

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

(подпись)

К.Р. Донец

Преподаватель

(подпись)

Р.А. Парингер

АРХИТЕКТУРА

Данная лабораторная работа требует обработки видео из ТГ бота по средству приема чистого файла или url на видео. Логика пайплайна из задания перенесена в n8n для обработки каждой задачи по нодам.



Рисунок 1 – Схема пайплайна из задания

Общая логика - запуск schedule trigger на опрос TG API с помощью get Updates(Не использовался TG триггер из-за особенности его работы, так как требуется отправка на https домен). После чего идет проверка на тип данных, если не видео и не URL - отправляется сообщение пользователю о некорректном формате. В других случаях - либо скачивание видео по ссылке, либо с помощью getFile от тг получаем файл по его ID. Далее идет обработчик с сервера развернутого в контейнере. Используются маршруты - /extract_audio - для вытаскивания аудиодорожки из файла, /generate_subtitles - для конвертации аудио в текст, /translate_subtitles - для перевода субтитров,

/burn_subtitles - и наконец наложения субтитров на видео. После чего идет отправка видео обратно пользователю.

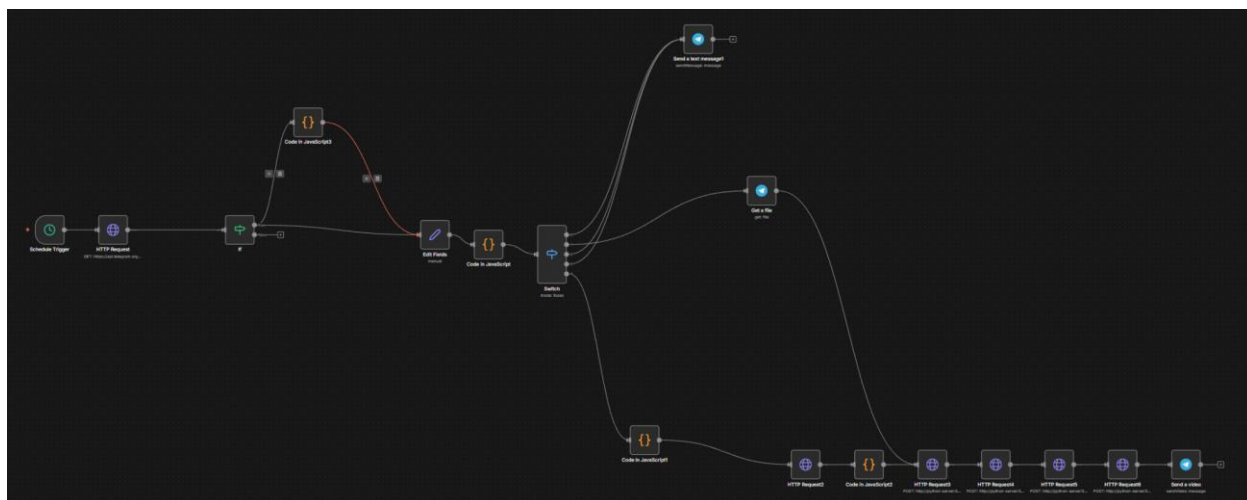


Рисунок 2 – Реализация workflow в n8n

Как исходник бралась песня Bring Me The Horizon - Throne, для теста отправлял выкаченное видео напрямую и через ссылку, которую получил путем размещения на гугл диске. Ниже представлен пример работы.

