AUDIO MODULE

SOFTWARE REQUIREMENT SPECIFICATION

VERSION 2.0

DECEMBER 10, 2021

CALIFORNIA STATE POLYTECHNIC UNIVERSITY, POMONA

3801 W. TEMPLE AVE., POMONA, CA 91768

REVISIONS

Version	Primary Author(s)	Description of Version	Date Completed
Initial Draft	Gershman, Maribel Perdoza, Abrahan Patel, Nachiket Garcia, Diego	Draft version	11/01/2021
1.0	Gershman, Maribel Perdoza, Abrahan Patel, Nachiket Garcia, Diego	Using AWS as the speech to text API	11/23/2021
2.0	Gershman, Maribel Perdoza, Abrahan Patel, Nachiket Garcia, Diego	Changed from AWS to Google Speech To Text API	12/05/2021

CONTENTS

INTRODUCTION

The Software Requirements Specification (SRS) describes the global requirements of the Audio Module (AM). The requirements were developed specifically for the Mobile Interview Application being developed in the Computer Science (CS) 4800 Software Development Fall 2021 Course, but it is believed to be suitable for many other applications such as those involving applications for the hearing impaired.

This SRS includes only those requirements that were deemed necessary to perform the desired outcome, connect with the Web Services Team and support both the Front-End Teams and the Back-End Teams.

Most of the requirements in this document originated from the Class Project description, Professor Zaidi, and work descripted in specific modules such as the Web-Services, Front-End, and Back-End Modules.

PURPOSE

The purpose of this SRS document is to collect and analyze requests as to define the AM requirements to best support the Mobile Interview Application. It will define and predict how the AM will behave and response, outline concepts which may be developed later (beyond our class time frame), document ideas being considered.

The purpose of this SRS document is to provide a detailed overview of our module, its parameters and goals. This document describes the module's target audience and user interface, hardware and software requirements. It defines how our client, team and audience see the module and its functionality. It also helps the developers assist with the software delivery lifecycle (SDLC) process.

SCOPE

This SRS is aimed at specifying requirements of the Audio Module. Specifically:

- System interfaces
- Acquiring an audio file
- Generating a text file
- Generating a confidence value
- Delivering audio analysis

The specifications focus on functional characteristics. Data structures and user interfaces will be developed using the agile iterative method.

DEFINITIONS, ACRONYMS, AND ABBREVIATIONS

SRS	Software Requirements Specification
MIA	Mobile Interview Application
AM	Audio Module
FEM	Front End Module
BEM	Back End Module
WSM	Web Services Module
KBF	KanbanFlow
GST	Google Speech to Text
API	Application Programming Interface
FAQ	Frequently Asked Questions

REFERENCES

The references are:

- Use Case Diagrams
- Activity Diagrams
- Domain Object Model
- Context Diagram
- Architecture Layout
- Component Diagram
- Class Hierarchy & Relationship Diagrams
- Sequence Diagrams
- Test Strategy/Plan
- Traceability Matrix
- Test Cases

- Test Results Summary (not sure if this one belongs in here)

OVERVIEW

The Audio Module will receive a video file from the Web Services Team and convert it to an audio only file, generate a transcription of the audio, generate a confidence number of the speaker in the audio and deliver both the transcription and confidence level to the Speech Analysis Team for further processing.

OVERALL DESCRIPTION

The Mobile Interview Application (MIA) needs to process and analyze the video recorded answers produced by the interviewee during a mobile interview session. The AM will process the video and support the analysis by converting the audio portion of the video into text and determine the interviewee's confidence in the response(s) given.

PRODUCT PERSPECTIVE

While the AM is self-contained it does require two things:

- The video file's location should be public and be clearly specified or made directly available to the AM.
- The location of the destination of the results be public and be specified or the default location will be within the current Google Bucket in the Google Cloud Drive labeled CS4800.

PRODUCT FUNCTIONS

The module shall be able to access the file from an external public site.

They module shall accept video files in mp4 format.

The module shall notify the user should an mp4 file is not found.

The module shall notify the user of an mp4 file of size 0.

