# 목표

* 사용자의 자연어 질의에 해당하는 youtube 영상의 구간들을 찾아 줌
* 각 Video별, 구간별 설명 추가, 자동 영상 편집
* 하나의 mp4 파일로 제작
* 해당 영상 구간들을 선택적으로 play 할 수 있도록 Web에 개시 (미구현)

# 활용 기술

* Langchain
* Langsmith
* Ffmpeg, OpenCV
* TTS (pyttsx3)

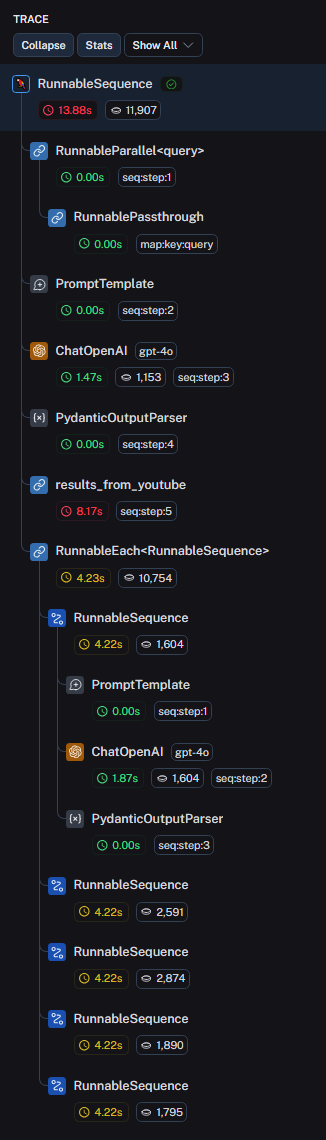
# 기능 구성

1. **Load Environment Variables**
   * **설명**: 환경 변수를 로드하여 API 키와 같은 중요한 설정을 관리합니다. 이 단계에서는 .env 파일에서 설정을 불러옵니다.
2. **Create LLM Object**
   * **설명**: 언어 모델 객체를 생성합니다. 이 객체는 후속 작업에서 텍스트를 생성하거나 분석하는 데 사용됩니다.
3. **Extract YouTube Query and LLM Prompt from User Request - First Chain**
   * **설명**: 사용자의 요청을 분석하여 YouTube 검색어와 LLM 프롬프트를 추출합니다. 첫 번째 체인을 통해 이 작업이 수행됩니다.
4. **Define Function to Retrieve Video Metadata and Captions from YouTube**
   * **설명**: YouTube API를 사용하여 동영상 메타데이터와 자막을 가져오는 함수를 정의합니다. 검색 결과에서 자막이 포함된 동영상의 메타 정보를 수집합니다.
5. **Analyze YouTube Captions to Extract User Request Content - Second Chain**
   * **설명**: 수집된 자막을 분석하여 사용자가 원하는 내용을 추출합니다. 두 번째 체인을 통해 이 작업이 수행됩니다.
6. **Create Full Chain - Combine First and Second Chains**
   * **설명**: 첫 번째 체인과 두 번째 체인을 결합하여 전체 체인을 생성합니다. 이 체인은 사용자 요청에서부터 최종 결과를 얻는 과정을 포함합니다.
7. **Execute Full Chain**
   * **설명**: 전체 체인을 실행하여 사용자 요청에 대한 결과를 생성합니다. 이 단계에서는 모든 과정이 연속적으로 수행됩니다.
8. **Produce Videos**
   * **설명**: 최종 동영상을 생성합니다. YouTube에서 다운로드한 동영상과 생성된 자막 및 설명을 사용하여 비디오를 제작합니다.

### **Produce Videos** 의 세부 단계 설명

* **Download YouTube Content and Create MP4 Files**
  + **설명**: YouTube에서 동영상을 다운로드하여 MP4 파일로 저장합니다.
* **Slice Videos to Create Segment Video Files**
  + **설명**: 동영상을 잘라 세그먼트 비디오 파일을 생성합니다. 각 세그먼트는 특정 내용을 포함하는 구간입니다.
* **Create Description Videos for Each YouTube Content**
  + **설명**: 각 YouTube 콘텐츠에 대한 설명 비디오를 생성합니다. 이 단계에서는 비디오에 텍스트를 추가하고 TTS를 사용하여 음성이 추가됩니다.
* **Create Explanation Videos for Each Segment and Insert Text in Videos**
  + **설명**: 각 세그먼트에 대한 설명 비디오를 생성하고 비디오에 텍스트를 삽입합니다.
* **Concatenate Segment Videos**
  + **설명**: 생성된 세그먼트 비디오들을 하나의 비디오로 이어 붙입니다. 최종 결과물은 모든 세그먼트가 연결된 하나의 연속적인 비디오입니다.

