

Модуль речевого ввода данных с LLM-корректором

Информация о программе

Работа на здоровье

INTERIN
ТЕХНОЛОГИИ

2026 г.

Модуль речевого ввода данных с LLM-корректором

Информация о программе

Документ разработан ООО «Интерин технологии» (©).

Все права защищены. Никакая часть настоящего документа не может быть воспроизведена или передана в какой бы то ни было форме и какими бы то ни было средствами, будь то электронные или механические, включая фотокопирование, запись на магнитный носитель, электронную почту и публикации в Интернет, если на то нет письменного разрешения автора.

Контактная информация

Группа компаний «Интерин»

Web: www.interin.ru

E-mail: info@interin.ru

Тел: +7 (495) 2208235

МОДУЛЬ РЕЧЕВОГО ВВОДА ДАННЫХ С LLM-КОРРЕКТОРОМ. О ПРОГРАММЕ

Полное наименование: Модуль речевого ввода данных с LLM-корректором.

Краткое наименование: Модуль.

Разработчик и правообладатель: ООО «Интерин технологии».

Модуль представляет собой программу для транскрибации речевого потока в текст с последующей коррекцией ошибок транскрибации с помощью больших языковых моделей (LLM). На входе программа получает речевой поток, на выходе выдает транскрибированный и исправленный текст в формате медицинского документа.

Эксплуатационным назначением модуля является применение в здравоохранении для формирования специализированным программным обеспечением медицинских документов заданной структуры при голосовом вводе данных.

Одним из барьеров, препятствующих широкому распространению голосового ввода медицинских данных в медицинских информационных системах (МИС), являются недостаточные потребительские качества текстов, получающихся после транскрибации. Не все медицинские термины и слова общего лексикона распознаются корректно, нарушается согласование слов по роду, числу и падежам, текст недостаточно хорошо форматирован с точки зрения грамматики. Всё это требует дальнейшей доработки текста.

Ещё одной проблемой видится необходимость приведения текста к структуре медицинского документа в МИС. Структура документа может быть достаточно сложной, содержать много элементов, иметь требования к типу и формату данных в элементах структуры.

Но речевой ввод может лишь частично использоваться для формирования документа, а недостающие данные могут быть взяты из пользовательского шаблона. Для решения указанных проблем следует использовать LLM в качестве корректора результатов транскрибации речи, интегратора речевых и текстовых данных из шаблона и формирователя структуры результирующих данных.

1) **Коррекция результатов транскрибации.**

Из-за ошибок распознавания медицинских терминов и слов общего лексикона, нарушений грамматического согласования и слабого форматирования текст после транскрибации требует доработки. ПО должно автоматически исправлять эти недостатки с использованием большой языковой модели (LLM). Конкретные задачи коррекции:

- точное распознавание и написание медицинских терминов;
- грамматическая проверка и исправление ошибок согласования (род, число, падеж);
- приведение текста к единому стилю и улучшение его форматирования.

2) **Структурирование данных.**

ПО должно преобразовывать свободный текст, полученный после коррекции, в строго заданную структуру медицинского документа МИС. Структура может включать множество элементов с жёсткими требованиями к типу данных (текст, число, дата, код) и их формату.

3) Интеграция данных из разных источников.

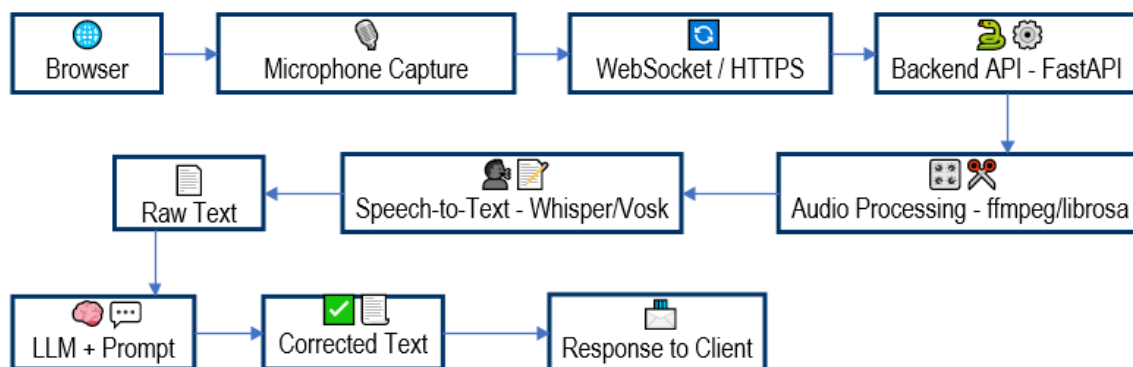
Поскольку в ходе речевого ввода информация может быть представлена не полностью, ПО должно дополнять транскрибированные данные недостающей информацией из заранее подготовленного пользовательского шаблона. LLM должна выступать в роли интегратора, объединяя речевые и шаблонные данные в единый структурированный документ.

