



НОВЫЕ ЛИНГВИСТИЧЕСКИЕ
ТЕХНОЛОГИИ

PROMT Translation Factory

Уникальная платформа для
повышения эффективности
переводческих бизнес-процессов:
новые возможности для роста

Юлия Епифанцева
Директор по развитию бизнеса

Современные ВЫЗОВЫ И ТРЕНДЫ

- Миграция с OS Windows на свободное программное обеспечение на базе ОС Linux и использование ПО из Единого реестра отечественного ПО
- Сложности с приобретением, продлением лицензий на системы автоматизированного перевода (CAT tools) от зарубежных производителей или ограничения по их использованию по соображениям конфиденциальности
- На российском рынке нет ни одной CAT-системы
 - ✓ созданной российскими разработчиками,
 - ✓ работающей под управлением российских ОС на базе Linux,
 - ✓ работающей на локальном сервере клиента,
 - ✓ зарегистрированной в Едином реестре отечественного ПО

Ситуация на рынке

Trados, memoQ

О решении

- Разработчик - иностранная компания
- Решение работают офлайн, все данные хранятся на локальной машине или локальном сервере
- Работает на ОС Windows

Минусы

- Разработчик приостановил весь свой бизнес (поставка услуг и технологий) в России в 2022 году
- Нет возможности приобрести ПО и активировать лицензии
- Пользователи ранее приобретенных версии больше не могут:
 - обновить существующую версию и перейти на новые версии;
 - обратиться в службу поддержки.
- Решение не входит в Единый реестр отечественного ПО
- ОС Linux не поддерживается



Ситуация на рынке

Phrase (Memsource), Smartcat

О решении

- Разработчик - иностранная компания
- Решение работает онлайн

Минусы

- Все данные хранятся в облаке, дата-центры не в РФ
- Решение не входит в Единый реестр отечественного ПО



Как все начиналось

Компетенция PROMT

Общая компетенция

- Опыт в разработке лингвистического ПО более 30 лет
- Команда разработчиков в штате компании
- Знание и технологии работы с основными форматами документов (MS Office, Open Document, PDF, RTF и др.)
- Знание и технологии работы с большими объемами текстовой информации (поиск, анализ, сегментация, фильтрация, загрузка/выгрузка, редактирование и т.д.)
- Лингвистические компетенции, которые имеют ключевое значение для разработки и тестирования лингвистических алгоритмов продукта (fuzzy match, работа с конструкциями в разных языках)

Алгоритмы, БД

- Нейронные модели
- OCR модуль
- Собственная СУБД
- Алгоритмы работы с форматами

Опыт в разработке CAT

- Общее понимание устройства и принципов работы CAT систем
- Опыт разработки простой десктопной CAT системы
- Опыт разработки элементов CAT системы в переводчиках PROMT (работа с ТМ, с глоссариями, разработка плагинов Trados и Memsource/Phrase)
- Знание и технологии работы с основными форматами передачи данных в CAT системах (TMX, TBX, XLIFF)

Можно начинать 😊

ПОДГОТОВИТЕЛЬНЫЙ ЭТАП (апрель – сентябрь 2022)



- ✓ Подробное изучение зарубежных CAT систем
- ✓ Опросы, анкетирование, интервью с клиентами
- ✓ Составление ТЗ
- ✓ Планирование (команда, сроки, бюджет)

ПЕРВЫЙ ЭТАП

(октябрь - декабрь 2022)

- ✓ Реализация базового алгоритма Fuzzy Match
- ✓ Реализация полноценного макета (загрузка данных, поиск в ТМ с учетом FM, уточнение алгоритма FM, работа с конструкциями, разработка редактора, работа с форматными признаками)
- ✓ 4 разработчика + 4 лингвиста
- ✓ Опытная эксплуатация в компании



ВТОРОЙ ЭТАП (январь – май 2023)



- ✓ Разработка решения с основными функциями CAT (работа с форматами, машинный перевод, оптимизация поиска по ТМ)
- ✓ Подготовка дизайна (figma)
- ✓ 5 месяцев, 7 разработчиков (full/part-time), 3 лингвиста (part-time), 4 тестера (part-time), привлеченный эксперт по CAT системам (part-time)

ТРЕТИЙ ЭТАП

(июнь - июль 2023)

- ✓ Проведено 13 демонстраций клиентам и предоставлен демо-стенда или набор для локальной установки для самостоятельного тестирования
- ✓ Сбор фидбэка, систематизация, исследования, новые ТЗ
- ✓ Осознание необходимости архитектурных изменений

