Collaborative AI Audio/Video Editor

Guided by:
Prof. Dr. Ali Arsanjani

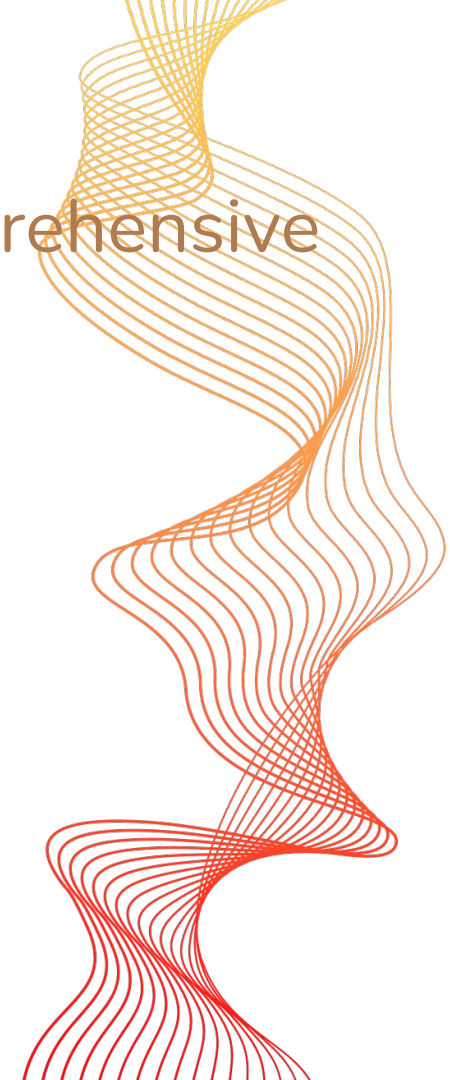
Presented by:
Avinash Ramesh
Yash Kamtekar
Nevil Shah

Acknowledgements

The entire team is deeply indebted to Professor Dr. Ali Arsanjani for his invaluable comments and assistance in the preparation of this study

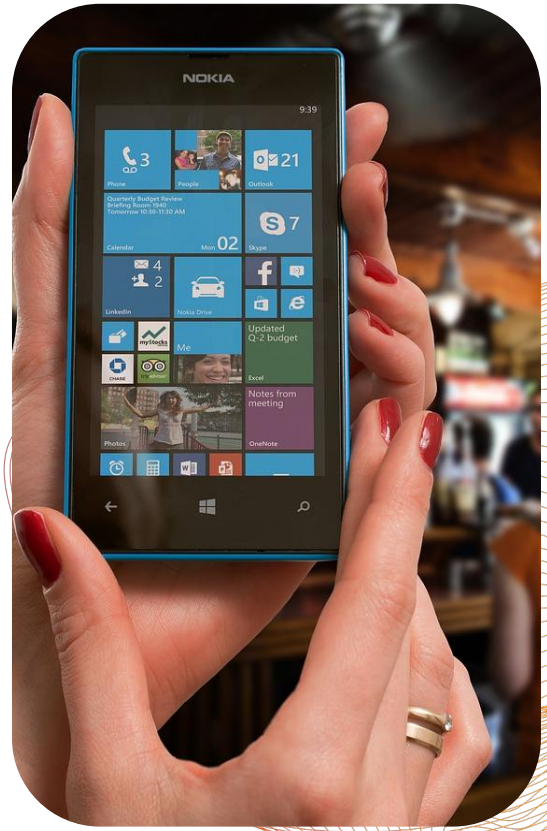
AI Transforms Video Editing: A Comprehensive Editing Tool

This presentation explains the impact of Artificial Intelligence (AI) on the video production industry through its role in expediting video editing workflows, generating realistic visuals, and reducing overall production costs. The presentation also introduces an AI based audio and video editing tool that offers a user-friendly experience and empowers both novice and experienced editors to achieve professional precision.