The module shall convert mp4 files into mp3 files.

The module shall notify the user of any errors resulting from converting the mp4 file into an mp3 file.

The module shall convert mp3 files into text files.

The module shall convert mp3 files into json files.

The module shall notify the user of any errors encountered while converting the mp3 file into text.

The module shall notify the user of any errors encountered while converting the mp3 file into a json file.

The module shall count how many filler words are contained within the text file.

The module shall generate a confidence number based on the audio's mp3 pitch.

The module shall deliver a text file containing the transcription and the confidence number.

The module shall notify the user of any errors occurring with the delivery of the text file.

USER CHARACTERISTICS

The user intended is the Web Services Team from the Mobile Interview Application. The user shall be a student in the CS 4800 Software Development course at California Polytechnic University, Pomona in the Fall of 2021.

CONSTRAINTS

The following are items which will constrain the design and development of the AM:

GST word transcription confidence level averages 90%

The user must have access to back4App and GitHub

Must be able to communicate with the Web Team

ASSUMPTIONS AND DEPENDENCIES

Accepted filed types.

Filler word list.

Video file retrieval location.

Text file delivery location.

APPORTIONING OF REQUIREMENTS

These requirements will be delayed until a future version of the system:

Anything we do not finish due to time constraints, meaning we ran out of time.

SPECIFIC REQUIREMENTS

This should describe all requirements at a sufficient level of detail for developers to design a system satisfying the requirements and testers to verity that the system satisfies requirements. Basically, we are waiting on the Web Team to tell us what they need to connect with us.

EXTERNAL INTERFACE REQUIREMENTS

To Be Determined

Waiting for External Interface Requirements from the other modules.

USER INTERFACES

To Be Determined

Waiting for External Interface Requirements from the other modules.

HARDWARE INTERFACES

To Be Determined

Waiting for External Interface Requirements from the other modules.

SOFTWARE INTERFACES

To Be Determined

Waiting for External Interface Requirements from the other modules.

COMMUNICATIONS INTERFACES

To Be Determined

Waiting for External Interface Requirements from the other modules.

