ТОО «Т3 Solutions»

Техническое задание

Распознавание речи на казахском и русском языках в условиях промышленного шума для голосового ввода в мобильных проверках HSE

1. Общие сведения

ТОО «Т3 Solutions» развивает платформу HSSE и мобильные инструменты для проверок охраны труда и промышленной безопасности на объектах.

Инспекторам и мастерам на площадке неудобно заполнять формы вручную: перчатки, СИЗ, погодные условия. Нужен голосовой ввод, который работает в шуме и понимает смешанную казахско-русскую речь и отраслевую терминологию.

2. Цель

Получить модуль распознавания речи (speech-to-text) для мобильного приложения, устойчивый к промышленному шуму, с поддержкой казахского и русского языков и HSE-терминологии.

3. Объем работ

- Выбор и дообучение модели распознавания речи (например, на базе Whisper или аналогов) под шумные условия и смешанную речь.
- Словарь отраслевой терминологии: виды работ, СИЗ, оборудование, типы нарушений.
- Шумоподавление и обработка речи на устройстве или на сервере в РК.
- Преобразование распознанного текста в структурированные поля формы проверки (тип нарушения, место, описание, ответственный).
- SDK или API для встраивания в мобильное приложение (Android).

4. Требования к исполнителю

- Опыт работы с моделями распознавания речи, особенно для казахского языка.
- Опыт подготовки аудиоданных и дообучения моделей.
- Готовность подписать NDA и обрабатывать данные на территории РК.

5. Ожидаемый результат

Голосовой ввод в мобильных проверках с приемлемой точностью на площадке; сокращение времени на оформление нарушений; больше зафиксированных замечаний за смену.

6. Порядок взаимодействия

Пилот на ограниченном наборе сценариев и записей, оценка точности (WER), далее развитие. Условия и стоимость обсуждаются по итогам пилота.

7. Пилот (MVP)

- Объект: мобильное приложение проверок HSE T3 Solutions (Android), тестирование на площадке пилотного заказчика.
- Объем: 50 типовых сценариев фиксации нарушений; словарь 300 HSE-терминов; казахский, русский и смешанная речь; шум до 85 дБ.
- Сроки: 10 недель.
- Критерии успеха: WER не более 15% на тестовом наборе записей в шуме; распознанный текст правильно раскладывается по полям формы не менее чем в 80% случаев; время ответа не более 2 секунд.

8. Источники данных и доступы

- Аудиоданные: не менее 10 часов записей в условиях промышленного шума и словарь HSE-терминов готовит и передает T3 Solutions; при необходимости открытые корпуса казахской речи (например, KSC2).
- Интеграция: мобильное приложение проверок HSE (Android) через SDK или REST API; развертывание на сервере в РК или на устройстве. Тестовая сборка и описание API предоставляются T3 Solutions.