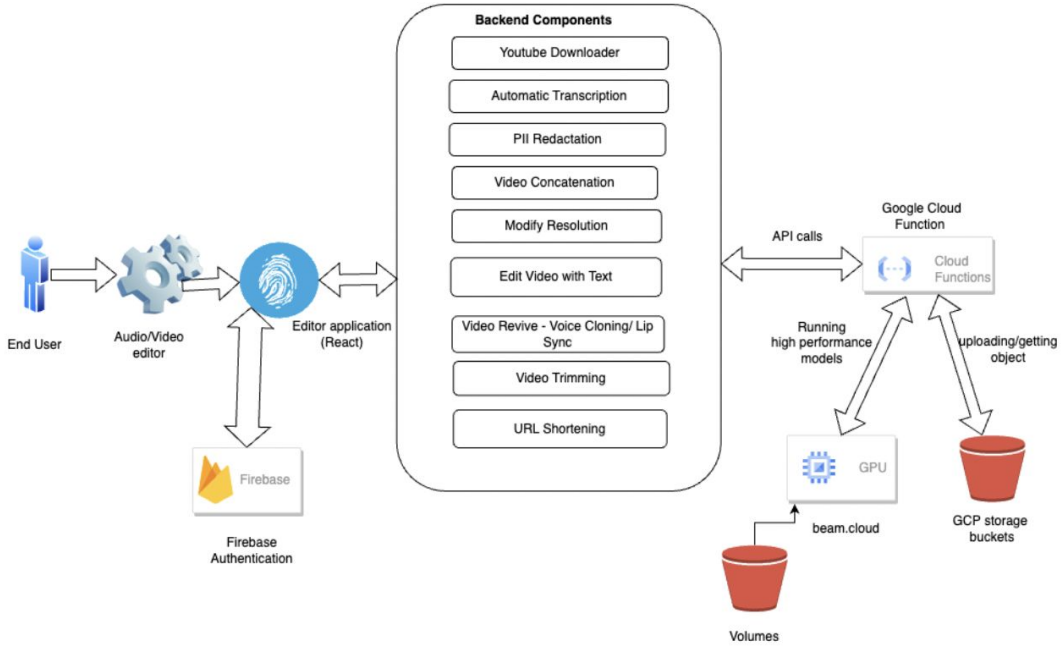


The AI-Based Audio and Video Editing Tool

- Our project introduces an intuitive AI-based audio and video editing tool that offers a comprehensive and user-friendly experience similar to editing a Google Doc.
- The tool harnesses text extracted from transcription to streamline the editing process.
- Cutting-edge state-of-the-art models, transfer learning techniques and features like automatic transcription, one-click removal of filler words, and video segment trimming are employed.



Project Architecture

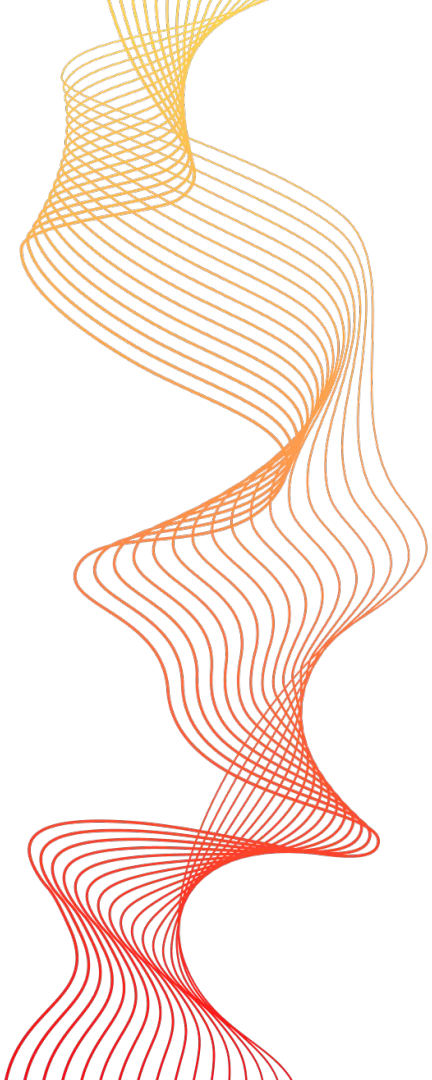


Technology Stack Overview

Client Tier technologies	Data Tier technologies
React	Python
Redux	Firebase
Material UI	Google Cloud Functions (for backend APIs)
	Flask
	Pretrained models : ● GPT Turbo 3.5 ● Tortoise TTS(for personalized voice cloning) ● Wav 2 Lip (for Lip Syncing)
	Beam.cloud(Accelerated AI Model Deployment)

