать продуманную работу с данными и грамотную настройку МТ



У нас есть компетенция и технические возможности реализации подобного проекта



# КАСТОМИЗИРОВАННАЯ СИСТЕМА МАШИННОГО ПЕРЕВОДА EGO-PROMT



## 3 фактора успеха при внедрении МП

### **I** **Технология**

Передовые достижения современной науки (лингвистики) позволяют получить потрясающий результат машинного перевода

### **II** **Продукт**

Гарантия конфиденциальности переводимых данных, гибкая настройка под переводческие бизнес-задачи

### **III** **Данные**

Возможность настройки для перевода терминологии и узкоспециализированных текстов для полноценного отражения специфики клиента

## Обучение на параллельных данных

Настройка системы машинного перевода на перевод терминологии и стилистических клише, уникальных для предметной области заказчика, на параллельных данных.

Параллельные данные

оригинальные тексты и их переводы, выровненные по предложению;

переводы, выполненные профессиональными переводчиками;

на практике могут содержать значительное кол-во ошибок технического и лингвистического характера

## Обучение СМП PROMT

Специализированная модель перевода

Модель подключается в PROMT Neural Translation Server или десктопное решение PROMT Expert через профиль перевода. Профиль перевода доступен для перевода текста, документа, сайта в CAT-инструменте (Memsources, SDL Trados)

# ПРЕИМУЩЕСТВА КАСТОМИЗИРОВАННОЙ СМП EGO-PROMT



Удовлетворяет потребности компаний на том уровне, в той области и той языковой паре, которая необходима



Обучена наиболее точному терминологическому подбору



Учитывает предпочтения по стилю и синтаксису



Гарантирует корректное использование профессиональной терминологии

- Сбор и подготовка лингвистических данных, формирование баз переводов
- Индивидуальная настройка по вашим данным и параметрам
- Нелимитированный объем обрабатываемого перевода
- Минимизация количества ошибок без постредактирования
- Экономия до 50% времени на перевод

# ПРЕИМУЩЕСТВА КАСТОМИЗИРОВАННОЙ СМП EGO-PROMT

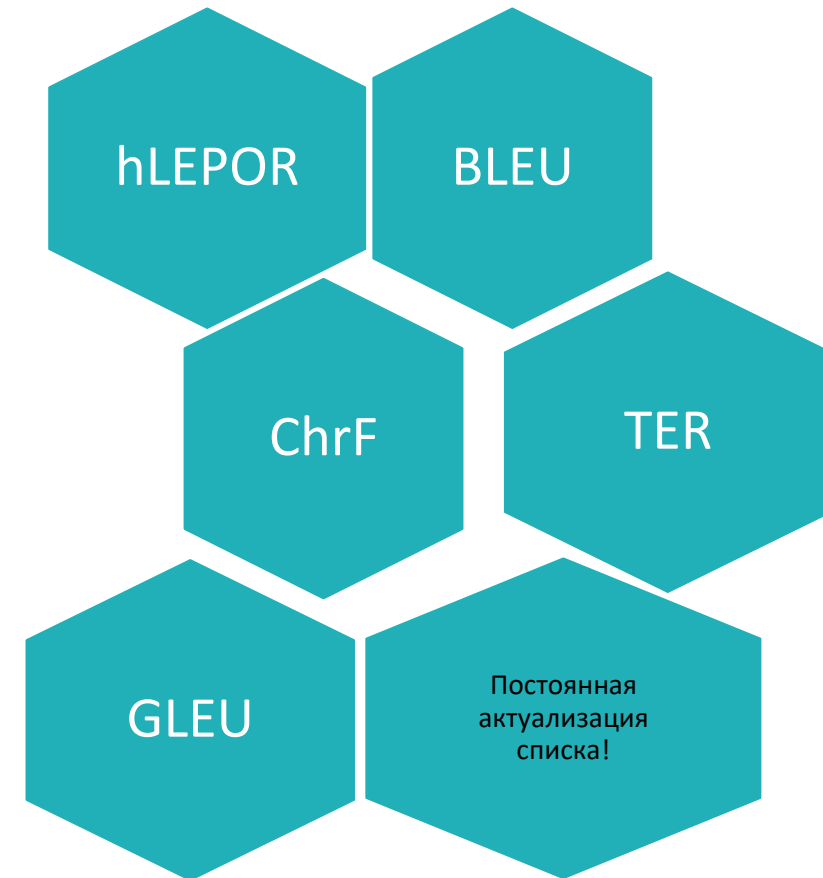
## Переводческие компании

- 01 Сжатые строки:**  
выполнение срочных заказов теперь возможно
- 02 Дополнительная прибыль и маржа:**  
снижается количество упущенных заказов из-за дисбаланса в цене и сроках
- 03 Технологичность:**  
ваш голубой океан внутри алого
- 04 Возможность значительного повышения объемов и высвобождения времени под другие задачи**
- 05 Профильная специализация:**  
повышение качества, объемов и прибыли от профильного направления

