

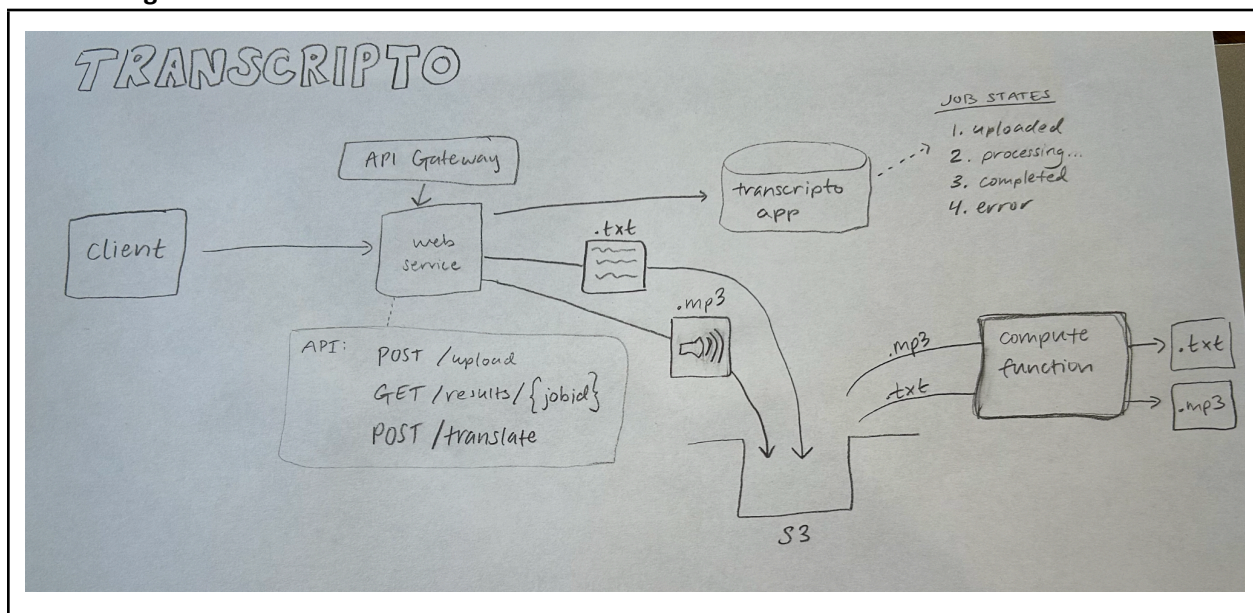
Ellen Tomlins

CS 310 Final Project: **Transcripto**

My project is called Transcripto, a web service designed to bridge the gap between different text and speech mediums as well as languages. The main problem I wanted to tackle is the annoyance of getting transcriptions from audio files. The general use case of the app is as follows: say I am an aspiring speechwriter, and I really like a quote in a speech I heard, but I only have the .mp3 file of the speech. Using Transcripto, I can convert that audio file into a transcript (.txt file), then I can translate that transcript into any language I want to share the quote with the world. Then, say I want to write my own speech, but before delivering it myself, I'd like to hear it read back to me to hear how it sounds out loud. Using Transcripto again, I can upload my .txt file and Transcripto will output an audio file dictating that speech to me.

The three non-trivial operations that my web service performs is 1) transcribe audio into text using Amazon Transcribe, 2) turn text into an audio file using Amazon Polly, and 3) translate the audio/text into different languages using Amazon Translate. I implemented this web service using AWS Lambda, API gateway, S3, and RDS.

Server Design Overview



API operations:

1. POST /upload

This API operation uploads either an .mp3 or a .txt to S3 bucket, which triggers a lambda function to process the file and convert .mp3 to .txt or .txt to .mp3, respectively. The client prompts the user for the local path to the file he wants to upload, then asks them to enter the job_type they want to perform (transcription or text_to_speech). This assembles the payload and query parameters (shown below):

```
payload = {
    "filename": "gettysburg_address.mp3",
    "data": "<base64-encoded-file>"
}
params = {"job_type": "transcription" } # or text_to_speech
```

Status code 200 indicates a success and returns the jobid generated for the task in the database.

Status code 400 indicates that there was a bad request, triggered when the file is not a .mp3. However, because this lambda function is triggered only when a .mp3 function is uploaded to the S3, this should never return a 400. To prepare future iterations of Transcripto, though, this is good error handling. The same holds for the text_to_speech lambda function with .txt files.

Status code 500 indicates an internal server error.

2. GET /results/{jobid}

This operation fetches the status and results of a given jobid. The only path parameter is jobid, entered by the user using the client (`url = f"{baseurl}/results/{jobid}/"`), jobid being the path parameter. This triggers the `finalproj_download` function, which retrieves the status of the jobid and downloads the results if they are available.

Status code 400 means the jobid was not found in the database.

Status code 480 means the status is uploaded.

Status code 481 means the status is processing.

Status code 482 means the status is an error getting the response from AWS Polly or Transcribe.

Status code 500 indicates an internal server error.

Status code 200 means that the status is completed, the output is sent back to the user. For `text_to_speech`, the url with the results is sent to the client, and for transcription the text of the transcription is sent to the client.

```
# for text_to_speech:
return {
    'statusCode': 200,
    'body': json.dumps({
        'results_url':
f"https://{bucketname}.s3.amazonaws.com/{results_file_key}"
    })
}

# for transcription:
return {
    'statusCode': 200,
    'body': json.dumps(transcript_text)
}
```

3. POST /translate

This API call translates a .txt file into another language using AWS Translate. The client prompts the user for a .txt file and the language they want to translate the .txt file into. The user must provide the language code for the language they want to change the text into. The payload is shown below.

```
{
  "text": "Four score and seven years ago...",
  "target_language": "es"
}
```

Status code 400 means that the body was sent without text or target_language.

Status code 500 indicates an internal server error.

Status code 200 means that the operation succeeded, and the packet sent back to the user is the jobid and the URL of the text file where the translated text is stored in the S3 bucket (shown below).

```
'body': json.dumps({
  'job_id': jobid,
  'results_url':
  f"https://{bucketname}.s3.amazonaws.com/{filename}"
})
```

Database design:

This database keeps track of the jobs uploaded to the web service. When a job is uploaded via the API call /upload, it is inserted into the database with status 'uploaded,' and the status is updated accordingly as the job is processed, ending with either 'completed' or 'error.'

jobid	status	jobtype	originaldatafile	datafilekey	resultsfilekey
1001	'uploaded', 'processing...' , 'completed', 'error'	'transcription', 'text_to_speech', 'translation'	gettysburg_address .mp3	'transcriptoapp/ gettysburg_address- cffee65c-e036-4f8d-92b4-ecb3b76bdd78.mp3'	'gettysburg_address_0b25ab9c_27fd_4626_9520_6aaa5d03bba4_39ffe34f-0732-4f8d-adb9-58a45aacbd9a.txt'