Project Features

Sr. No	Feature
1	Youtube Video Downloader
2	Automatic Transcription (Using Video Model from Google)
3	Speaker Diarization
4	PII Redaction (Using Text Davinci Model from OpenAI)
5	Video Concatenation
6	Modify Resolution
7	Personalized voice cloning (using Tortoise TTS model)
8	Lip Syncing Video (using Wav 2 Lip model)
9	Video trimming
10	Video Cutting
11	URL Shortening
12	Download of SRT file

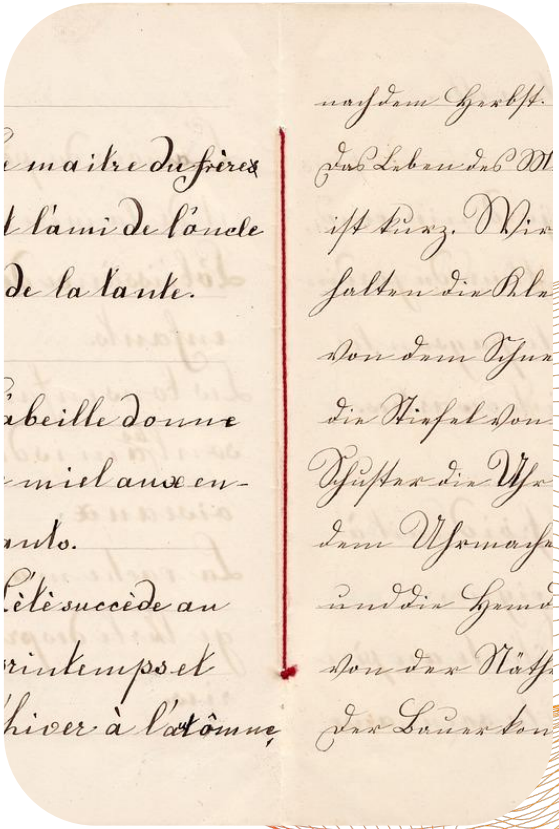


The Impact of AI on Video Production

1) Time Efficiency: AI video editing software provides automated capabilities that drastically cut down on the time needed to complete video editing activities like video trimming, changing the resolution of edited movies, and sharing edited videos.
2) Enhanced Productivity: Users of AI video editing software can automate tedious activities, increasing their productivity for the features including creating srt files, converting speech to text, and eliminating filler words
3) Advanced Editing Capabilities: Even if they lack substantial editing knowledge, content creators can get professional-quality outcomes with the help of cutting-edge technologies like customized voice cloning, PII redaction, lip synchronization, and many more.