## Компании с регулярной потребностью в переводах (средний и большой объем)

- 01 Независимость от сторонних компаний:**  
ежедневной перевод документации любого объема «здесь и сейчас»
- 02 Экономия: финансов, времени и человеческих ресурсов**
- 03 Максимальная автоматизация и минимизация ошибок**
- 04 Конфиденциальность корпоративной информации**
- 05 Расширение:**  
настройка системы под себя с высоким качеством и минимальным привлечением человеческих ресурсов, возможность других языков
- 06 Выход из зоны внутреннего применения**

- ✓ Большой спектр метрик: учитывают разные факторы, применимы для разных задач и выводов.
- ✓ Каждая отдельная метрика решает конкретную задачу и отвечает на конкретный вопрос, следовательно нужно рассматривать результаты нескольких метрик.
- ✓ Результаты метрик говорят о динамике и дают первичную оценку, уменьшая «воронку» анализируемых параметров.
- ✓ Следует проводить комплексную оценку качества перевода СМП, включая экспертную оценку для финализации выводов.





## Показатели по метрикам на примере специализированной модели «Нефть и газ»:

Метрика	Показатели по стоковой модели	Показатели кастомизированной модели
BLEU	0,39	0,46
ChrF	0,37	0,47

Дополнительная информация по использованию автоматических метрик качества при анализе машинного перевода на [www.egotech.tech](http://www.egotech.tech)

**Общая тенденция прироста высокая  $\approx$  10%!**

### Результаты экспертного тестирования:

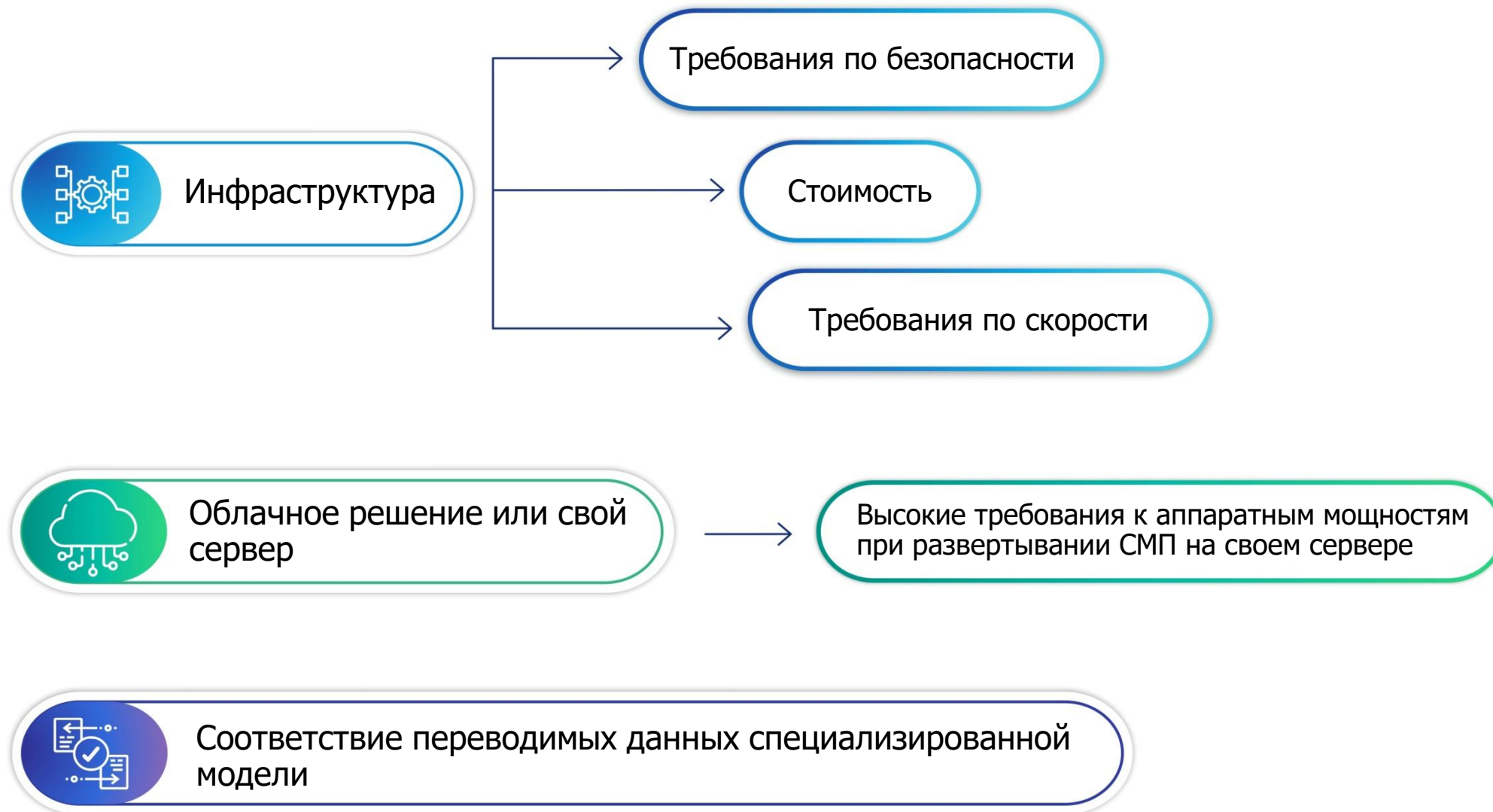
- С точки зрения комплексной оценки качества перевода, модель «Нефть и газ» показала большую результативность и эффективность при переводе отраслевой терминологии, в предметной области.
- Уменьшилось количество ошибок словоупотребления.
- Уменьшилось число встречаемых ошибок грамматического несоответствия.
- Общее количество ошибок, стандартных для машинного перевода, снизилось  $\approx$  на 20%