Шаблон структуры документа. Как правило, медицинские документы имеют структуру. Эта структура видна в печатных копиях документов, структура отражается в объектах хранения и представления документов в базе данных МИС. Обычно структура представляет собой дерево. Информационной моделью структуры документа может служить xml или Json объект. Обратное отображение этой структуры в дерево тривиально при условии, что все узлы структуры имеют уникальные имена. Именно такие шаблоны структуры документа должны использоваться при реализации.

Шаблон заполнения документа. Такие шаблоны используются врачами повсеместно. Шаблон заполнения может быть совмещён со структурой документа – атрибуты узлов xml или пары ключ – значение Json. Основная сложность применения шаблонов заполнения как дополнительного к речевому источнику данных для LLM заключается в сложности «умного» синтеза данных из двух источников. Врачи при работе с данными из шаблона заполнения могут действовать по-разному с каждым из элементов данных: 1) оставлять неизменным (insert); 2) удалять данные заполнения (delete); объединять данные заполнения с дополнительными данными вводимыми пользователями (append); могут отредактировать данные заполнения (edit).

Модуль речевого ввода должен обрабатывать полученный текст и шаблон, осуществляя «умный» синтез данных из речевого ввода с данными из шаблона заполнения.

Архитектура Модуля:



Пример Prompt для LLM при использовании Модуля:

1. Запрос (инструкции, требования)
2. Шаблон диагностического протокола
3. Запрос (требование)

<p>В ходе транскрибации голосового ввода УЗИ почек получен текст.</p> <p>Исправь в этом тексте ошибки транскрибации.</p> <p>Размеры почек указывай как на примере: "11,1x4,8x4,9 см".</p> <p>Обрати внимание и исправь в протоколе возможные ошибки в словах: "паренхимы", "гиперэхогенные", "дилатирована", "нефроптоз".</p> <p>Сформируй протокол УЗИ почек строго по указанному шаблону в случае, если значение поля шаблона отсутствует в тексте обязательно выводи это поле в протоколе, но без указания значения.</p>	<p>Шаблон протокола:</p> <p>"ОСМОТР ПОЧЕК:</p> <p>Правая почка (состояние и размеры):</p> <p>Контуры:</p> <p>Толщина паренхимы:</p> <p>Структура паренхимы почки:</p> <p>В почечном синусе:</p> <p>Чашечно-лоханочная система:</p> <p>Левая почка (состояние и размеры):</p> <p>Контуры:</p> <p>Толщина паренхимы:</p> <p>Структура паренхимы почки:</p> <p>В почечном синусе:</p> <p>Чашечно-лоханочная система:</p> <p>Проекция надпочечников:</p> <p>УЗ-ангиография:</p> <p>Заключение:</p> <p>Рекомендации:</p> <p>Патологии:</p> <p>Диагноз: "</p>	<p>Строго придерживайся информации, указанной в тексте, не добавляй лишней информации, которой в тексте нет.</p>
---	---	--

ОПИСАНИЕ ФУНКЦИОНАЛЬНЫХ ХАРАКТЕРИСТИК

Программа выполняет функции:

- отправка получаемого с микрофона аудиопотока в сервис транскрибации (ASR) и возвращение полученного от ASR текста;
- отправка полученного от ASR текста и шаблона ответа (медицинского документа) на обработку в LLM, возвращение обработанного текста в запрошенном формате медицинского документа;
- задание шаблонов для формирования документов;
- настройка программы на выбор сервисов (могут быть использованы как облачные, так и локальные сервисы транскрибации и LLM).

Процесс строится следующим образом:

Шаг	Вход	Действие	Выход
1	Микрофон	Захват аудио	PCM / Opus чанки
2	Аудио чанки	WebSocket отправка	Бинарные сообщения на сервер
3	Сжатое / raw аудио	Декодирование (ffmpeg)	Моно 16 кГц int16
4	Очищенное аудио	STT (Whisper/Vosk)	Текст с ошибками
5	Текст с ошибками + промпт	LLM	Исправленный текст
6	Исправленный текст	LLM	Структурированный документ
7	Структурированный документ	WebSocket отправка клиенту	Показ пользователю

УСТАНОВКА, ВВОД В ЭКСПЛУАТАЦИЮ И ЭКСПЛУАТАЦИЯ МОДУЛЯ РЕЧЕВОГО ВВОДА ДАННЫХ С LLM-КОРРЕКТОРОМ

Порядок действий

Установка и ввод в эксплуатацию Модуля речевого ввода данных с LLM-корректором в медицинской организации включает следующие этапы:

- 1) Поставка права использования программного обеспечения на основе лицензионного соглашения.
- 2) Предоставление дистрибутива Программы с инструкцией для самостоятельной установки (опционально, при отсутствии договора на внедрение).
- 3) Установка программного обеспечения.
- 4) Организация взаимодействия между Модулем речевого ввода данных с LLM-корректором и МИС МО.
- 5) Обучение персонала МО работе с Модулем.
- 6) Ввод программы Модуля речевого ввода данных с LLM-корректором в эксплуатацию в составе МИС МО.
- 7) Поддержка пользователей и администратора программы на начальном этапе.