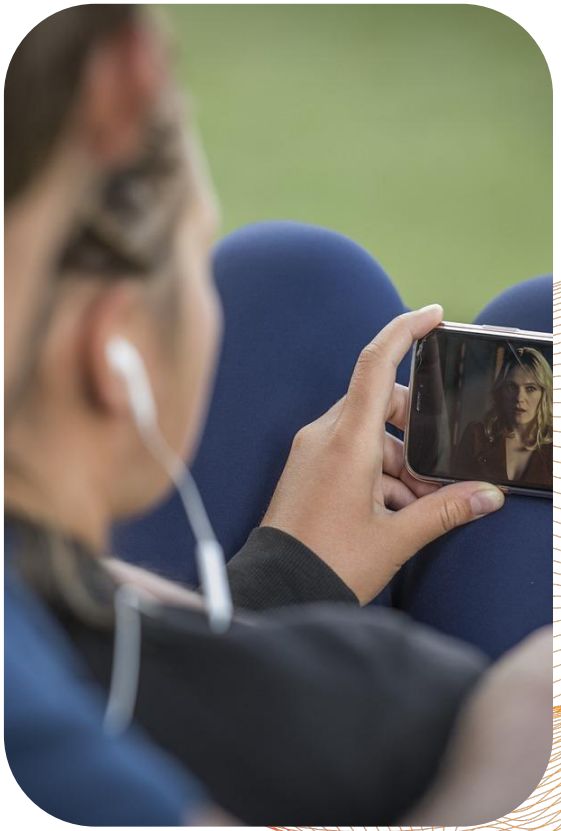
Automatic Transcription

- With the help of Google Speech-to-Text technology, our application features a speech-to-text transcription capability.
- With the help of this technology, users can quickly turn spoken audio into text, making audio material more accessible and searchable.
- Now, users only have to upload or capture audio recordings, and the program will convert the speech into text that can be altered or downloaded for use elsewhere.



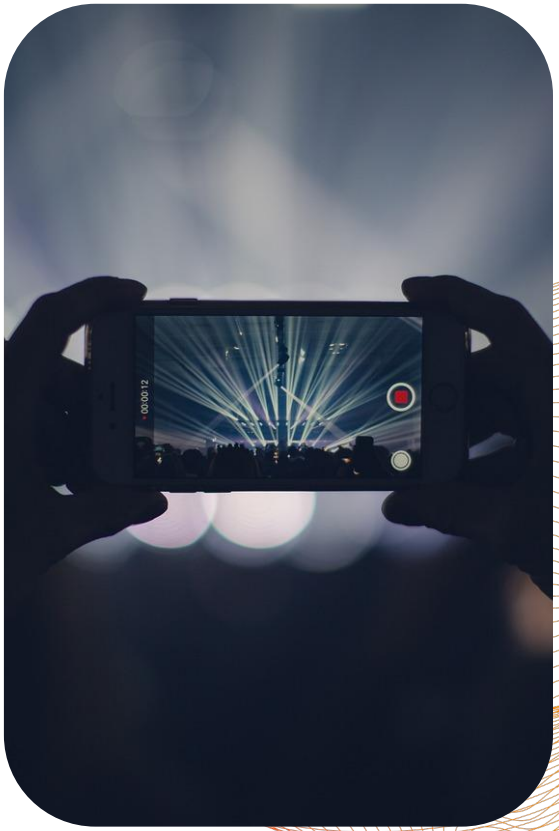
PII Redaction

- This application utilizes GPT-3.5 Turbo model for PII redaction feature which helps users to protect their private information. The technology automatically mutes the matching segments of audio recordings that contain personally identifiable information (PII), such as names, addresses, or phone numbers, thereby preventing the disclosure of sensitive information.
- A new video output is created by smoothly integrating the muted audio parts thanks to the PII redaction capability. This guarantees a fluid viewing experience while preserving data security. Users can post or share the redacted video output with confidence knowing that their personal information and the privacy of others are safeguarded.



Video Editing Guided by the Transcription

- The transcription serves as a guide when editing the video, which eliminates the need to watch the video repeatedly while making modifications.



Video Segment Trimming

- The AI tool enables users to trim video segments accurately and precisely.

