ITIS - 6177 System Integration

FINAI PROJECT

Microsoft Azure Text to Speech API Documentation

Name: Abhishekji Dumala
Student ID: 801307903
Email-ID: Abhishekji Dumala

Introduction:

This is a document which explains the total process of generating the text to speech using Microsoft Azure API key. Deep neural networks are used in this process to generate the texts into the speech format.

Text to Speech Requirements:

This is built by using Microsoft Azure Text to Speech API. For this initially we have to create an account in the Azure portal.
https://azure.microsoft.com/en-us/products/cognitive-services/text-to-speech/

1. **Setup:**

- Creating the speech resource in the azure microsoft account for getting the subscription key and endpoint.
- Installation the code through git clone link
- cd ITIS-6177-Final-Project.
- Create .env file in this directory and include the following data:
    - API_AUTH_URL = https://eastus.api.cognitive.microsoft.com/sts/v1.0/issuetoken
    - API_URL = https://eastus.tts.speech.microsoft.com/cognitiveservices/v1
    - API_KEY = <your-api-key>
    - SERVICE_NAME_NEURAL = (en-US, JessaNeural)
    - VOICE_URL = https://eastus.tts.speech.microsoft.com/congnitiveservices/voices/list
    - VOICE_HOST = eastus.tts.speech.microsoft.com

2. **Installation dependencies:**

- npm install - nodejs package manager.
- node app.js - JavaScript run engine.

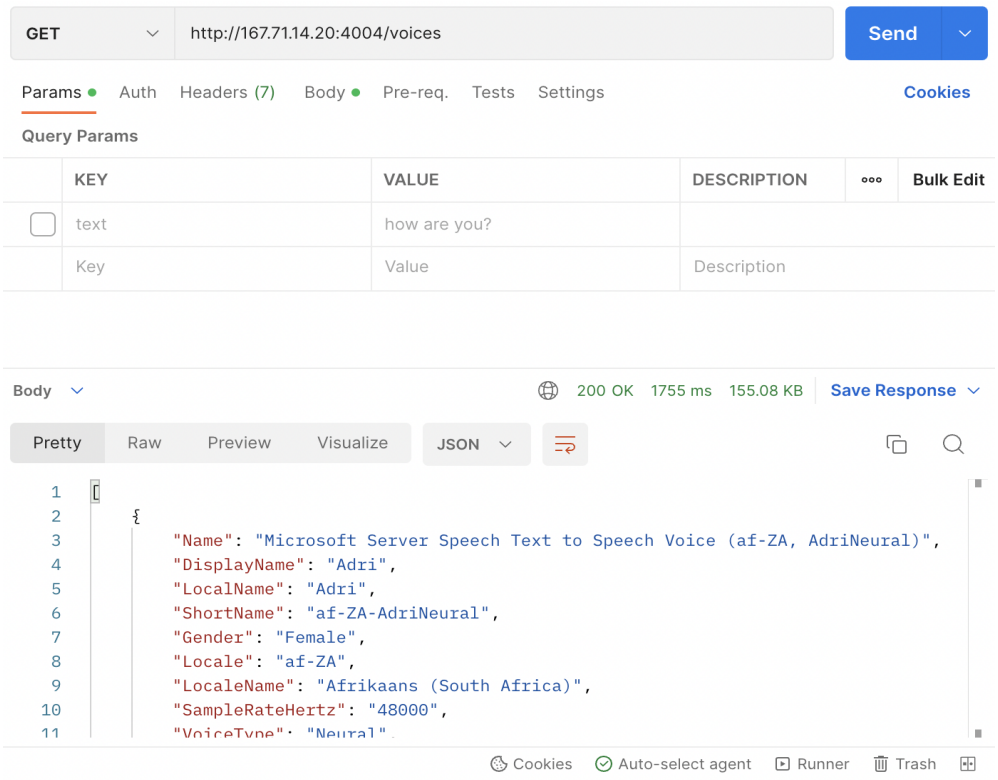
3. **Applications:**
    - Node js for running the js program.
    - Terminal for running the code.
    - Postman for testing purposes.

