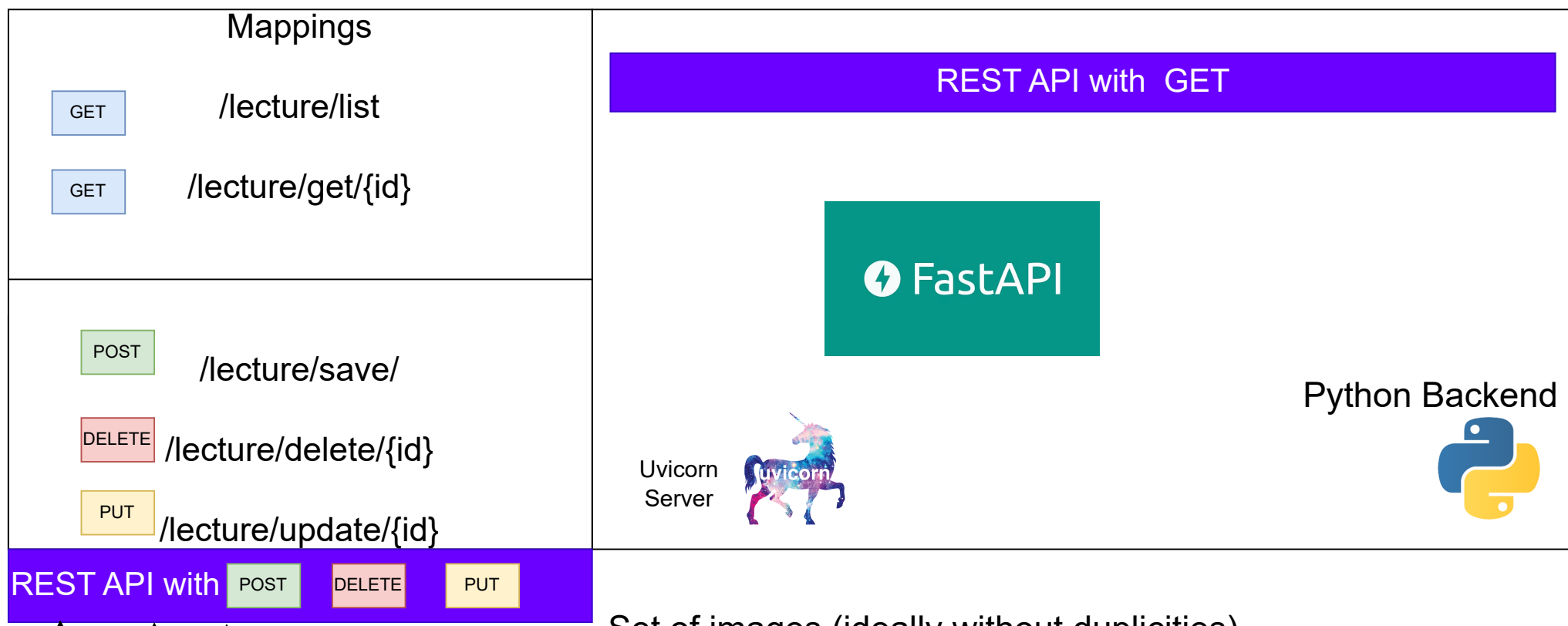
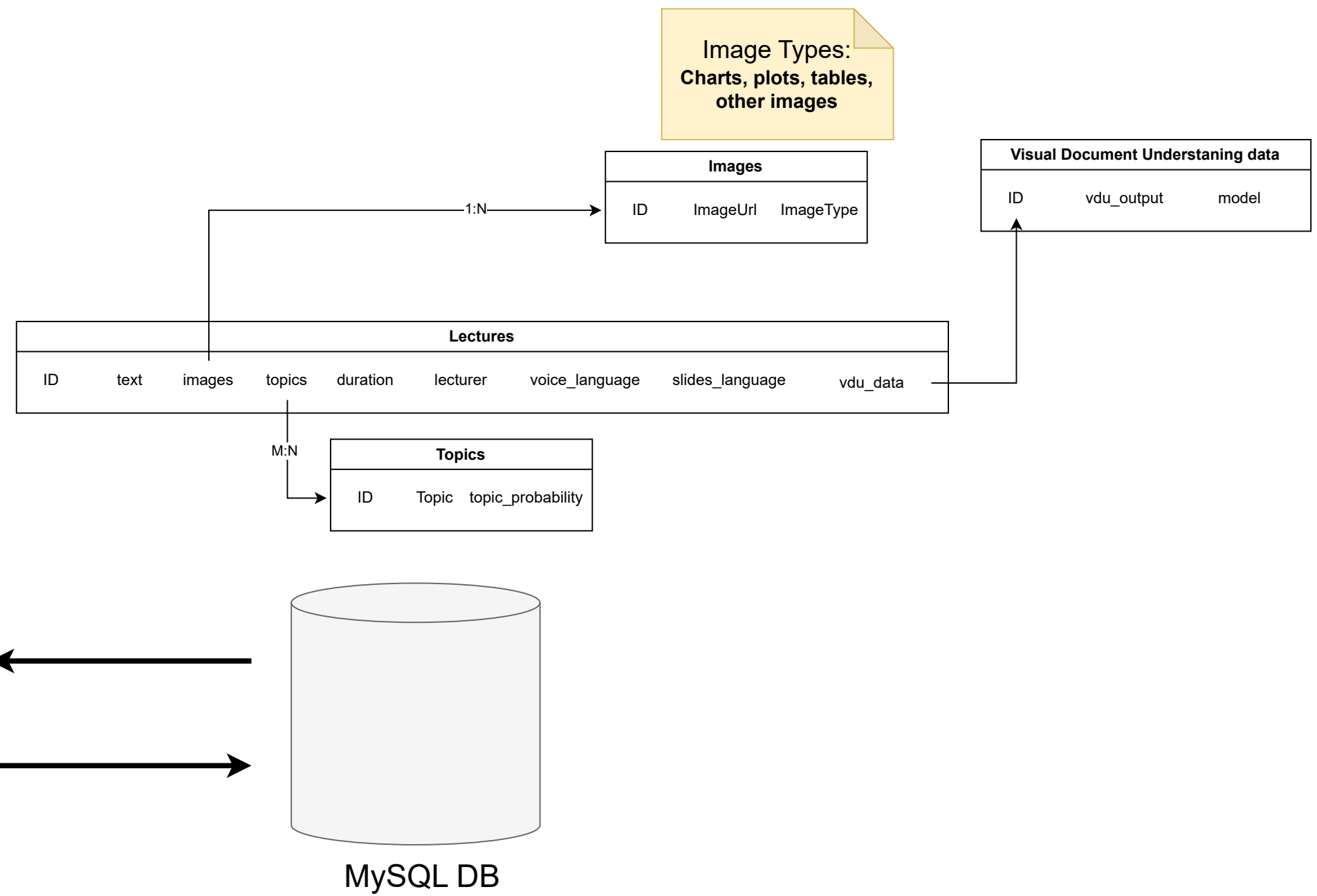