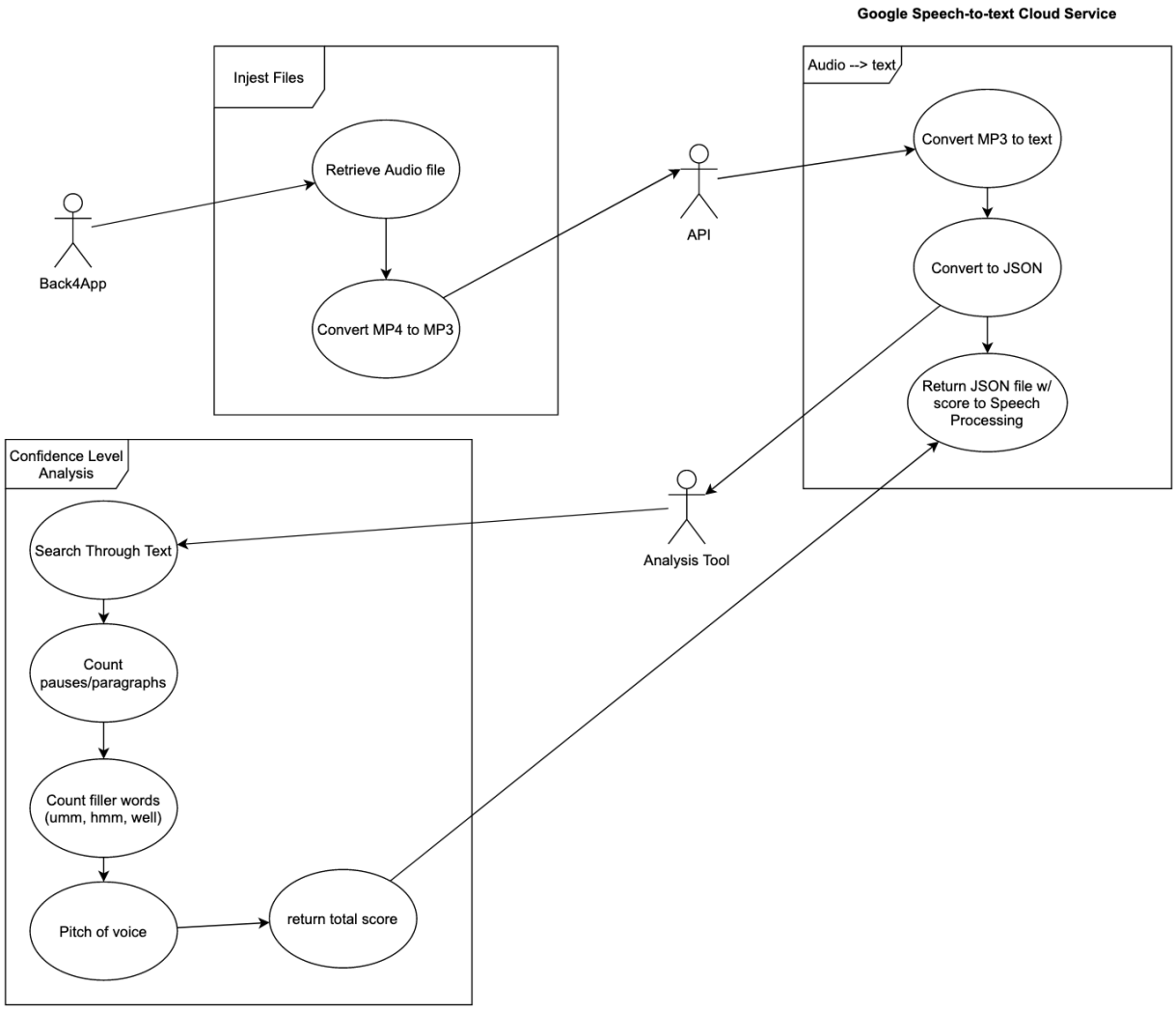
SOFTWARE PRODUCT FEATURES

VIEW AVAILABLE NETWORK CONNECTIONS

PURPOSE

To allow users to view the high-level functions and scope of the AM. This also identifies the interactions between the AM and the other modules and defines the context and requirements of the entire AM.

