

CLOUD COMPUTING – CS623

A Cloud Based Smart Meeting Transcription & Analysis Platform

Murali Krishna Kancheti – yr1529
Karthik Veeravalli – df8025
Sunnyhita Guda – dj2361

Project Overview:

This project offers an end-to-end solution to capture, transcribe, store, and analyze speech recorded in meetings. It leverages Google Cloud Speech-to-Text API, Google Cloud Storage, OpenAI's API for text analysis, and incorporates FFmpeg for media handling.

Description:

The project creates an integrated solution for managing meeting recordings by utilizing Google Cloud Speech-to-Text API, Google Cloud Storage, OpenAI's API, and FFmpeg. Audio from meetings is first processed and optimized using FFmpeg, ensuring quality and compatibility. It is then transcribed into text using Google's Speech-to-Text service and stored securely in Google Cloud Storage. This textual data is analyzed by feeding it to OpenAI's ChatGPT to generate summaries or identify specific to-do tasks. The entire workflow creates an automated, efficient, and intelligent system for handling and deriving insights from meeting recordings, integrating various cutting-edge technologies to maximize accuracy and utility.

Output:

The output is in form of a text file which contains AI generated response for the specified prompts.

```
Transcript for the given file: ["hello everyone thank you guys for coming to our weekly Student Success meeting and let's just get started so I have our list of chronically absent students here and I've been noticing a troubling trend I'm a lot of students are skipping on Fridays does anyone have any idea what's going on I've heard some of my mads he is talking about how it's really hard to get out of bed on Fridays it might be good if we did something like a pancake breakfast to encourage them to come I think that's a great idea let's try that next week", "it might also be because a lot of students have been getting sick now that it's getting colder outside of having numbers didn't come by my office with symptoms like sniffing and coughing we should put up posters with tips for not getting sick since it's almost Lucy's and I like to wash your hands after the bathroom stuff like that I think that's a good idea and it'll be a good reminder for the teachers as well as one of the things I want to talk about there's a student I've noticed here John Smith he's missed seven days already and it's only November does anyone have an idea what's going on with him I might be able to fill in the gaps there I talked to John today and he's really stressed out he's been dealing with helping his parents take care of his younger siblings during the day it might actually be a good idea if you spoke to the guidance counselor a little bit I could talk to John today if you want to send him to my office after you meet with him a lot to deal with for a middle schooler", "great thanks and I can help out with the family child care needs I'll look for some free or low-cost resources in the community to share with John and he can share them with his family today thanks for coming and if no one has anything else I think we can wrap up"]
Response from Model(GCP): The meeting is about the chronically absent students. The attendees discussed the reasons for students skipping school and brainstormed solutions. They also discussed John Smith, a student who has missed a lot of school. The meeting was productive and ended with a plan to help John.
Response from CHAT GPT: In a weekly Student Success meeting, the participants discuss the issue of chronically absent students, specifically the trend of students skipping on Fridays. They propose having a pancake breakfast to encourage attendance. The discussion also includes the topic of students getting sick due to the colder weather, and the idea of putting up posters with tips on staying healthy, including washing hands. A particular student, John Smith, who has already missed seven days in November, is mentioned. It is discovered that John is dealing with family responsibilities and stress, so the suggestion is made to have him speak with the guidance counselor. The meeting concludes with a plan to find resources for John's family and a reminder to wrap up if there are no further issues to discuss.
Transcript file for the given file generated successfully
(base) muralikrishnakancheti@Muralis-MacBook-Pro SpringML_Summary_POC %
```

gin ⌘ Configure ↩ Cloud Code ↗ murali-pde

Ln 32, Col 69 Spaces: 4 UTF-8 LF Python 3.11.1 64-bit

Key Components:

Media Handling (FFmpeg)

- Tool: FFmpeg
- Function: Processes and prepares audio files for transcription.
- Features: Supports various formats and codecs, enables conversion, filtering, and optimization of audio files.

Speech Transcription

- Tool: Google Cloud Speech-to-Text API
- Function: Converts audio into text.
- Features: Multilingual support, real-time transcription.

Data Storage

- Tool: Google Cloud Storage
- Function: Stores the textual transcripts securely.
- Features: High durability, various storage classes.

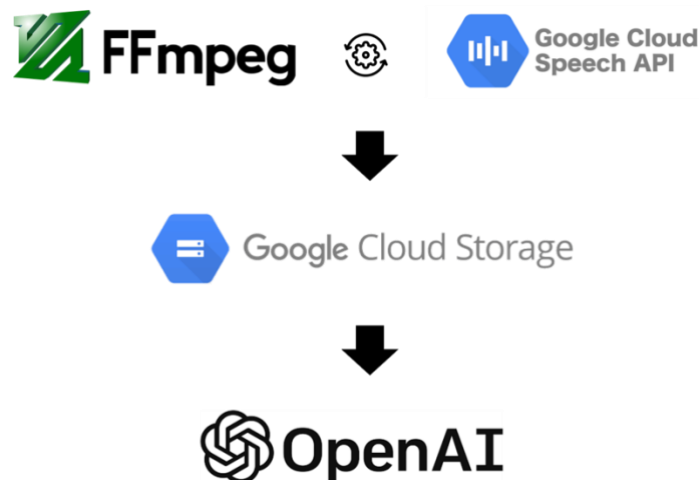
Text Analysis

- Tool: OpenAI's ChatGPT through OpenAI API
- Function: Summarizes transcripts or identifies to-do tasks.
- Features: Advanced natural language processing.