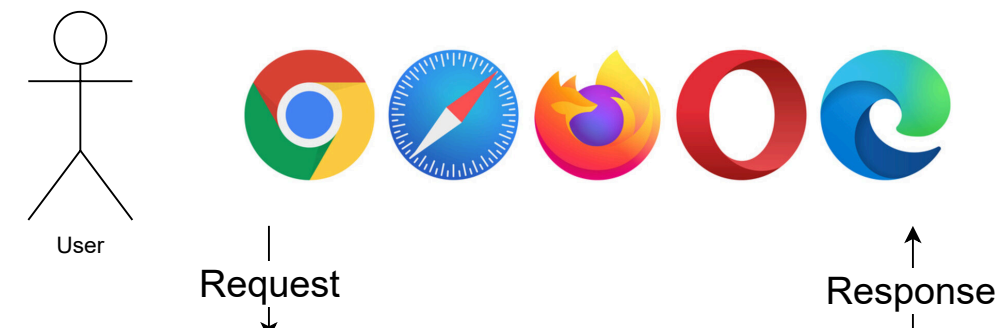
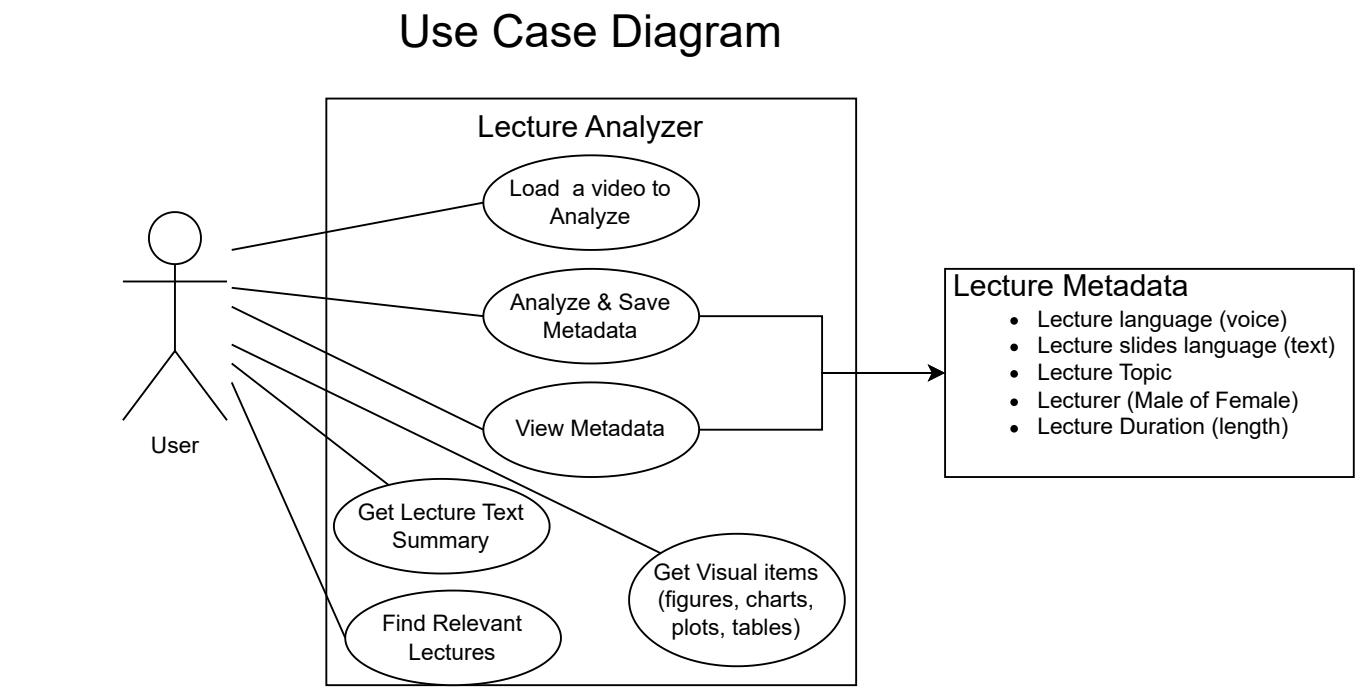


