Task 3: Forming an audio protocol

The essence of the problem

Develop a system that automatically generates a structured meeting protocol from an audio recording.

In real processes (meetings, conferences, dissertation councils) it is necessary to record the progress of the discussion and create minutes. Manual transcript occupies significant time .

System must :

- transform speech to text ;
- break up the flow of speech into lines;
- divide speaking ;
- to form readable protocol ;
- allow the user to set participant roles.

Important: the system should not determine the identity of the speaker, only distinguish between participants.

The task comes down to: **building a processing pipeline audio → text → structure.**

Entrance data

Formats :

- `.wav ,.mp3`

Audio Maybe contain :

- some participants ;
- noise ;
- intersections speeches ;
- various quality .

Requirements for the solution

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

```
POST / build_protocol
```

- automatic processing .

Format result

```
{
"transcript": [
{
"speaker": "",
```

```
" start _time ": 0,
" end _time ": 0,
"text ": " "
}
],
"protocol ": {
"participants": [],
"discussion": [],
"decisions": []
},
"metadata ": {
" duration _ms ": 0
}
}
```

Criteria assessments

- Quality recognition speeches - 30%
- Separation speakers - 25%
- Structure protocol - 20%
- Productivity - 15%
- Convenience improvements - 10%

Expected result

A system that generates clear and structured meeting minutes based on audio recordings. Welcome improvements that improve convenience and practicality value solutions .