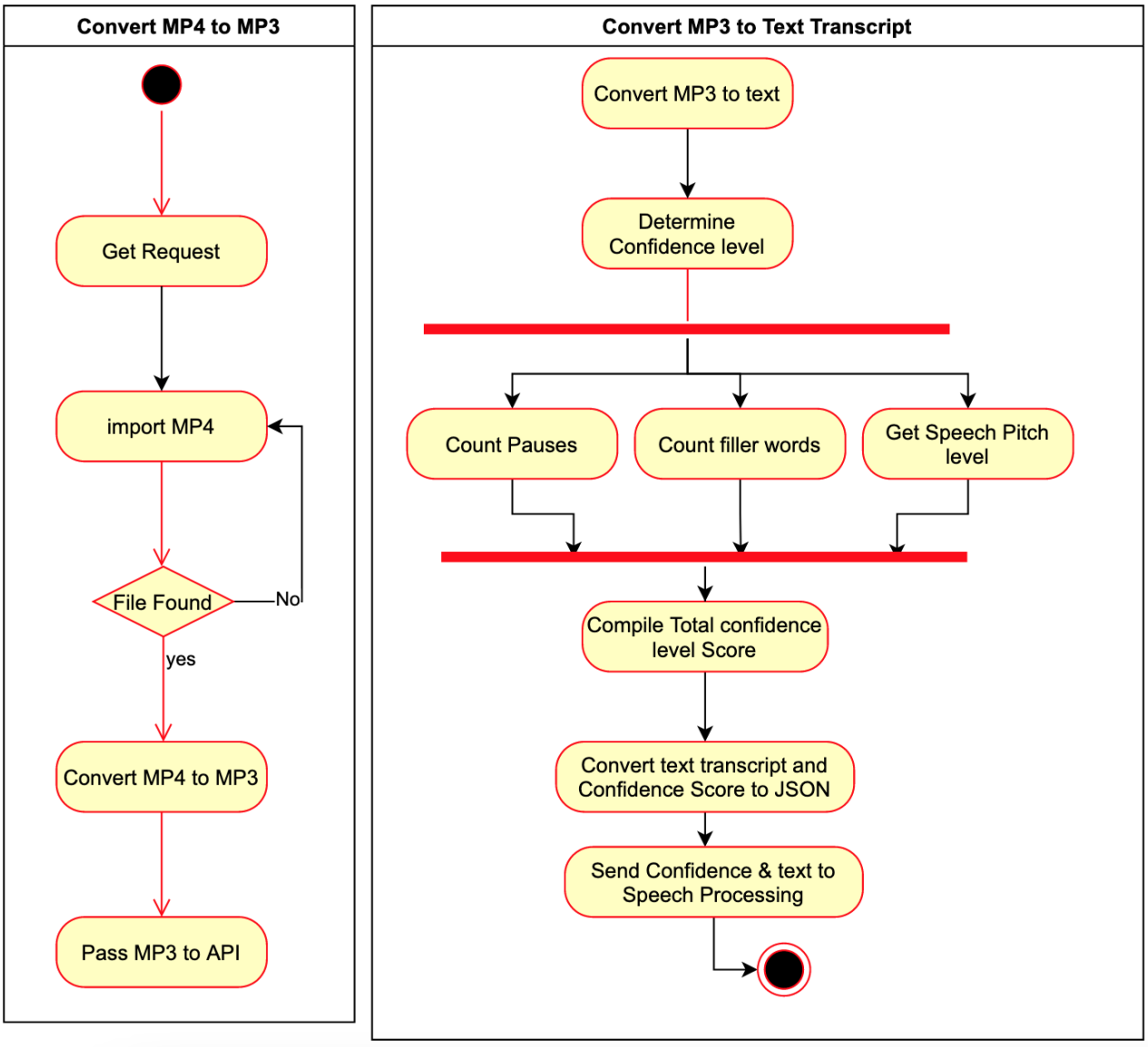
USE CASE



ASSOCIATED FUNCTIONAL REQUIREMENTS

NETWORK CONNECTIONS

Use system calls to analyze interviewee's confidence in their responses.

