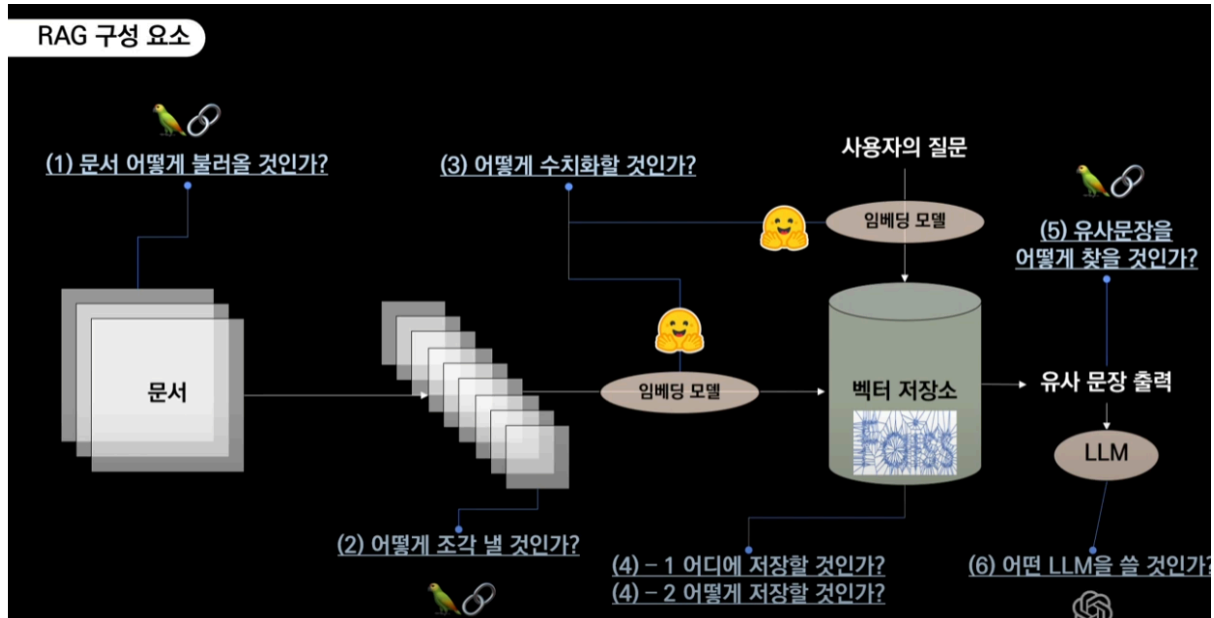


스트림잇 프로젝트 배포하기

<https://github.com/kmdadoo/langchain.git> GhatAPI 활용 샘플



배포를 위해서는 `requirements.txt` 파일이 필요함.

