

Speech to Text Transcription with the Cloud Speech API

허성실(Seongsil Heo)

overview

첫번째 퀵랩과 유사..

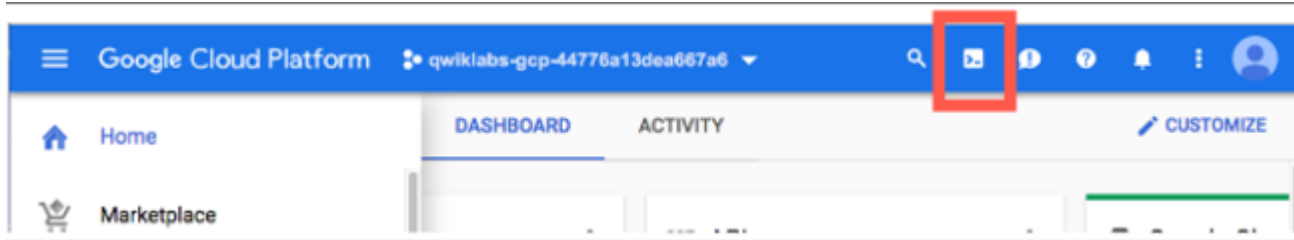
- The Cloud Speech API는 오디오 파일을 80개 넘는 언어로 텍스트화 시킬 수 있다. 이 랩에서는 Cloud Speech API에 오디오 파일을 보내 텍스트화 시키는 작업을 한다.
- 우리가 배울것?
 - Speech API 요청을 생성하고 curl로 API를 부른다.
 - Speech API로 오디오 파일을 다른 언어들로 부른다.
- 필요한것?
 - A Google Cloud Platform Project
 - A Browser such as Chrome or Firefox

Setup and Requirements

- 1. Start lab 버튼 클릭
- 2. 왼쪽에 username 복사해서 Open Google Console을 클릭
- 3. 사용자 계정에 Use Another Account 선택
- 4. 복사한 username 입력, password도 복사 후 입력
- 5. Next 버튼 클릭
- 6. terms and conditions accepts 하기, recovery option은 temporary 계정이니 건너뛰고 free trials을 sign up 하지 않기

The Google Cloud Shell

- 8. 빨간 박스 클릭하여 Google Cloud Shell 프로그램 실행

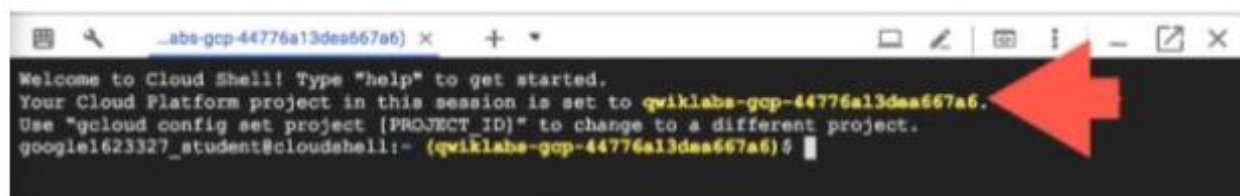


- 9. STRAT CLOUD SHELL 클릭



The Google Cloud Shell

- 10. 본인의 PROJECT_ID로 프로젝트가 생성될 것이다.

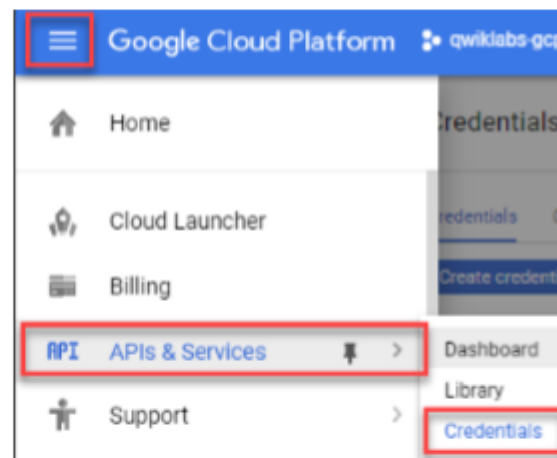


The screenshot shows a terminal window titled "...abs-gcp-44776a13dea667a6". The terminal text reads: "Welcome to Cloud Shell! Type 'help' to get started. Your Cloud Platform project in this session is set to **qwiklabs-gcp-44776a13dea667a6**. Use 'gcloud config set project [PROJECT ID]' to change to a different project. google1623327_student@cloudshell:~ (qwiklabs-gcp-44776a13dea667a6) \$". A red arrow points to the project ID "qwiklabs-gcp-44776a13dea667a6".