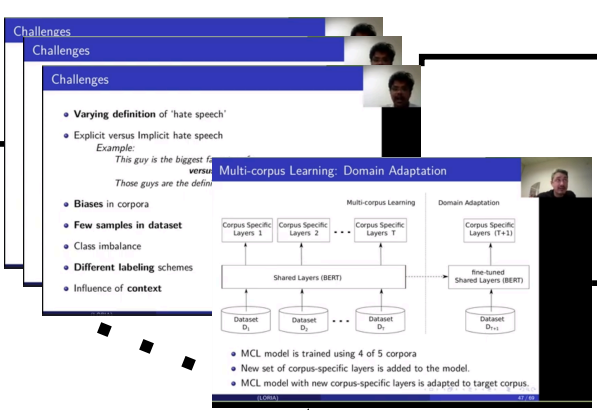
Questions and problems to discuss

Audio Investigator model

How to do text summarization

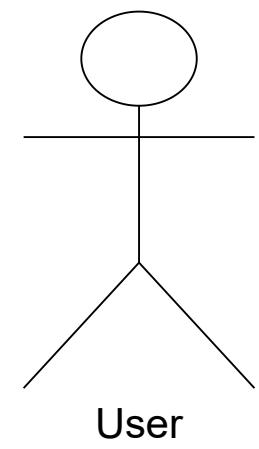
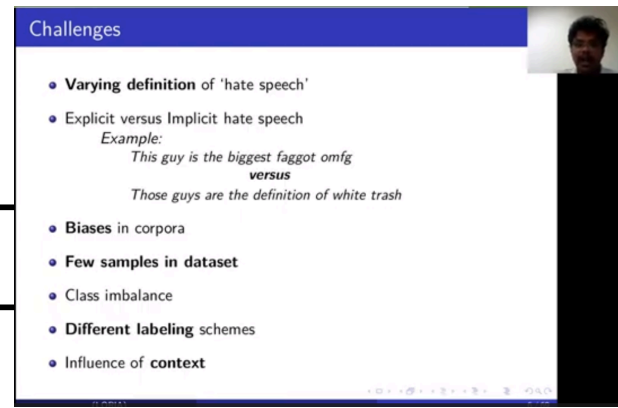


Set of images (ideally without duplicities)



ImageSamplerService

AudioInvestigatorService



User

Input lecture video (MP4, WAV, AVI)

VDUService

OCRService

Text

VDUPostprocessing Pipeline

Postprocessing Pipeline

Postprocessing steps:

- 1) Removing poorly recognized words
--> keeping only meaningful words
- 2) Deduplication (due to the slides)
- 3) Interpunction removal
- 4) Lowercasing

VDU Data

Cleaned Text

Text ClassificationService

TextSummarizationService

Cleaned Text

List[Topic] with probabilities

Summary

