

Задание 3: Формирование протокола по аудио

Суть задачи

Разработать систему, которая по аудиозаписи автоматически формирует структурированный протокол заседания.

В реальных процессах (заседания, совещания, диссертационные советы) требуется фиксировать ход обсуждения и формировать протокол. Ручная расшифровка занимает значительное время.

Система должна:

- преобразовывать речь в текст;
- разбивать поток речи на реплики;
- разделять говорящих;
- формировать читаемый протокол;
- позволять пользователю задать роли участников.

Важно: система не должна определять личность говорящего, только различать участников.

Задача сводится к: **построению pipeline обработки аудио → текст → структура.**

Входные данные

Форматы:

- .wav, .mp3

Аудио может содержать:

- несколько участников;
- шум;
- пересечения речи;
- различное качество.

Требования к решению

- ASR + сегментация;
- speaker diarization;
- HTTP API:

POST /build_protocol

- автоматическая обработка.

Формат результата

```
{
  "transcript": [
    {
```

```
    "speaker": "",
    "start_time": 0,
    "end_time": 0,
    "text": ""
  }
],
"protocol": {
  "participants": [],
  "discussion": [],
  "decisions": []
},
"metadata": {
  "duration_ms": 0
}
}
```

Критерии оценки

- Качество распознавания речи — 30%
- Разделение говорящих — 25%
- Структура протокола — 20%
- Производительность — 15%
- Удобство доработки — 10%

Ожидаемый результат

Система, формирующая понятный и структурированный протокол заседания на основе аудиозаписи. Приветствуются доработки, повышающие удобство и практическую ценность решения.