Администрирование Модуля осуществляет администратор МИС МО. Порядок работы с Модулем описан в документе «Модуль речевого ввода данных с LLM-корректором. Руководство пользователя».

Необходимая инфраструктура

Модуль использует следующие сторонние компоненты (не ограничивающие по условиям лицензирования возможности использования ПО):

Библиотека	Версия ¹	Лицензия	Примечания
Aiofiles	25.1.0	MIT	Стандартная разрешительная лицензия
Aiohttp	3.13.3	Apache 2.0	Разрешительная, с явным указанием патентных прав
aiohttp-cors	0.8.1	MIT	Совместима с лицензией aiohttp
Av (PyAV)	16.1.0	BSD 2-Clause	Разрешительная
boto3	1.42.57	Apache 2.0	Стандартная для AWS SDK
ctranslate2	4.7.1	Apache 2.0	Совместима с другими компонентами ML
faster-whisper	1.2.1	MIT	Основан на Whisper от OpenAI (MIT)
Huggingface_hub	1.4.1	Apache 2.0	

¹ Или последние на момент выпуска продукта стабильные версии

Библиотека	Версия ¹	Лицензия	Примечания
Httpx	0.28.1	BSD 3-Clause	Разрешительная
langchain-classic	1.0.1	MIT	Разрешительная
langchain-community	0.4.1	MIT	Аналогично langchain-classic
langchain-core	1.2.16	MIT	
langchain-text-splitters	1.1.1	MIT	
Langsmith (SDK-библиотека для Python)	0.7.7	MIT	Свободна для использования
Numba	0.64.0	BSD 2-Clause	
Onnxruntime	1.24.2	MIT	Разрешительная
openai-whisper	20231117	MIT	
Orjson	3.11.7	Apache 2.0	Разрешительная лицензия
Pydantic	2.12.5	MIT	
pydantic-settings	2.13.1	MIT	
Pydub	0.25.1	MIT	Разрешительная
python-dotenv	1.2.1	BSD 3-Clause	
PyYAML	6.0.3	MIT	
Requests	2.32.5	Apache 2.0	
Rich	14.3.3	MIT	
Scipy	1.17.0	BSD 3-Clause	
Sounddevice	0.5.5	MIT	
SQLAlchemy	2.0.47	MIT	
Srt	3.5.3	MIT	Разрешительная
Tenacity	9.1.4	Apache 2.0	
Tiktoken	0.12.0	MIT	
Tokenizers	0.22.2	Apache 2.0	
Torch (PyTorch)	2.10.0	BSD 3-Clause	Разрешительная
Tqdm	4.67.3	MIT	
Typeer	0.24.1	MIT	
Vosk	0.3.45	Apache 2.0	Разрешительная
Websockets	16.0	BSD 3-Clause	
Модель Qwen	2.5-32B-Instruct (квантованная 4-8 бит)	Apache 2.0	Разрешительная
Модели Vosk	например, vosk-model-ru-0.22, vosk-model-small-ru-0.22 и другие	Apache 2.0	Разрешительная
Модели Whisper	например, whisper-small, whisper-large-v3 и другие	MIT	Разрешительная

Для функционирования серверной части Модуля речевого ввода необходимо следующее системное программное обеспечение:

- Системы на базе Debian: Debian 12 и выше Ubuntu 24 и выше.
- Для функционирования клиентской части Модуля речевого ввода на рабочей станции пользователя должен использоваться следующий набор системного программного обеспечения:

1) Веб-браузеры актуальных версий (любой из списка):

- Яндекс.Браузер;
- Спутник;
- Microsoft Edge;
- Google Chrome;
- Mozilla Firefox;
- Opera.

Модуль речевого ввода должен функционировать на технических средствах с характеристиками не ниже следующих.

Сервер

Характеристика	Значение
CPU	От Intel Xeon 2.0 ГГц 8 ядер
GPU	От RTX 4090 24 ГБ
RAM	От 32 ГБ (рекомендуется 64 ГБ для комфортной работы с Qwen2.5-32B)
Дисковое пространство	От 500 ГБ (RAID 10)

Рабочие станции

Параметры и характеристики классической рабочей станции:

Устройство	Характеристики
CPU	От 2 ГГц
RAM	От 4 ГБ
Дисковое пространство	От 80 ГБ
Видеокарта	Поддержка видеорежима 1280 x 1024 и выше
Монитор	От 19.0" LCD монитор

Принтер (опционально)

Сетевая инфраструктура

Клиенты взаимодействуют с серверами на скорости не ниже 100 Мбит/сек.

МОДУЛЬ РЕЧЕВОГО ВВОДА ДАННЫХ С LLM-КОРРЕКТОРОМ. СТОИМОСТЬ И ПОРЯДОК ПРИОБРЕТЕНИЯ

Стоимость ПО рассчитывается индивидуально по запросу заинтересованного лица, направленному письменно или по телефону.

Для приобретения программного обеспечения «Модуль речевого ввода данных с LLM-корректором» можно:

Оставить заявку на сайте: www.interin.ru

Написать письмо в свободной форме по e-mail: info@interin.ru

Позвонить по телефону: +7 495 220-82-35