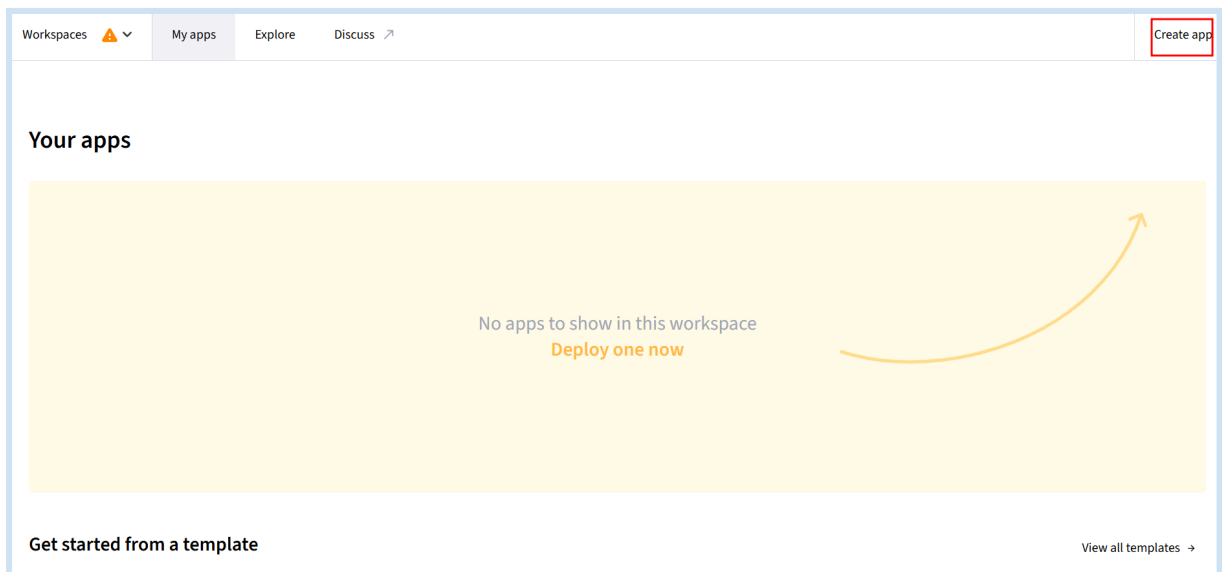
스크립터에서 정의한 여러가지 라이브러리들이 스트림 클라우드 PC 상에도 다운이 되도록 만들어주는 파일.

이 파일에 있는 라이브러리들을 자동으로 **PIP** 인스톨 해줘서 우리가 만든 스크립트를 무리없이 구동하게 해줌.

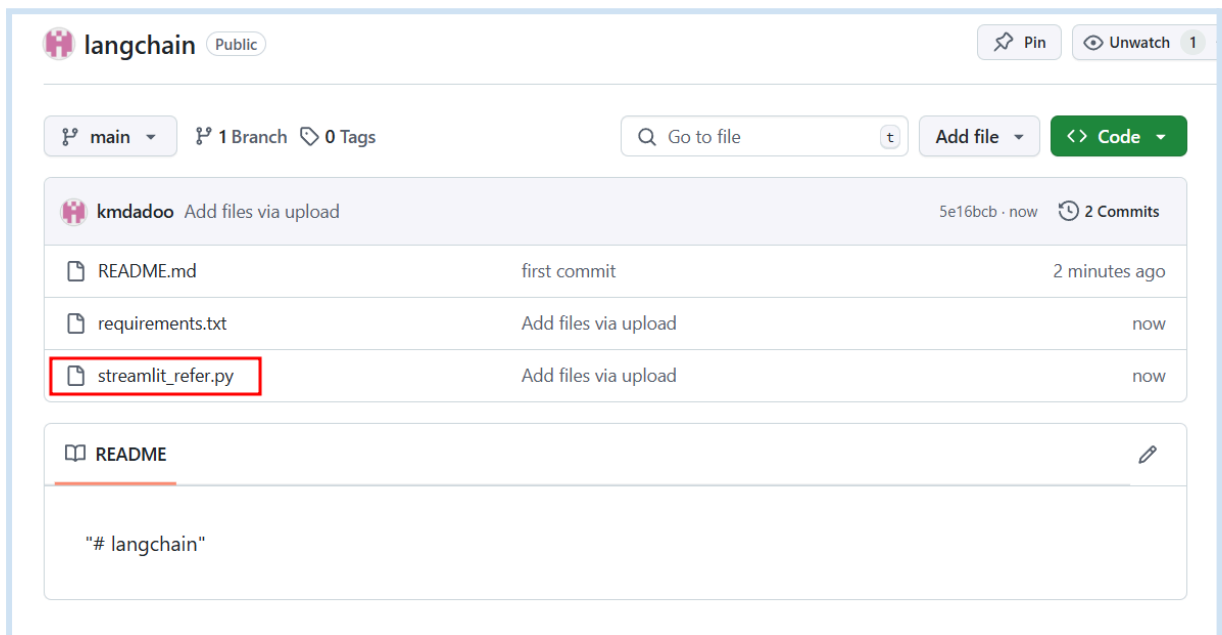
`requirements.txt`

```
pysqlite3-binary  
langchain  
chromadb  
unstructured  
sentence-transformers  
faiss-cpu  
tiktoken  
openai  
pypdf  
loguru  
docx2txt
```

이 두파일이 준비가 다되면 share.streamlit.io 에 들어가서 회원가입하면 다음 화면이 나옴.