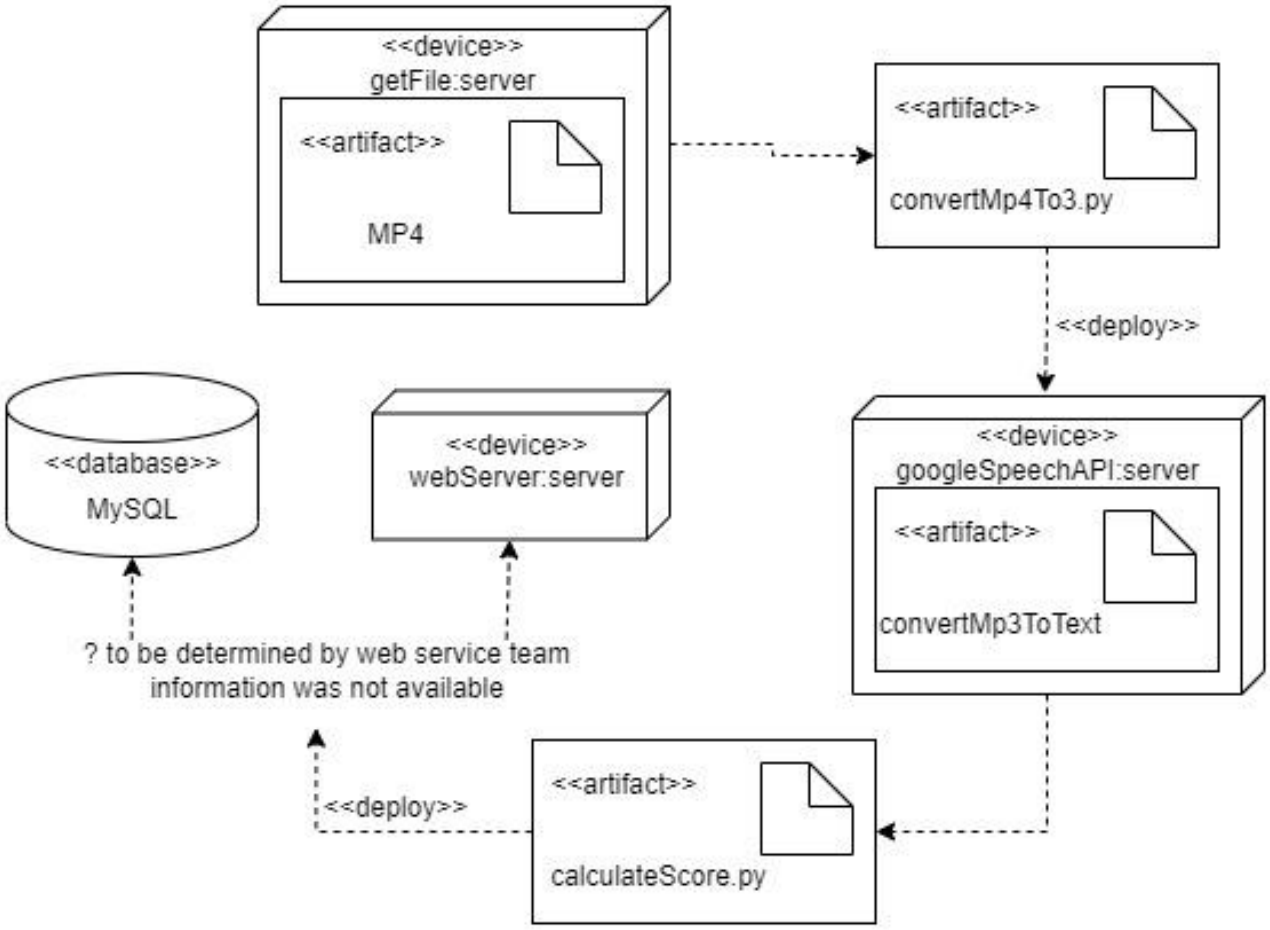

Convert MP4 to MP3
Get Request
import MP4
File Found
No
yes
Convert MP4 to MP3
Pass MP3 to API
Convert MP3 to Text Transcript
Convert MP3 to text
Determine Confidence level
Count Pauses
Count filler words
Get Speech Pitch level
Compile Total confidence level Score
Convert text transcript and Confidence Score to JSON
Send Confidence & text to Speech Processing

