
KnowledgeGraph와 Large Language Model

허 홍 수



주제

- 지식그래프에 대한 개요
- LLM과 지식그래프

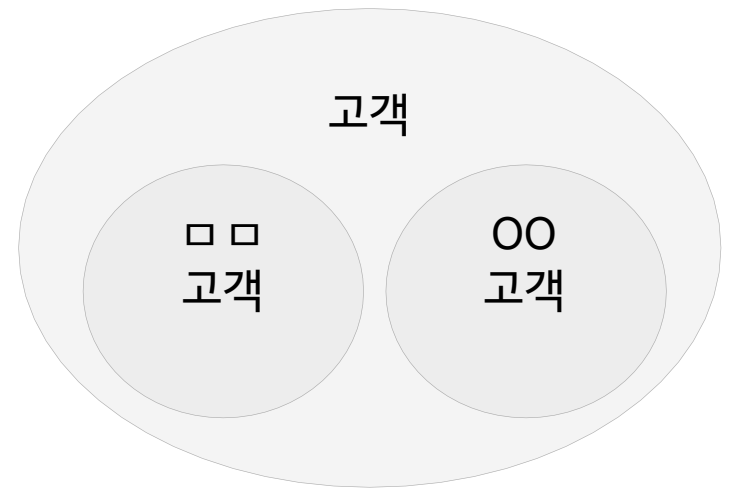
지식그래프? WHY?



