

Google Speech To Text API

1. 문제점

- 기존 방식대로 Power shell로 post 요청을 보내는 방식으로 할 경우 정상적으로 작동은 하는데 아무런 결과가 나오지 않음

```
PS C:\Users\memen> Invoke-WebRequest `
>> -Method POST `
>> -Headers $headers `
>> -ContentType: "application/json; charset=utf-8" `
>> -InFile C:\Users\memen\OneDrive\바탕 화면\인턴 - 무지개연구소\구글\SpeechToText\sync_request.json `
>> -Uri "https://speech.googleapis.com/v1/speech:recognize" | Select-Object -Expand Content
{}
```

1. 문제점

- 기존 방식대로 Power shell로 post 요청을 보내는 방식으로 할 경우 정상적으로 작동은 하는데 아무런 결과가 나오지 않음

```
PS C:\Users\memen> Invoke-WebRequest `
>> -Method POST `
>> -Headers $headers `
>> -ContentType: "application/json; charset=utf-8" `
>> -InFile C:\Users\memen\OneDrive\바탕 화면\인턴 - 무지개연구소\구글\SpeechToText\sync_request.json `
>> -Uri "https://speech.googleapis.com/v1/speech:recognize" | Select-Object -Expand Content
{}
```

1. 문제점

오디오를 Base64로 인코딩 한 후 요청 시 정상 작동하는 듯 보이지만 아무런 결과가 반환되지 않음

TTS변환 음성 및 직접 녹음한 음성 모두 아무런 결과가 나오지 않았음.

본인의 발음은 안 좋으니 이해가 되지만,
TTS를 인식 못하는 점이 이해가 되지 않음.

다음 슬라이드에서 사용한 Google Storage 예제 음성파일을 다운받아 변환하여 사용해 보고 싶었지만, 다운 받을 수 없어서 더 이상 확인해 볼 수 없었음.

[illegible]

```
PS C:\Users\memem> Invoke-WebRequest `
>> -Method POST `
>> -Headers $headers `
>> -ContentType: "application/json; charset=utf-8" `
>> -InFile C:\Users\memem\OneDrive\바탕 화면\인턴 - 무지개연구소\#구글#SpeechToText#sync_request.json `
>> -Uri "https://speech.googleapis.com/v1/speech:recognize" | Select-Object -Expand Content
{}

```

2. 예시 코드 -> Google Storage에 저장 후 URI로 액세스

그래서, Document에 있는 예시대로 Google Storage URI를 이용하기로 함.

예제 URI인 `gs://cloud-samples-tests/speech/brooklyn.flac` 를 그대로 이용하였음.

그랬더니 정상 작동함을 알 수 있다.

결과

음성을 번역한 텍스트와 정확도가 도출된다.

```
sync_request.json
스키마: <선택된 스키마가 없음>
1 {
2   "config": {
3     "encoding": "FLAC",
4     "sampleRateHertz": 16000,
5     "languageCode": "en-US",
6     "enableWordTimeOffsets": false
7   },
8   "audio": {
9     "uri": "gs://cloud-samples-tests/speech/brooklyn.flac"
10  }
11 }
```

```
PS C:\Users\memen\speech> $env:GOOGLE_APPLICATION_CREDENTIALS="C:/Users/memen/google_api_key.json"
PS C:\Users\memen\speech> $cred = gcloud auth application-default print-access-token
PS C:\Users\memen\speech> $headers = @{"Authorization" = "Bearer $cred"}
PS C:\Users\memen\speech> Invoke-WebRequest `
>> -Method POST `
>> -Headers $headers `
>> -ContentType "application/json; charset=utf-8" `
>> -InFile C:\Users\memen\OneDrive\바탕 화면\인턴 - 무지개연구소\구글\SpeechToText\sync_request.json `
>> -Uri "https://speech.googleapis.com/v1/speech:recognize" | Select-Object -Expand Content
{
  "results": [
    {
      "alternatives": [
        {
          "transcript": "how old is the Brooklyn Bridge",
          "confidence": 0.98314303
        }
      ]
    }
  ]
}
```