먼저 깃허브에 파일을 업로드 한다.



그럼 다음 create app 을누르면 깃로그인을 하면

Deploy an app

Repository ?

[Paste GitHub URL](#)

kmdadoo/langchain

Branch

main

Main file path

streamlit_refer.py

App URL (optional)

langchain-8aq9jzmioew6pn5yuoiz4u

.streamlit.app

Domain is available

[Advanced settings](#)

Deploy

Branch가 안나오면 cmd 창dp

```
c:/>git commit --allow-empty -m "Empty commit"
```

입력하고 다시들어가면 보인다. URL 은 기본 값을 사용.

Advanced settings 은 API키를 등록하면 다른 사람은 보지 못하고 개발한 사람만 볼수 있게 변수를 설정할 수있음.

Advanced settings

Python version

3.12

Secrets

Provide environment variables and other secrets to your app using [TOML](#) format. This information is encrypted and served securely to your app at runtime. [Learn more about Secrets in our docs](#). Changes take around a minute to propagate.

```
OPENAI_API_KEY = 'YOUR_API_KEY'
```

Save

이렇게 하면 APP이 등록 되고 배포를 사용할 수 있다.

그럼 아까 받았던 URI 로 들어가면 볼수 있다.

<https://langchain-8aq9jzmioew6pn5yuoiz4u.streamlit.app>

Upload your file

Drag and drop files here
Limit 200MB per file • PDF, DOCX

Browse files

IROS_366_수입식품안...
342.9KB

BOK 이슈노트_시 알고...
8.5MB

BOK 이슈노트_시 알고...
8.5MB

OpenAI API Key

Process

- 불규칙한 문법: 한글은 영어와는 다른 문법 규칙을 가지고 있습니다. 따라서 한글 텍스트를 처리할 때는 이러한 불규칙성을 고려하여 처리해야 합니다.
- 데이터의 특성: 한글 텍스트 데이터는 오타, 줄임말, 인터넷 신조어 등의 노이즈가 많이 포함되어 있을 수 있습니다. 이러한 노이즈를 처리하고 정제하는 과정이 필요합니다.
- 자원의 부족: 한글 자연어 처리에 필요한 자원이 영어에 비해 상대적으로 부족할 수 있습니다. 따라서 한글 텍스트 분석을 위한 모델과 도구의 개발이 더 어려울 수 있습니다.

이러한 어려움들을 극복하기 위해서는 한글 자연어 처리 기술의 발전과 함께 충분한 데이터와 자원의 확보, 효과적인 전처리 기법 등이 필요합니다.

gpt를 활용한 텍스트 데이터 분석은 어떤?

GPT(Generative Pre-trained Transformer)는 거대한 인공신경망 모델로, 대용량 텍스트 데이터의 문장 패턴을 학습시킵니다. GPT는 텍스트 데이터를 입력으로 받아 다양한 목적에 맞게 활용할 수 있습니다. 예를 들어, 감성분석이나 문장 생성과 같은 분야에서 상업화가 가능한 수준까지 발전하고 있습니다. GPT를 활용한 텍스트 데이터 분석은 텍스트 데이터를 입력으로 주고, 모델이 해당 데이터의 패턴을 학습하여 다양한 분석 결과를 도출하는 방식으로 이루어집니다.

질문을 입력해주세요.

파일을 먼저 업로드하면 우리가 질문한 것에대한 답변의 나오는 것을 알 수 있다.