4. **Testing process:**
    - We can test this the postman workspace api by giving this link: http://167.71.14.20:4004/

    - The main two endpoints for this API are:
        - ➢ /voices
        - ➢ /create

1. **GET Method:** This is the link used for getting the list of voices to be viewed.

    http://167.71.14.20:4004/voices



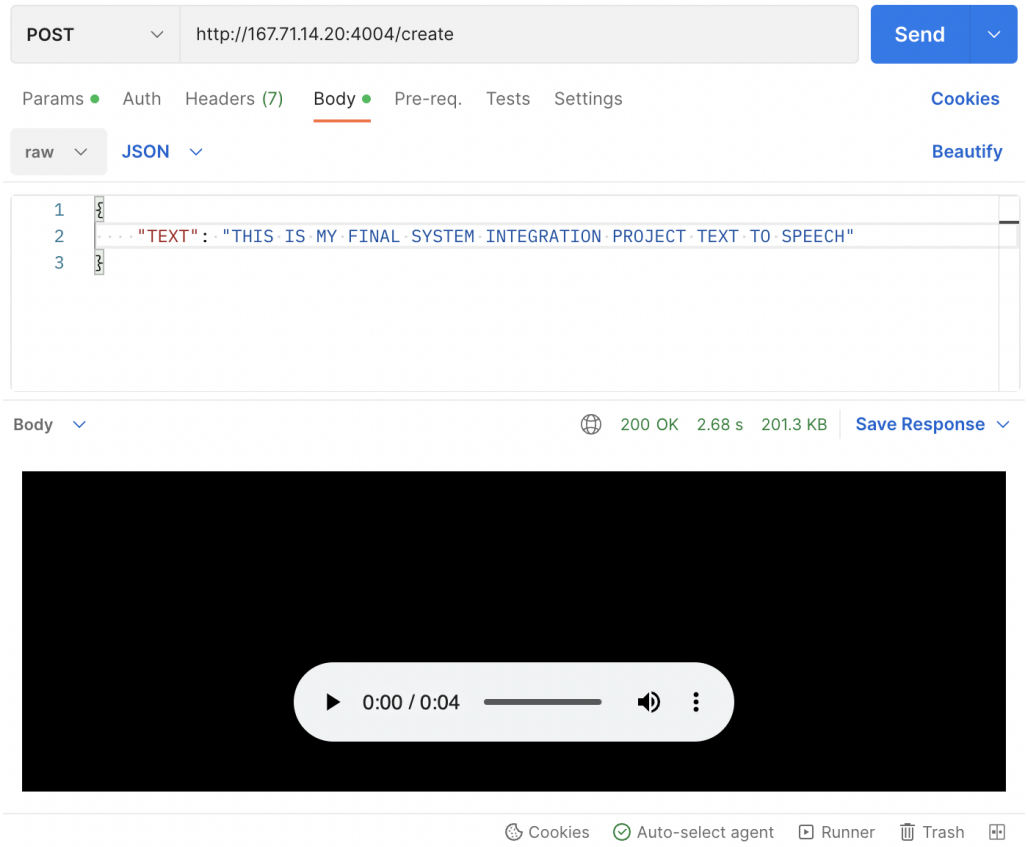
Response:
Code Description
200 Successful
400 Bad data
500 Internal Server Error

2. **POST Method:** This is the link used for posting our request to convert text to speech in json format in Postman application.

http://167.71.14.20:4004/create

POST > Body > raw > JSON

```
{
    “TEXT”: “Enter the text which you want to convert to speech? ”
}
```