<LangSmith Trace>



# 프롬프트 예시

1. Query 생성 Prompt (첫번째 chain의 prompt)

"""

System :

너는 사용자가 원하는 youtube 컨텐츠를 찾아주는 AI야.

사용자의 query를 분석해서 다음 3가지 정보를 생성해줘.

 - 사용자가 원하는 YouTube 컨텐츠를 찾을 수 있도록 YouTube 검색어를 만들어줘.

 - 찾아야 할 컨텐츠가 한글인 경우 'ko' 영어인 경우 'en'이라고 표시해줘.

 - YouTube 컨텐츠가 담고있는 Text 정보에서 사용자가 원하는 부분을 찾기 위한 자연어 질의를 만들어줘

   예를 들어, 채널명이나 필터 정보 등 YouTube 검색에 사용되는 내용은 제외해야되.

검색어를 만들때 tips 를 참고해.

tips :

{tips}

query :

{query}

Format :

{format}

"""

1. Youtube 클립 추출 prompt (두번째 chain의 prompt)

"""

System :

너는 동영상 편집자이다.

먼저 captions 내용을 100자 이내로 요약해.

요약 결과는 네가 시청자에게 동영상을 소개하는 말투여야 해.

그리고 captions에서 사용자 query를 잘 설명하고 있는 내용을 뽑아내고

그 내용이 포함된 구간을 찾아내.

여기서 구간이란 captions 내의 하나의 세부 구간이 아니라 연속되는 여러개의 세부 구간의 합을 의미한다.

예)

세부 구간 1 :

  start : 112.0, duration : 5.0

세부 구간 2 :

  start : 117.0, duration : 4.0

세부 구간 3 :

  start : 121.0, duration : 6.0

세부 구간 4 :

  start : 127.0, duration : 3.0

네가 찾은 하나의 구간이 세부 구간 1 ~ 4 인 경우

구간의 start, duration은 다음과 같이 계산 하면 된다.

  start : 112.0 (세부 구간 1 의 start)

  duration : 18.0 ( 5.0 + 4.0 + 6.0 + 3.0, 세부 구간 4개 duration 의 합계)

해당 구간의 내용을 20자 이내로 요약해서 구간 제목을 만들어.

그 구간은 없을 수도 있고 여러개 일수도 있다.

그 구간이 4개 이상이라면 가장 적합한 것 3개만 찾아내라.

video\_id 값은 동영상을 식별하는 id이다. 변경하지 말고 그대로 출력하라.

channel\_name은 동영상을 소유한 채널 이름이다. 변경하지 말고 그대로 출력하라.

query :

{query}

video\_id :

{video\_id}

channel\_name :

{channel\_name}

captions :

{captions}

Format :

{format}

"""

1. 사용자 입력 쿼리

input\_query = """

한국인 중에 기네스북에 오른 사람들은 누가 있으면 어떤 분야에서 올랐는지 알려줘

"""

# 실행 결과

1. 사용자 입력 쿼리

input\_query = """

한국인 중에 기네스북에 오른 사람들은 누가 있으면 어떤 분야에서 올랐는지 알려줘

"""

1. 첫번째 Chain의 completion 결과

# youtube\_query : YouTube 검색용 query

# llm\_query : LLM용 query (자막 내용으로 In-context learning방식)

# language : YouTube 검색 시 자막의 언어

{

  "output": {

    "youtube\_query": "한국인 기네스북 기록",

    "llm\_query": "한국인 중 기네스북에 오른 사람들과 그들이 어떤 분야에서 기록을 세웠는지",

    "language": "ko"

  }

}

1. Youtube 컨텐츠 추출정보 (Execute Full Chain 수행 결과)

[

 Summary(video\_id='Je0sxeRPD58', channel\_name='MBCNEWS', description='기네스북에 오른 턱수염 장신구 기록 보유자 소개', segments=[Segment(title='턱수염 장신구 기네스 기록', start=4.034, duration=43.22)]),