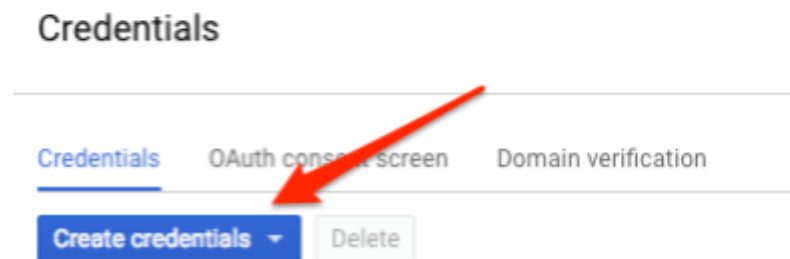
- 11. 다음의 명령어로 활성화된 계정 이름들을 볼 수 있다.
gcloud auth list
- 12. Project ID들을 볼 수 있다.
gcloud config list project

Create an API Key

- 13. Speech API의 요청을 전송하기 위해 curl을 이용할 것이기 때문에, URL 요청을 받기 위해서는 API 키를 생성해야 한다. 본인의 PROJECT_ID로 프로젝트가 생성될 것이다.
- 14. Navigation menu -> APIs & Services -> Credentials

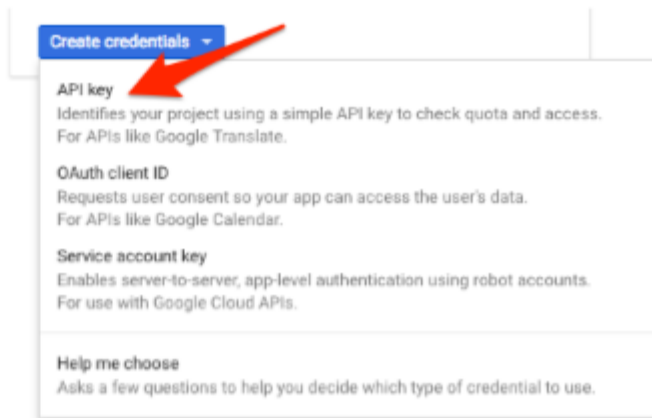


- 15. Create credentials 클릭



Create an API Key

- 16. 드롭다운 메뉴에서 API key 선택 -> 생성된 키 복사



- 17. 환경변수 설정해주기, 복사한기를 <YOUR_API_KEY> 부분에 붙여넣기

```
export API_KEY=<YOUR_API_KEY>
```

Create your Speech API

- 18. “gs://cloud-samples-tests/speech/brooklyn.flac” 다음의 링크는 텍스트로 변환할 예제 파일임
- 19. Google Cloud Shell에서 “request.json” 파일 생성
- 20. 선호하는 커맨드 라인 에디터를 이용하여 request.json 파일을 연다.(nano, vim, emacs, glolud) (저는 vim request.json명령어 사용)
- 21. 아래의 내용을 채워 넣는다.

```
touch request.json
```

```
{
  "config": {
    "encoding": "FLAC",
    "languageCode": "en-US"
  },
  "audio": {
    "uri": "gs://cloud-samples-tests/speech/brooklyn.flac"
  }
}
```

-config- 요청 처리 환경 설정
Encoding: API로 보낼 때 쓰이는 오디오 파일 형식
Sample rate: API로 보낼 때 오디오 데이터의 Herz 정보
Language code: 변환 대상에서 쓰인 언어
- audio- 변환할 오디오 파일 url

Call the Speech API

- 22. Google Cloud Shell 에서 다음의 curl 명령어를 입력한다.

```
curl -s -X POST -H "Content-Type: application/json" --data-binary  
@request.json \  
"https://speech.googleapis.com/v1/speech:recognize?key=${API_KEY}"
```

- 23. 요청 직후의 반응은 다음과 같다.

```
{  
  "results": [  
    {  
      "alternatives": [  
        {  
          "transcript": "how old is the Brooklyn Bridge",  
          "confidence": 0.98267895  
        }  
      ]  
    }  
  ]  
}
```

Transcript: Text로 변환한 결과
Confidence: 정확도

Speech to text transcription in different languages

- 24. “gs://speech-language-samples” 다음의 링크는 텍스트로 변환할 예제 파일임
- 25. Google Cloud Shell에서 “request.json” 이라는 파일 생성 (직전 파일 삭제 “rm request.json”)
`touch request.json`
- 26. 선호하는 커맨드 라인 에디터를 이용하여 request.json 파일을 연다.(nano, vim, emacs, glolud) (저는 vim request.json 명령어 사용)
- 27. 아래의 내용을 채워 넣는다.

```
{
  "config": {
    "encoding": "FLAC",
    "languageCode": "fr"
  },
  "audio": {
    "uri": "gs://speech-language-samples/fr-sample.flac"
  }
}
```

Speech to text transcription in different languages

- 28. 요청 직후의 반응은 다음과 같다.

```
{
  "results": [
    {
      "alternatives": [
        {
          "transcript": "maître corbeau sur un arbre perché tenait en son  
bec un fromage",
          "confidence": 0.9710122
        }
      ]
    }
  ]
}
```