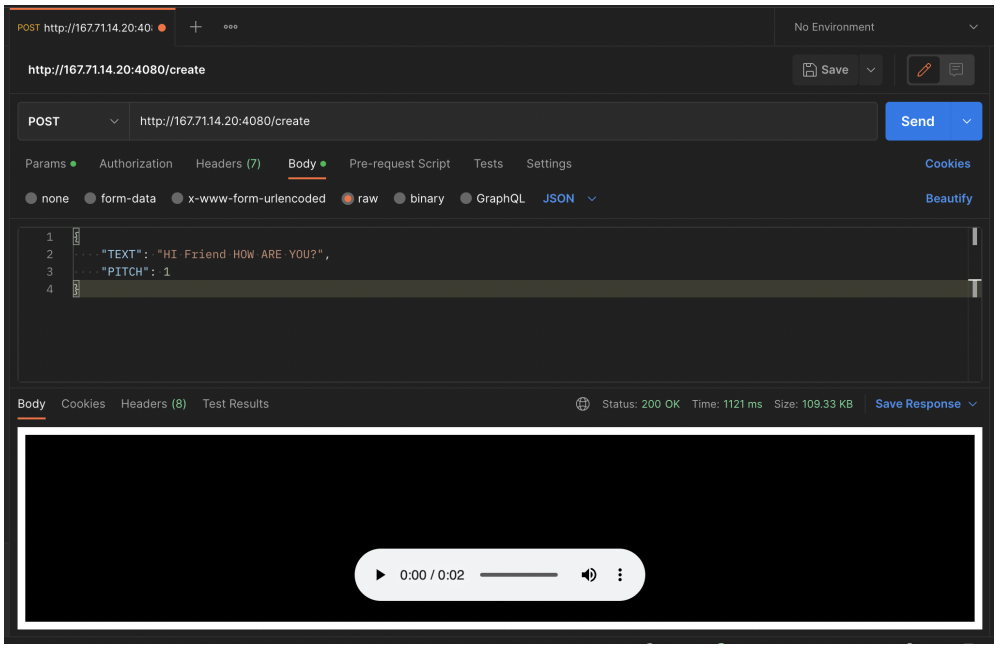
Here Default given voices in the app.js file.

Response:
Code Description
200 Successful
400 Bad data
500 Internal Server Error

PITCH:
In addition to this, there are a few features like Pitch to the voice. Where it takes the pitch value and converts it into accordingly.

POST > Body > raw > JSON

```
{
  “TEXT”: “Enter the text which you want to convert to speech? ”,
  “PITCH”: 1
}
```

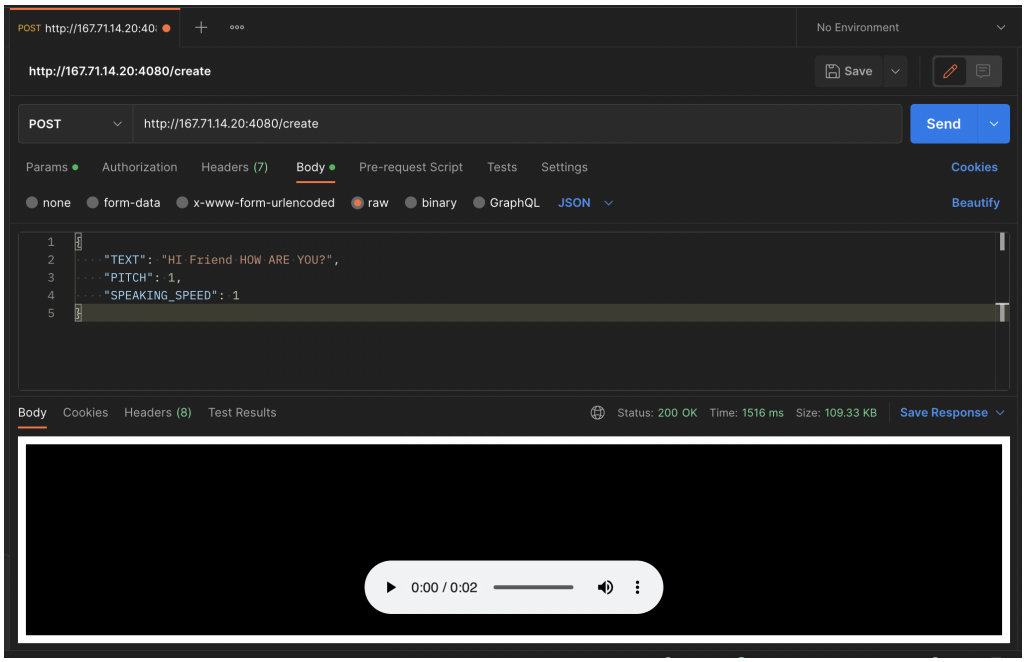


SPEEKING_SPEED:
In addition to this, there are a few features like SPEEKING_SPEED to the voice. Where it takes the speaking speed value and converts the voice speed according to the given value.

