

Cloud Natural Language API: Qwik Start

허성실(Seongsil Heo)

Overview

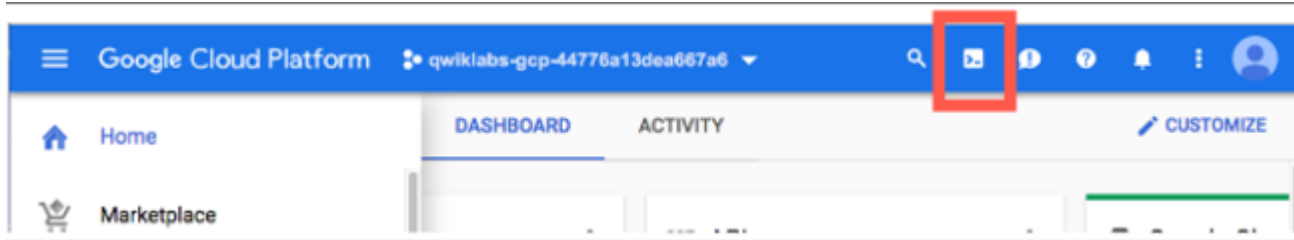
- Google Cloud Natural Language API는 텍스트 문서에 언급된 사람, 장소, 이벤트 등에 관한 정보를 추출해준다. 소셜 미디어에서의 제품에 대한 정서를 파악하거나, 콜센터에서 고객 대화의 의도를 분석할 수 있다. 분석을 위한 텍스트 문서를 업로드 할 수도 있다. 주요 기능은 다음과 같다.
 - Syntax Analysis : 토큰 및 문장을 추출하고, 품사(PoS)를 식별하고, 각각의 문장들간의 관계를 나타내는 트리를 생성한다.
 - Entity Recognition : 엔티티를 식별하고 사람, 조직, 위치, 이벤트, 제품 및 미디어와 같은 타입별로 레이블을 지정한다.
 - Sentiment Analysis : 텍스트 블록에서 표현된 전반적인 감정을 이해한다.
 - Content Classification: 사전 정의된 700개 이상의 카테고리로 문서를 분류한다.
 - Multi-Language : 다중 언어 지원
 - Integrated REST API: REST API 통합 지원
- 우리가 할 일
 - 텍스트로부터 엔티티(사람, 장소, 이벤트) 추출

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

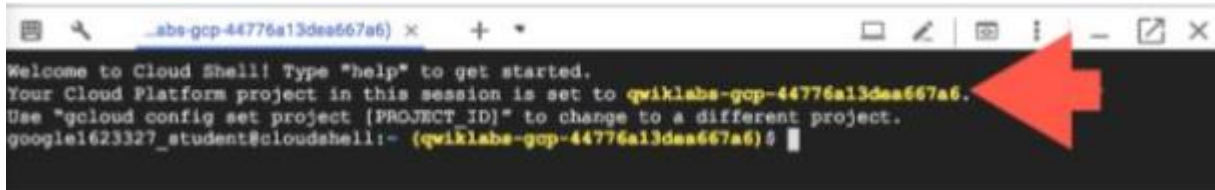


- 9. STRAT CLOUD SHELL 클릭



Create an API Key

- 10. 본인의 PROJECT_ID로 프로젝트가 생성 될 것이다.



```
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) $
```

- 11. 다음의 명령어로 활성화된 계정 이름들을 볼 수 있다.

`gcloud auth list`

- 12. Project ID들을 볼 수 있다.

`gcloud config list project`

- 13. 현재 PROJECT_ID를 불러온 값을 GOOGLE_CLOUD_PROJECT 환경변수로 등록한다.

```
export GOOGLE_CLOUD_PROJECT=$(gcloud config get-value core/project)
```

- 14. NL API 접근을 위한 새로운 서비스 계정을 생성한다.

```
gcloud iam service-accounts create my-natlang-sa \
  --display-name "my natural language service account"
```

Service Account 내부 이름

Service Account 외부 이름

Create an API Key

- 15. 새로 만들어진 IAM Service Account를 통해 API 키를 생성하고, 이를 JSON file로 저장한다.

```
gcloud iam service-accounts keys create ~/key.json \  
  --iam-account my-natlang-  
sa@${GOOGLE_CLOUD_PROJECT}.iam.gserviceaccount.com
```

- 16. 생성된 API 키를 환경 변수에 등록한다.

```
export GOOGLE_APPLICATION_CREDENTIALS="~/key.json"
```

JSON 파일이 있는 디렉토리 위치
입력, "pwd" 명령으로 확인 가능

Make an Entity Analysis Request

- 17. 아래의 gcloud 명령 수행

```
gcloud ml language analyze-entities --content="Michelangelo Caravaggio,  
Italian painter, is known for 'The Calling of Saint Matthew'."
```

NL API의 어떤 기능을 쓸 것인지?
여기서는 개체 분석

나머지는 분석할 텍스트 입력!
Or
--content-file={file name} 으로
파일을 분석할 수 있음

NL API요청을 수행하겠다.

- 18. 결과 정보가
출력된다.

- 19. End Lab 버튼
클릭

```
{  
  "entities": [  
    {  
      "name": "Michelangelo Caravaggio",  
      "type": "PERSON",  
      "metadata": {  
        "wikipedia_url": "http://en.wikipedia.org/wiki/Caravaggio",  
        "mid": "/m/020bg"  
      },  
      "salience": 0.83047235,  
      "mentions": [  
        {  
          "text": {  
            "content": "Michelangelo Caravaggio",  
            "beginOffset": 0  
          },  
          "type": "PROPER"  
        },  
        {  
          "text": {  
            "content": "painter",  
            "beginOffset": 33  
          },  
          "type": "COMMON"  
        }  
      ]  
    },  
    {  
      "name": "Italian",  
      "type": "LOCATION",  
      "metadata": {  
        "wikipedia_url": "http://en.wikipedia.org/wiki/Italy",  
        "mid": "/m/026bm"  
      },  
      "salience": 0.16952765,  
      "mentions": [  
        {  
          "text": {  
            "content": "Italian painter",  
            "beginOffset": 14  
          },  
          "type": "PROPER"  
        },  
        {  
          "text": {  
            "content": "Italian",  
            "beginOffset": 33  
          },  
          "type": "COMMON"  
        }  
      ]  
    }  
  ]  
}
```

Name: 개체 이름
Type: 말그대로 타입
Metadata: 연관된 위키피디아 정보
나타냄
Salience: 전체 텍스트에서 얼마나
중요한 비중을 차지하나
Mentions: 전체 텍스트에서, 이
개체와 동일한 개체?