CONNECT TO MODULES

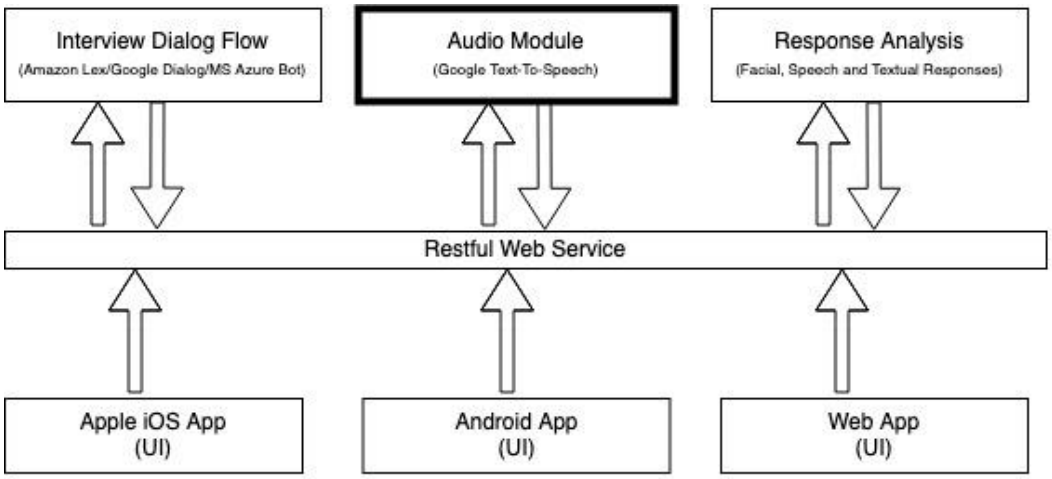
PURPOSE

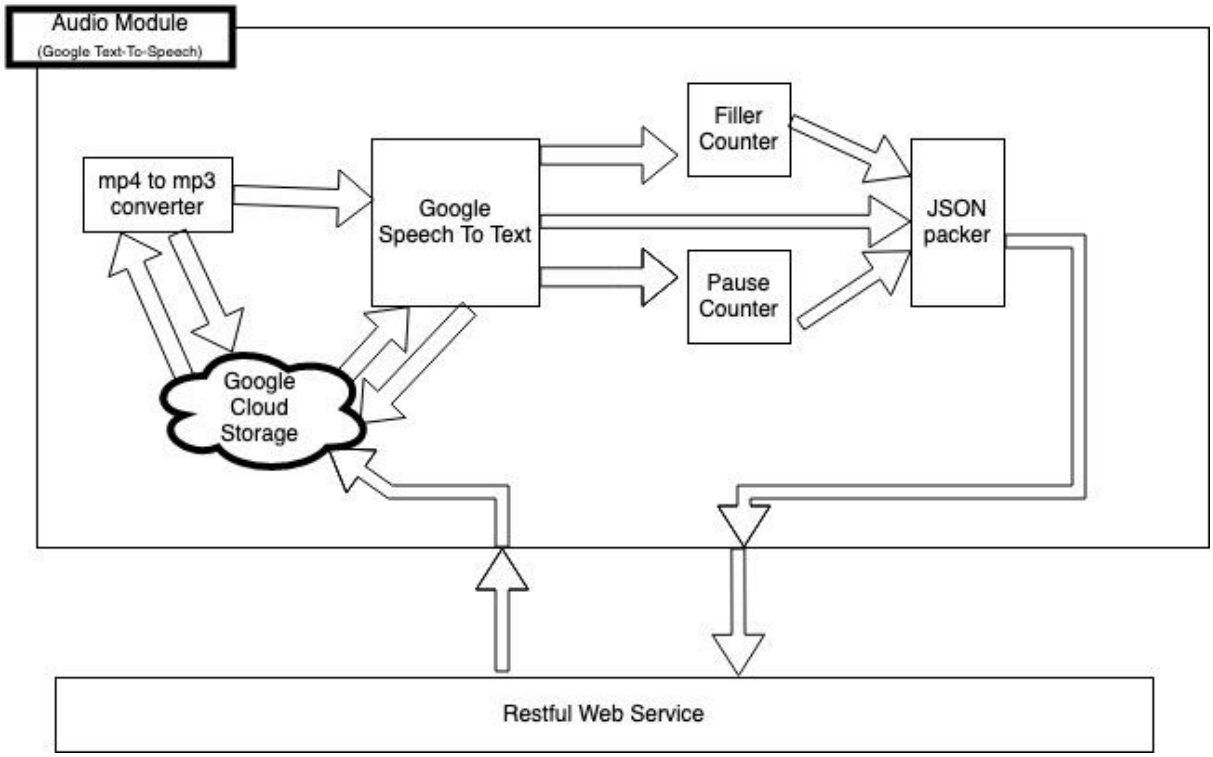
To allow users to connect to the Audio Module and understand what is needed to connect to it. To allow users to understand what the Audio Module will produce.