 Summary(video\_id='5ctys-Bgjow', channel\_name='SBS 뉴스', description='기네스북에 오른 다양한 기록들과 그 주인공들을 소개합니다.', segments=[Segment(title='기네스 세계 기록 발표', start=10.765, duration=7.974), Segment(title='진귀한 기록들', start=27.547, duration=10.011), Segment(title='기네스북에 오르는 이유', start=52.806, duration=15.554)]),

 Summary(video\_id='jx70ld3F7GA', channel\_name='SBS 뉴스', description='브라질 서퍼 마야 가베이라가 22.4m 파도로 기네스 신기록을 세운 이야기', segments=[Segment(title='브라질 서퍼의 기네스 신기록', start=0.1, duration=67.968)]),

 Summary(video\_id='IlGqIab-tRA', channel\_name='SBS Entertainment', description='한국 대표가 기네스 기록에 도전하는 모습을 담은 영상입니다.', segments=[Segment(title='기네스 기록 도전 소개', start=0.12, duration=13.76), Segment(title='기네스 기록 도전 시작', start=32.32, duration=16.362), Segment(title='기네스 기록 도전 성공', start=101.973, duration=12.968)]),

 Summary(video\_id='RLhoTII5enk', channel\_name='SBS 뉴스', description='기네스북에 오른 한국인과 그들의 기록을 소개합니다.', segments=[])

]

1. Segment 화된 video 파일들(Create Explanation Videos for Each Segment and Insert Text in Videos 단계 수행 결과)

# <UUID>\_<video\_id>\_c.mp4 : video 표지 영상

# <UUID>\_<video\_id>\_n\_c.mp4 : video n번 segment 표지 영상

# <UUID>\_<video\_id>\_n\_t.mp4 : video n번 segment 영상

[

 '3593b521-fdcc-42c1-9040-f6c9cda481cd\_Je0sxeRPD58\_c.mp4',

 '3593b521-fdcc-42c1-9040-f6c9cda481cd\_Je0sxeRPD58\_0\_c.mp4',

 '3593b521-fdcc-42c1-9040-f6c9cda481cd\_Je0sxeRPD58\_0\_t.mp4',

 '3593b521-fdcc-42c1-9040-f6c9cda481cd\_5ctys-Bgjow\_c.mp4',

 '3593b521-fdcc-42c1-9040-f6c9cda481cd\_5ctys-Bgjow\_0\_c.mp4',

 '3593b521-fdcc-42c1-9040-f6c9cda481cd\_5ctys-Bgjow\_0\_t.mp4',

 '3593b521-fdcc-42c1-9040-f6c9cda481cd\_5ctys-Bgjow\_1\_c.mp4',

 '3593b521-fdcc-42c1-9040-f6c9cda481cd\_5ctys-Bgjow\_1\_t.mp4',

 '3593b521-fdcc-42c1-9040-f6c9cda481cd\_5ctys-Bgjow\_2\_c.mp4',

 '3593b521-fdcc-42c1-9040-f6c9cda481cd\_5ctys-Bgjow\_2\_t.mp4',

 '3593b521-fdcc-42c1-9040-f6c9cda481cd\_jx70ld3F7GA\_c.mp4',

 '3593b521-fdcc-42c1-9040-f6c9cda481cd\_jx70ld3F7GA\_0\_c.mp4',

 '3593b521-fdcc-42c1-9040-f6c9cda481cd\_jx70ld3F7GA\_0\_t.mp4',

 '3593b521-fdcc-42c1-9040-f6c9cda481cd\_IlGqIab-tRA\_c.mp4',

 '3593b521-fdcc-42c1-9040-f6c9cda481cd\_IlGqIab-tRA\_0\_c.mp4',

 '3593b521-fdcc-42c1-9040-f6c9cda481cd\_IlGqIab-tRA\_0\_t.mp4',

 '3593b521-fdcc-42c1-9040-f6c9cda481cd\_IlGqIab-tRA\_1\_c.mp4',

 '3593b521-fdcc-42c1-9040-f6c9cda481cd\_IlGqIab-tRA\_1\_t.mp4',

 '3593b521-fdcc-42c1-9040-f6c9cda481cd\_IlGqIab-tRA\_2\_c.mp4',

 '3593b521-fdcc-42c1-9040-f6c9cda481cd\_IlGqIab-tRA\_2\_t.mp4'

 ]

1. 하나의 Video 파일로 이어 붙이기

기네스북\_한국인.mp4

* 1. Video 컨텐츠 별 간지 영상 예



* 1. segment 별 간지 영상 예



* 1. segment 영상 Capture