Workflow:

1. Process Audio Files:
Utilizes FFmpeg to process and prepare audio files from meetings.
2. Transcribe Audio:
Google Cloud Speech-to-Text API is used to transcribe the processed audio into text.
3. Store Transcripts:
Transcriptions are stored in Google Cloud Storage for later retrieval.
4. Analyze Transcripts:
Transcriptions are sent as prompts to ChatGPT to generate summaries or identify tasks.

5. Utilize Output:
Insights derived from OpenAI's API are used for follow-up actions.
6. AI-Generated Responses:
OpenAI's ChatGPT analyzes the prompts and generates responses. These responses might include concise summaries of the meeting or a clear list of tasks to be undertaken.
7. Text File Output:
The generated responses are consolidated into a text file. This file contains the AI's analysis of the meeting in the requested format, whether summaries, tasks, or other insights.



Technologies Used:

Google Cloud: Speech-to-Text API

Google Cloud Speech-to-Text API is a service that translates audio into written text. It supports over 125 languages and variants, and can be used in real-time or for batch processing. Known for its high accuracy, even in noisy environments, it provides features like automatic punctuation, word-level confidence scores, and customization for specific terms or accents.

Google Cloud: Storage

Google Cloud Storage is a scalable and highly durable object storage service provided by Google Cloud. It enables businesses and developers to store, access, and manage vast amounts of data effortlessly. With support for various types of data storage classes, ranging from high-frequency access (like multi-regional) to archival solutions (like coldline storage), it meets a wide array of storage needs.

Google Cloud: Storage API

Google Cloud Storage API is a service that provides scalable and flexible object storage for developers and enterprises. It allows for the storage and retrieval of any amount of data at any time, facilitating integration with various applications, websites, or data processing tasks. The API supports standard protocols like HTTP/HTTPS and offers various types of storage classes to cater to different needs, including regional, multi-regional, nearline, and coldline storage. Security is robust with fine-grained access controls, and the service is known for its high availability and durability.

Open AI API

OpenAI's API provides developers access to powerful machine learning models created by OpenAI. It allows for various natural language processing tasks including text generation, summarization, translation, and question-answering. Leveraging state-of-the-art models like GPT-3 and beyond, it enables the creation of intelligent applications and services capable of human-like text understanding and generation. The API is designed to be versatile and user-friendly, catering to both experienced AI practitioners and those new to the field. With clear documentation and supportive community, it's a widely used tool for incorporating advanced AI capabilities.

FFMPEG

FFmpeg is a widely used open-source software project that consists of a collection of tools and libraries for handling multimedia data. It enables users to convert, record, stream, and manipulate video and audio files in various formats. Known for its flexibility and robustness, FFmpeg supports a wide range of codecs, filters, and formats, making it a popular choice for professionals and hobbyists working with multimedia. Its command-line tools, such as *ffmpeg*, *ffplay*, and *ffprobe*, provide comprehensive control over media processing tasks. Being cross-platform, FFmpeg can be used on various operating systems including Windows, macOS, and Linux. Its extensive feature set and active community support make FFmpeg an essential tool in the field of multimedia processing.

Benefits:

1. **Enhanced Quality**
FFmpeg ensures that audio files are in the optimal format, enhancing transcription accuracy.
2. **Automation**
Facilitates a seamless flow from recording to analysis, reducing manual effort.
3. **Scalability**
Adaptable to various meeting sizes and content complexities.
4. **Integration**
Utilizes best-of-breed technologies for a holistic solution.

Challenges:

1. **Complexity**
Managing multiple technologies requires careful coordination and expertise.
2. **Customization**
Adapting the system to specific organizational needs or industry terminologies.
3. **Cost Considerations**
Managing associated costs with paid services.
4. **Privacy and Security**
Ensuring that sensitive information remains secure throughout the process.

Future Scope:

The future scope of this project, which integrates various technologies to process meeting recordings, is ripe for expansion. Potential directions include real-time transcription and analysis for immediate insights, customized summarization algorithms tailored to specific needs, integration with task management systems, and enhanced security measures. Scalability to handle diverse organizational needs and the addition of user-friendly interfaces could further enhance its applicability and efficiency. These enhancements could transform the project into a vital tool for intelligent organizational communication and decision-making.

Conclusion:

This project presents an innovative approach to automating the documentation and analysis of meetings by integrating FFmpeg, Google Cloud's Speech-to-Text and Storage, and OpenAI's ChatGPT. By including FFmpeg in the workflow, it ensures optimal audio quality, enhancing the overall effectiveness of the system. While it does present challenges in terms of complexity and cost, the benefits of streamlined efficiency, accuracy, and scalability make it a valuable asset for organizations seeking to leverage technology to enhance their meeting processes. Continued attention to customization and security will be essential for realizing the full potential of this integrated solution.