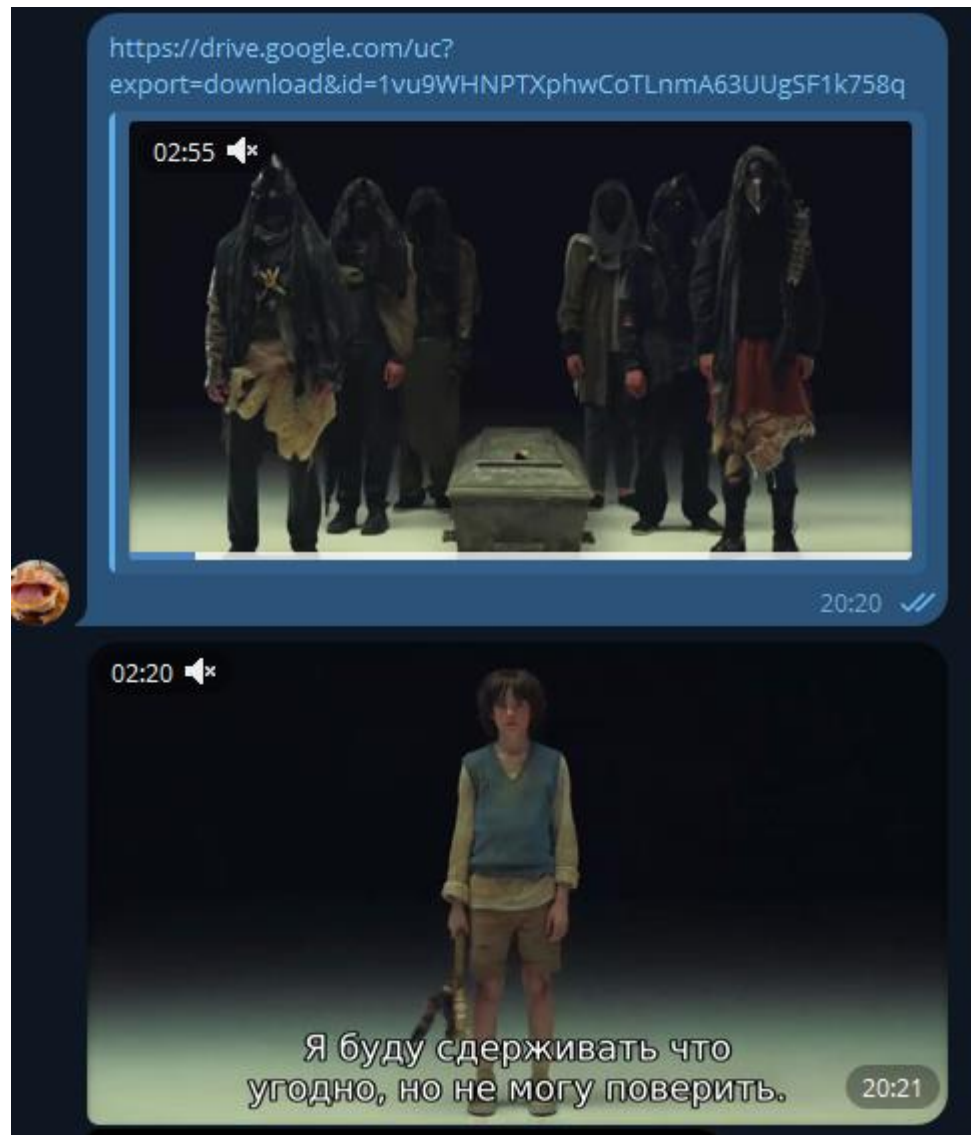


Рисунок 3 – пример переведенного видео

Теперь про сервер- использовалась предобученная модель Helsinki-NLP/opus-mt-en-ru для перевода, заранее скачанная и размещенная в докере. Так как напрямую качать не выходит, из-за ограничений сети:) Для извлечения субтитров из аудио использовалась whisper модель от OPEN AI.

ВЫВОДЫ

Самые сложные части работы - понять почему ничего не работает, хотя должно. Таких случаев было два - TG Trigger, и использование локальной модели для перевода. Как только были найдены способы это обойти - процесс явно стал идти быстрее. Данная лабораторная работа позволила познакомиться с n8n, что оказалось достаточно полезным, так как в рабочей задаче уже появился запрос о разворачивании n8n сервера для автоматизации работы компании. В целом - n8n удобная вещь, но нужно сильно больше опыта и сноровки для того чтобы писать микро сервисную архитектуру, еще и используя по максимуму возможности n8n - чего я явно не добился до конца, так как большинство возможностей пока что скрыты из-за неопытности. Задача была выполнена в соответствии с заданием, на вход получено видео(файл или url) на выход видео с переведенными субтитрами.