CONTEXT DIAGRAM

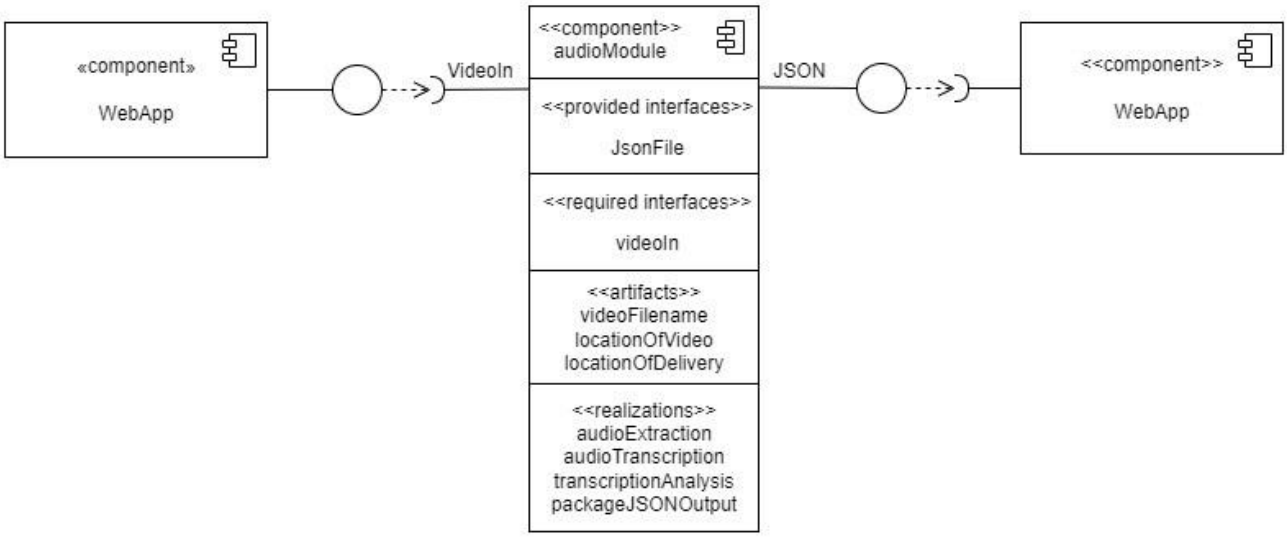


ARCHITECTURE LAYOUT DIAGRAM

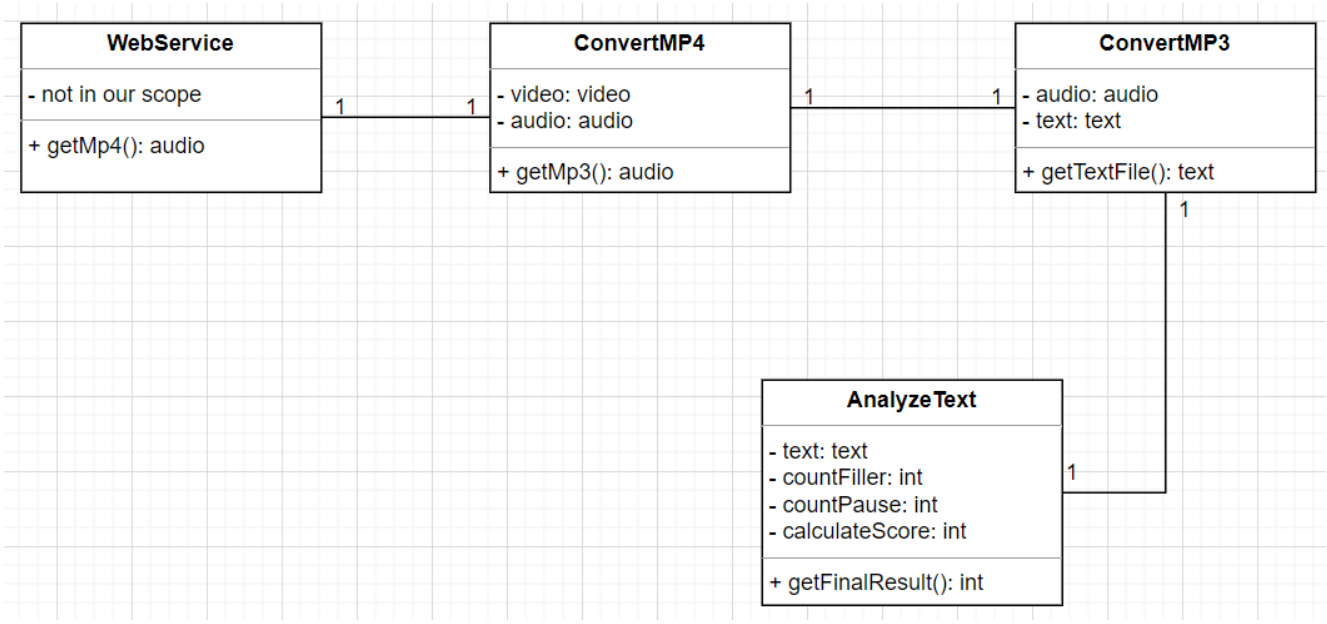




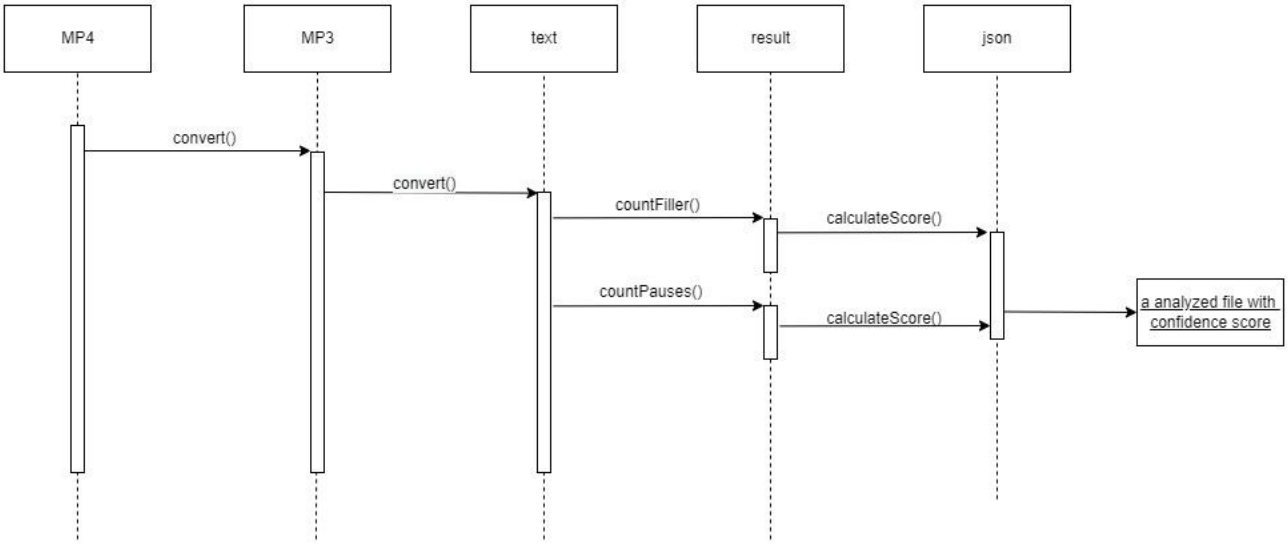
COMPONENT DIAGRAM



CLASS HIERARCHY & RELATIONSHIP DIAGRAM



SEQUENCE DIAGRAM



PERFORMANCE REQUIREMENTS

Only one audio file will be processed at any one time.

The Audio Module will supply a text file within 24 hours.

The name of any Google Bucket will not exceed 64 characters, if the name does not contain a dot, it can be up to 222 characters.

DESIGN CONSTRAINTS

Video file must be in an mp4 format.

Audio file must be in an mp3 format.

Audio files greater than 15 seconds long must be placed in the CS4800 bucket in the Google Cloud Drive.

The Google Bucket can have up to a total of 100 notification configurations.

SOFTWARE SYSTEM ATTRIBUTES

This Audio Module can be modified to transcribe:

- speech to text in real time
- languages other than U.S. English.

MAINTAINABILITY

The text file of filler word can be updated at any time and several custom lists can be maintained at the user's discretion.

The actual Google Speech to Text engine is maintained by Google.

LOGICAL DATABASE REQUIREMENTS

- The database must be publicly available.
- Filler word file must be of text format.
- Video files must be of mp4 format.
- Only one video file may be processed at a time.

OTHER REQUIREMENTS

None