# EGO-PROMT

## ПРИМЕРЫ РАСЧЕТА СТОИМОСТИ

# ОГРАНИЧЕНИЯ ПО ВНЕДРЕНИЮ СМП EGO-PROMT



## Нефть и газ


Качественно и быстро переведет техническую, проектную, научно-исследовательскую документацию, нормативы, инструкции и руководства к оборудованию и ПО, результаты испытаний и другую информацию в отрасли.

## Атомная энергетика

Атомная энергетика - одна из самых наукоемких отраслей энергетики. Переводы в этой отрасли должны быть максимально точными и грамотными. Специализированная модель - незаменимый помощник для технических специалистов и профессиональных переводчиков на предприятиях атомной отрасли.

## Языковые направления:

английский → русский, русский → английский

-  700 000+ сегментов в паре английский-русский и 300 000 сегментов в паре русский-английский, что гарантирует качественную настройку модуля «Нефть и газ»
-  800 000+ сегментов в паре английский-русский и 1 015 804 сегментов в паре русский-английский обеспечили качественную настройку модуля «Атомная энергетика»
-  Минимальные сроки внедрения при максимально быстром старте
-  Высокие показатели качества по метрикам качества и по результатам человеческой оценки

# МОДЕЛЬ «ИНДИВИДУАЛЬНАЯ»: КАСТОМИЗАЦИЯ НА ДАННЫХ ЗАКАЗЧИКА



## РАЗРАБОТКА

Разработка под ключ с учетом специфики вашей компании

## УНИКАЛЬНОСТЬ

Уникальность ваших данных обеспечивает уникальные высококачественные переводы

## КОНФИДЕНЦИАЛЬНОСТЬ

Конфиденциальность данных всегда под защитой

## УРОВЕНЬ

Исключительный уровень адаптации под ваши потребности (типы и виды документов)



# СМП EGO-PROMT: УНИВЕРСАЛЬНОСТЬ, КАСТОМНОСТЬ, БЫСТРЫЙ СТАРТ



**... и Вы / Ваши заказчики получаете кастомизированную систему машинного перевода «под ключ» для быстрого качественного перевода нелимитированных объемов**

# РАСЧЁТ ГОТОВОЙ СПЕЦИАЛИЗИРОВАННОЙ МОДЕЛИ СМП EGO-PROMT

Серверное решение / Windows / до 10 пользователей (5 одновременно)			Приоритетная техническая поддержка PROMT
PROMT Neural Translation Server	Специализированная модель перевода ( <u>нефть и газ</u> )	Плагин для Trados Studio*	20% от стоимости лицензии
140 000 руб.	40 000 руб.	9 000 руб.	

\* Для использования Memsource ничего приобретать требуется;

\*\* Дополнительно приобретается модуль для распознавания графических файлов – 40 000 рублей (для работы через CAT инструменты она не требуется).



**Специальные условия приобретения  
при размещении запроса  
до 31.03.2022**

# РАСЧЁТ КАСТОМИЗИРОВАННОЙ КОМПЛЕКТАЦИИ СМП EGO-PROMT

Серверное решение / Windows / до 10 пользователей (5 одновременно)				Приоритетная техническая поддержка PROMT
PROMT Neural Translation Server	Кастомизированная модель перевода (на материалах заказчика)		Плагин для Trados Studio*	20% от стоимости лицензии
	Техническая очистка ПБД, руб./10 000 сегментов (≈100-110 тыс. слов)	Полная очистка ПБД при наличии словаря, руб./10 000 сегментов (≈100- 110 тыс. слов)		
140 000 руб.	5 000 руб.	40 000 руб.	9 000 руб.	

\* Для использования Memsource ничего приобретать требуется;

\*\* Дополнительно приобретается модуль для распознавания графических файлов – 40 000 рублей (для работы через CAT инструменты она не требуется).



**Специальные условия приобретения  
при размещении запроса  
до 31.03.2022**



# Спасибо за внимание!

Информация об СМП EGO-PROMT  
Запрос демо-доступа  
Сбор и подготовка баз перевода  
Обучение МТ+МТРЕ

E-mail: [Julia.Epiphantseva@prompt.ru](mailto:Julia.Epiphantseva@prompt.ru)  
[corporate@prompt.ru](mailto:corporate@prompt.ru)

Телефон: +7 495 580 45 45