Video Concatenation

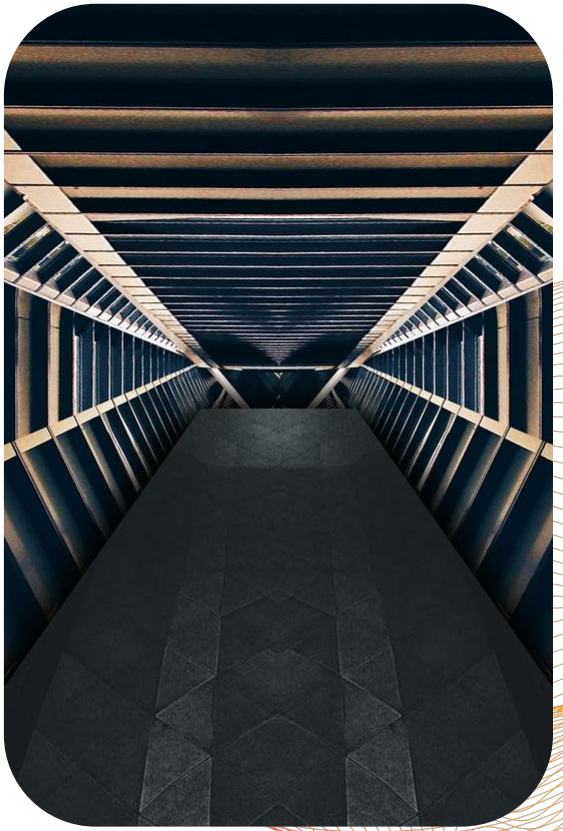
- The tool inserts edited video segments into the original footage with an efficient drag-and-drop feature.
- With the addition of this useful new functionality, our application can now combine two clips into a single output file.
- Users may choose and upload the clips that they want to merge with only a few clicks, and the software will seamlessly stitch them together into a new video.

Personalized Voice Cloning and Lip Syncing

- Our application now contains a powerful new function that allows users to substitute a word in the transcript with an individual's personalized voice utilizing the most recent Tortoise TTS model.
- We've also added Wav2Lip technology, which automatically syncs a video's lips to the new audio to create a seamless new video output. This feature provides users with an exciting new level of control over their audiovisual content, allowing for enhanced personalization and customization.

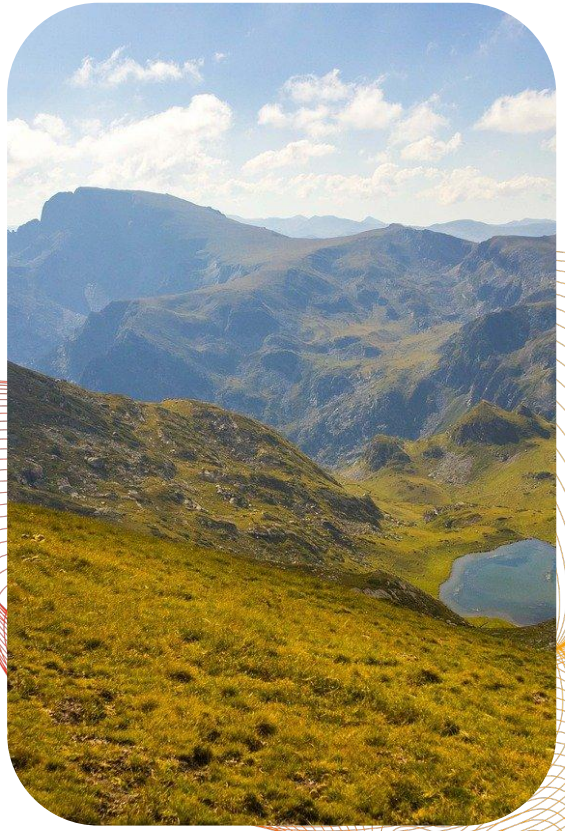
Challenges in Video Editing

- Compute Power Requirement: Demand for extensive GPU compute power to achieve faster execution of multiple video editing features.
- Managing Diverse Video Formats: Difficulty in efficiently handling various video formats while incorporating multiple editing features.
- Ensuring Accurate Results: Challenges in obtaining precise and reliable outcomes for a wide range of editing features within the AI video editor software.



Future of AI-based Video Editing

- Expansion of the application's editing tools, including automatic color correction, background noise reduction, and sophisticated video stabilization, through the integration of more AI and machine learning models.
- Research into more sophisticated voice cloning methods that would allow users to create synthetic voices that were both realistic and programmable for their edited content.
- Real-time co-editing, version control, and project management tools are all examples of collaboration features that can be used to improve teamwork and speed up the editing process for big projects



Thank you. Please feel free to ask any questions.

😄