POST > Body > raw > JSON

```
{
```

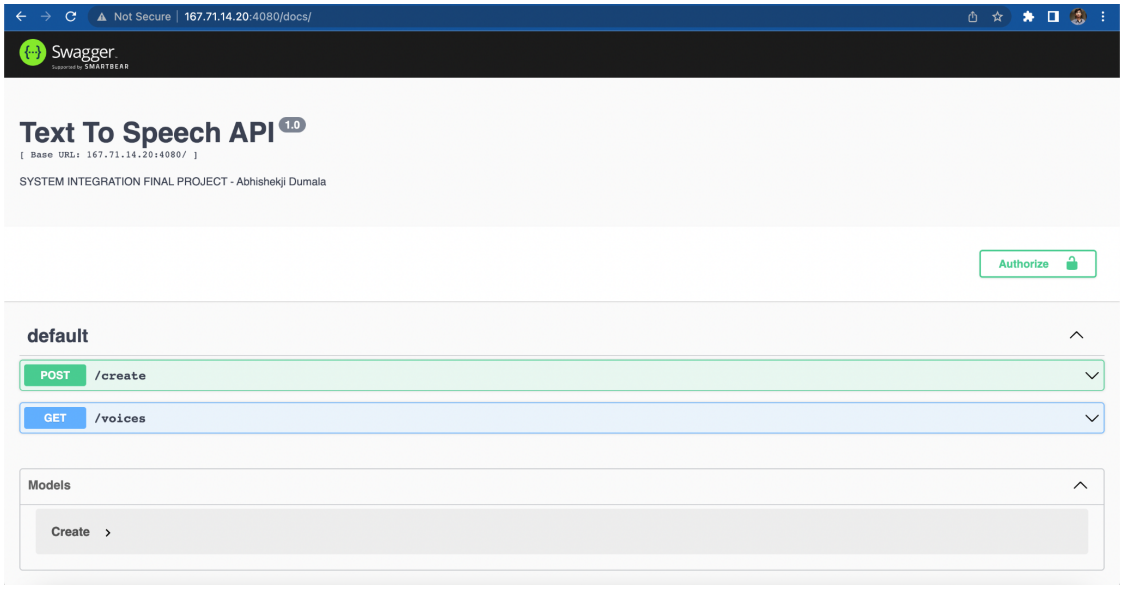
“TEXT”: “Enter the text which you want to convert to speech? ”,
“PITCH”: 1,
“SPEEKING_SPEED”: 1
}



5.Swagger:

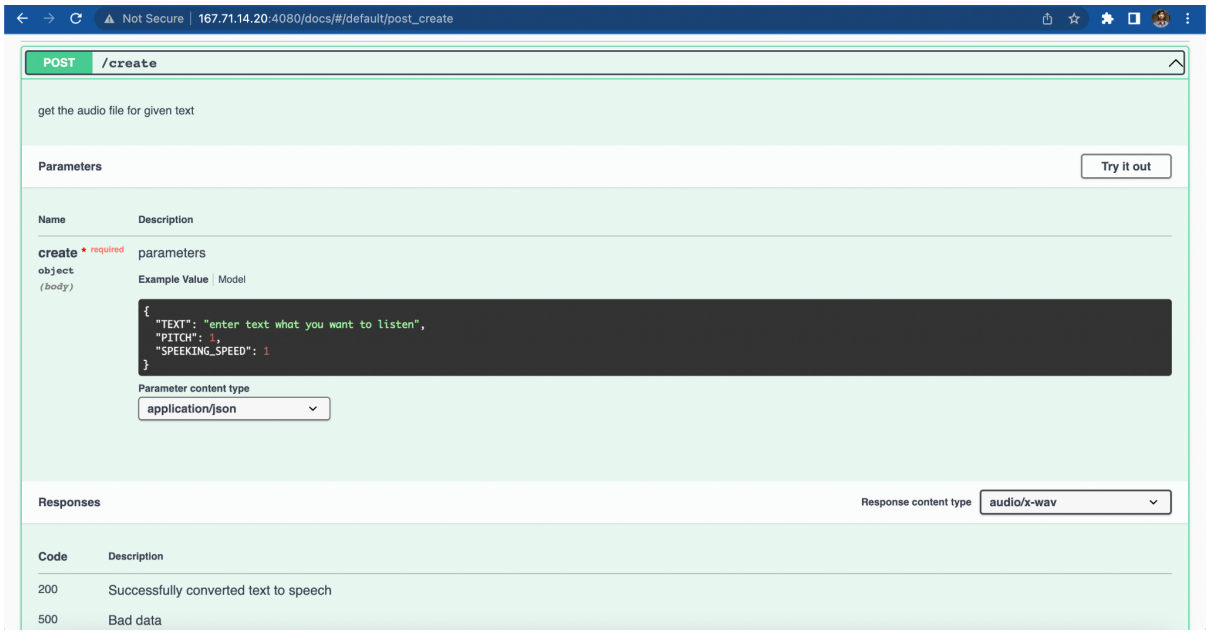
Follow the below link to access swagger application:

http://167.71.14.20:4080/docs/

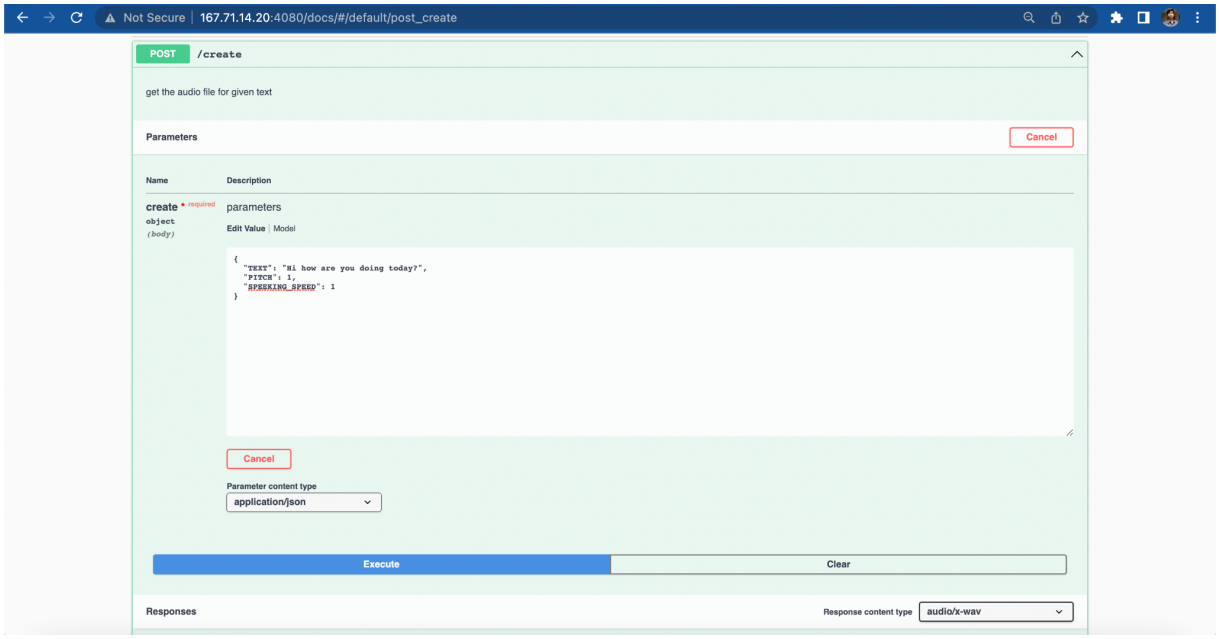


1. **/create**

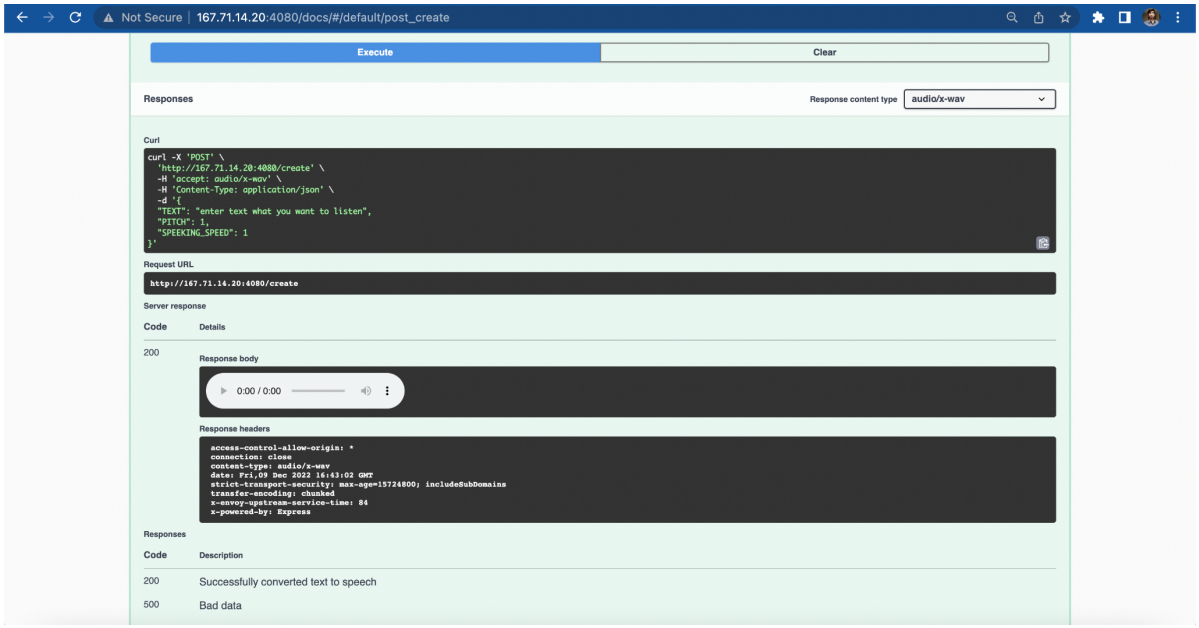
Click on try it out to post the values for getting the text to speech output of all the create.



Enter the text whatever you want to listen to in the speech format.

