

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

федеральное государственное автономное
образовательное учреждение высшего образования
«Самарский национальный исследовательский университет
имени академика С.П. Королева»
(Самарский университет)

Институт информатики и кибернетики
Кафедра технической кибернетики

ЛАБОРАТОРНАЯ РАБОТА № 2
Применение `p8n` для реализации пайплайна обработки видео с
использованием LLM моделей.

по курсу
Инженерия данных

Группа 6232

Студент _____ Н.В. Носов
(подпись)

Преподаватель _____ Р.А. Парингер
(подпись)

АРХИТЕКТУРА

Воркфлоу представляет собой набор нод в n8n. Он спроектирован для работы в контейнерной среде Docker, что обеспечивает изоляцию и воспроизводимость. На рисунке 1 представлена схема воркфлоу.



Рисунок 1 – Схема воркфлоу

Логическая схема процесса выглядит следующим образом:

- прием видео в виде ссылки или файла,
- извлечение аудиодорожки,
- генерация английских субтитров с помощью `auto_subtitle`,
- перевод их на русский язык с использованием LLM,
- вшив русских субтитров обратно в видео,
- отправка результат пользователю через telegram бота.

РЕЗУЛЬТАТЫ

На рисунке 2 представлен скриншот реализованного воркфлоу из n8n.

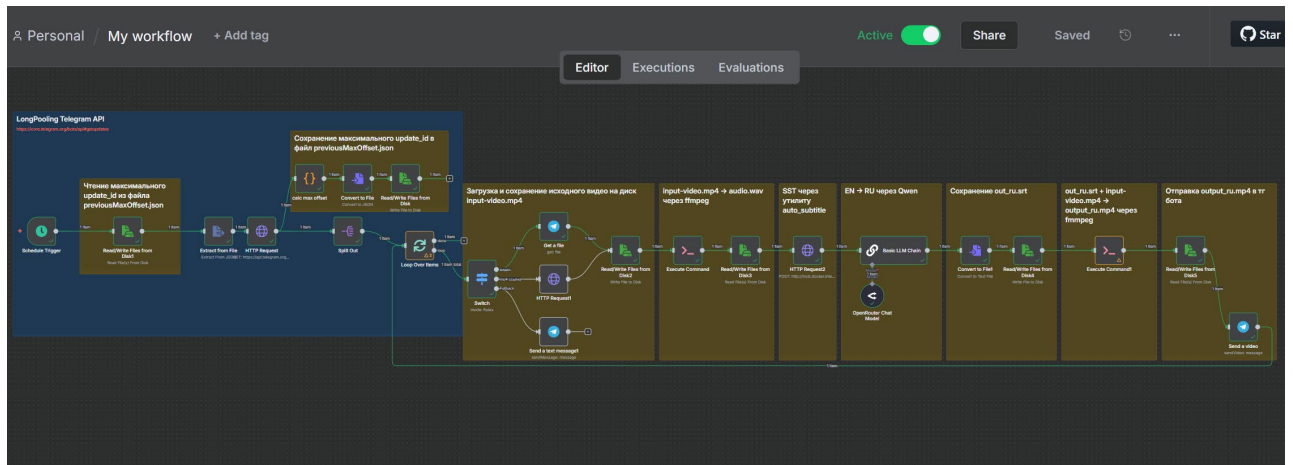


Рисунок 2 – Скриншот воркфлоу из n8n

На рисунке 3 представлен пример работы tg бота. На вход подаётся видео с речью на английском языке, в ответ пользователь получает это же видео, но с русскими субтитрами.

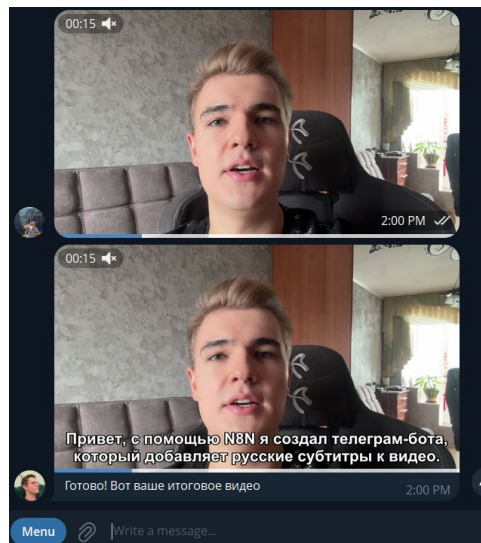


Рисунок 3 – Пример работы tg бота

На рисунке 4 представлены запущенные контейнеры, необходимые для корректной работы приложения.

<input type="checkbox"/>		auto-subtitle-service	Running (1/1)	0.03%	2 hours ago			
<input type="checkbox"/>		auto-subtitle-servi 31e9589f5f5d	auto-subtitle-service-au Running	0.03%	5000:5000			
<input type="checkbox"/>		n8n	Running (2/2)	0.26%	2 hours ago			
<input type="checkbox"/>		traefik-1 b17dc465d9db	traefik Running	0%	443:443			
<input type="checkbox"/>		n8n-1 a816496387fb	n8n-n8n Running	0.26%	5678:5678			

Рисунок 4 – Запущенные контейнеры

ВЫВОДЫ

Самым сложным при выполнении данной работы оказался этап SST. Сначала пытался добавить утилиту auto-subtitle (<https://github.com/m1guelpf/auto-subtitle>) сразу в Dockerfile n8n, чтобы использовать его через execute ноду, но данная утилита не работает на последних версиях питона. А образ n8n основан на alpinelinux, из репозитория которого уже удалены старые версии питона (https://pkgs.alpinelinux.org/packages?name=python3*&branch=edge), есть только 3.12, что нам не подходит. Поэтому было принято решение вынести утилиту auto-subtitle в отдельный docker контейнер с реализованным API.

Также из проблем оказалось то, что n8n, развёрнутый локально не даст просто так использовать встроенные вебхуки. Прием обновлений тг ботов можно было бы сделать одной нодой, но из-за того, что мой айпи-адрес не статический, сделать работающий извне локальной сети компьютера вебхук – непростая задача. Поэтому я решил сделать обработку сообщений не через вебхуки, а через LongPolling. Раз в 20 секунд я получаю все обновления по тг боту через API телеграмма, тем самым их обрабатываю.

Что касается ограничений, то тг не дает скачивать видео более 20 мб (примерно), поэтому на вход лучше подавать ролики, которые весят не выше данного порога. При скачивании видео через url такого ограничения нет.

В целом, работа оказалась полезной. Прокачал навыки работы с видео/аудио/субтитрами. Познакомился с новыми для себя утилитами ffmpeg и auto-subtitle. С n8n я уже был знаком, так как имею на нём коммерческий опыт на нескольких проектах.