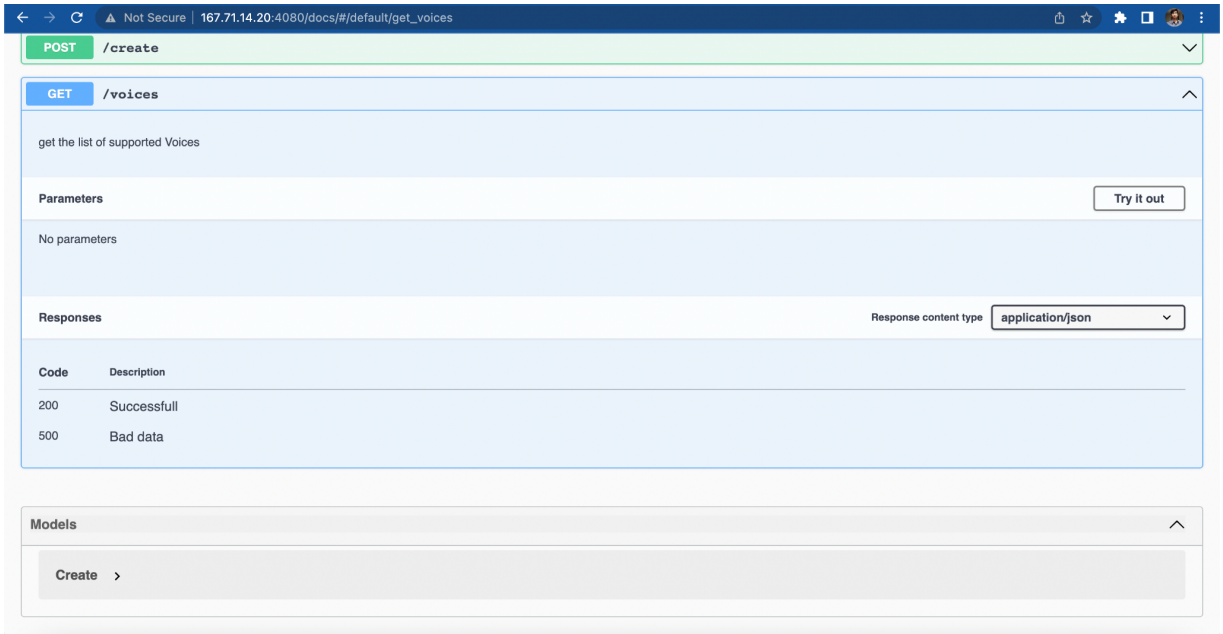


Click on execute to view the result.

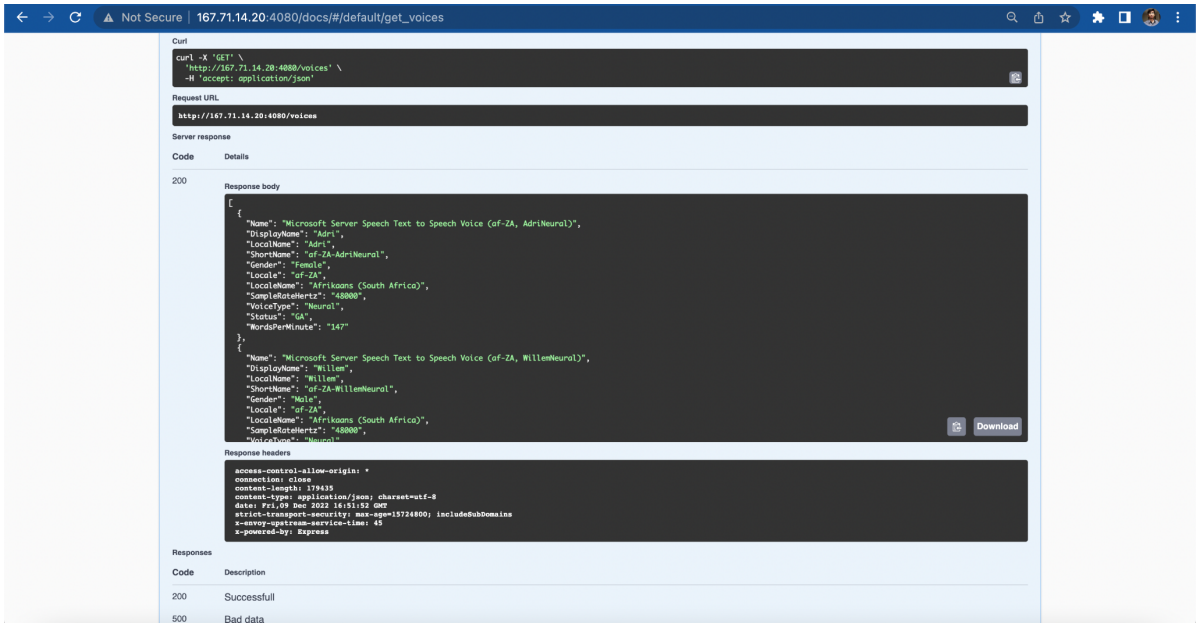


2. /voices

Click on try it out to get the list of values under voice to view which is getting retrieved from the microsoft azure.



Click on Execute to view the list of voices available in this API.



Response:

Code Description

200 Successful

400 Bad data

500 Internal Server Error

References:

https://azure.microsoft.com/en-us/products/cognitive-services/text-to-speech/

https://learn.microsoft.com/en-us/azure/cognitive-services/speech-service/index-text-to-speech

https://learn.microsoft.com/en-us/azure/cognitive-services/speech-service/get-started-text-to-speech