출처: <https://namu.wiki/w/바벨탑>



출처: <https://pxhere.com/ko/photo/540923>



지식그래프

그래프 형태의 지식베이스

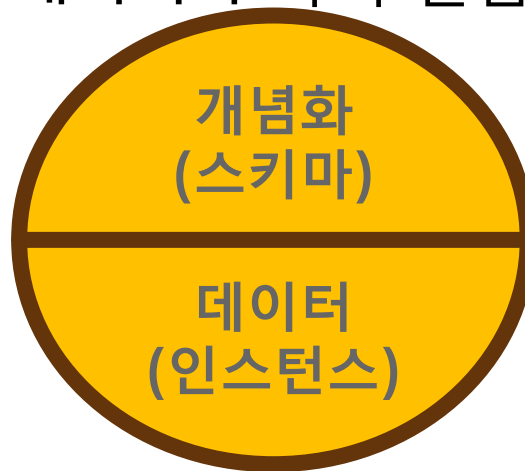
데이터 표현 형태

데이터간의
연결관계를 표현

지식의 체계화

데이터에 대한 개념과
상호간의 관련된 정보를
명시적으로 표현

데이터와 지식 결합



지식그래프 구성 기술

- 지식그래프를 구성하기 위해 사용하는 대부분의 기술과 표준은
- 시맨틱웹을 통해 만들어진 기술과 표준을 활용

URI



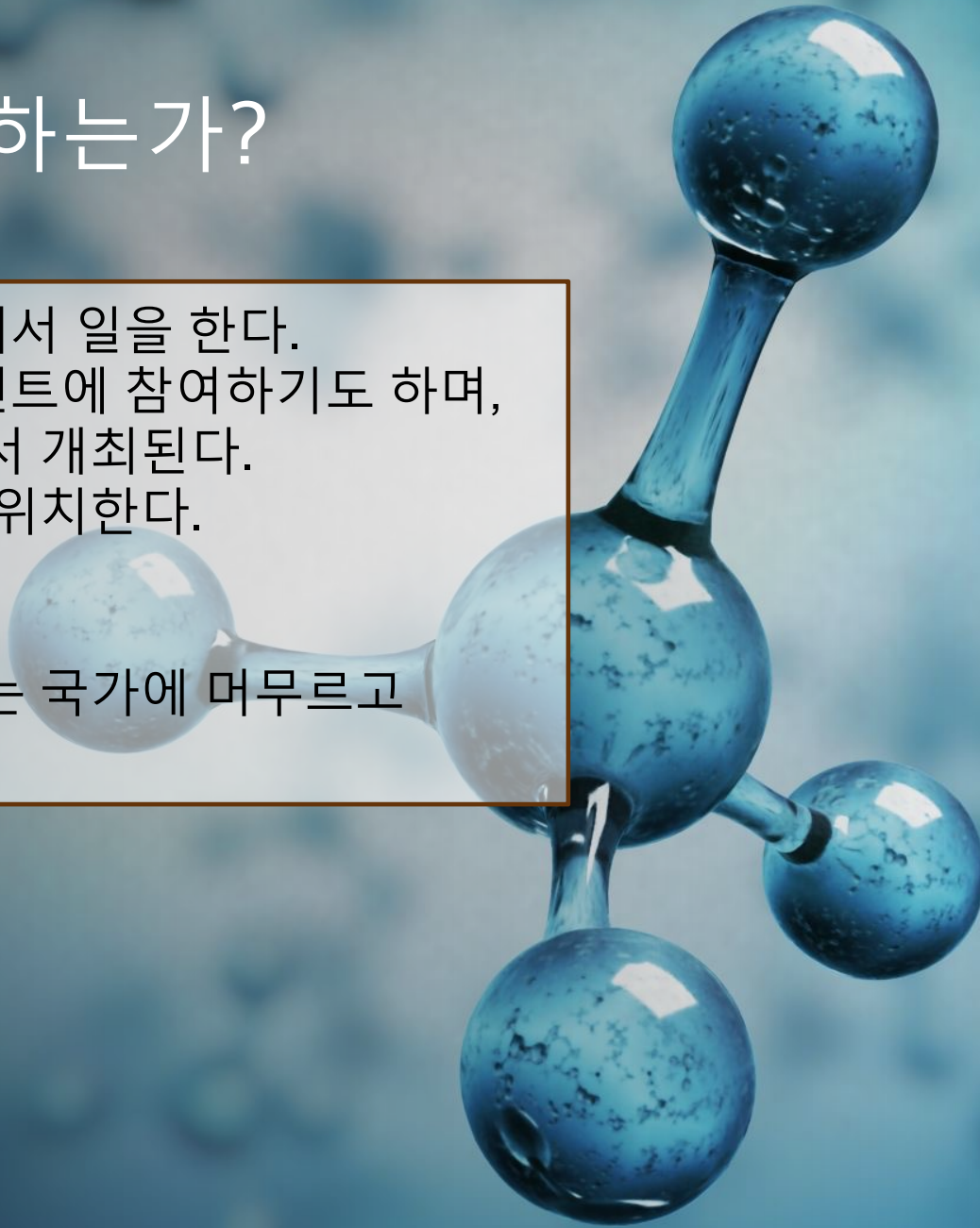
SKOS

SPARQL

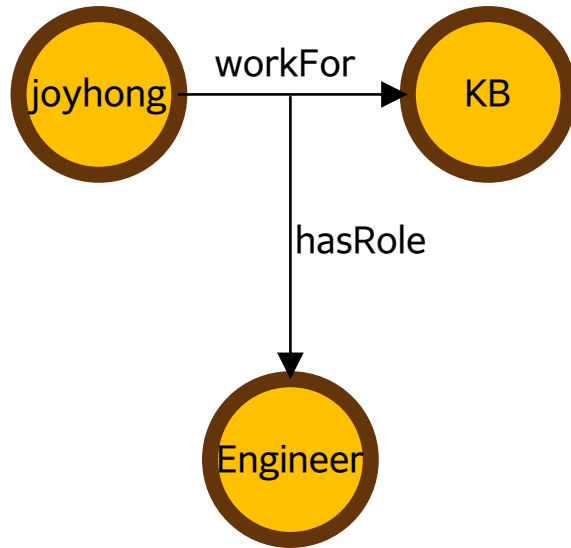
지식을 어떻게 추가하는가?

사람들은 역할을 가지고 회사에서 일을 한다.
사람들은 컨퍼런스와 같은 이벤트에 참여하기도 하며,
컨퍼런스는 특정 장소 인근에서 개최된다.
회사와 전철역은 특정 국가에 위치한다.

컨퍼런스에 참여하는 사람은
컨퍼런스 장소가 위치하고 있는 국가에 머무르고
있다는 것은 틀림없다.



RDF (Resource Description Framework)



PREFIX kg: <http://joyhong.tistory.com/gug/>

```
kg:joyhong a kg:Person;  
           kg:workFor { kg:hasRole kg:Engineer } kg:KB .
```

```
kg:KB a kg:Company ;  
      rdfs:label "KB국민은행" .
```


지식을 어떻게 추가하는가?

사람들은 역할을 가지고 회사에서 일을 한다.
사람들은 컨퍼런스와 같은 이벤트에 참여하기도 하며,
컨퍼런스는 특정 장소 인근에서 개최된다.
회사와 전철역은 특정 국가에 위치한다.

컨퍼런스에 참여하는 사람은
컨퍼런스 장소가 위치하고 있는 국가에 머무르고
있다는 것은 틀림없다.

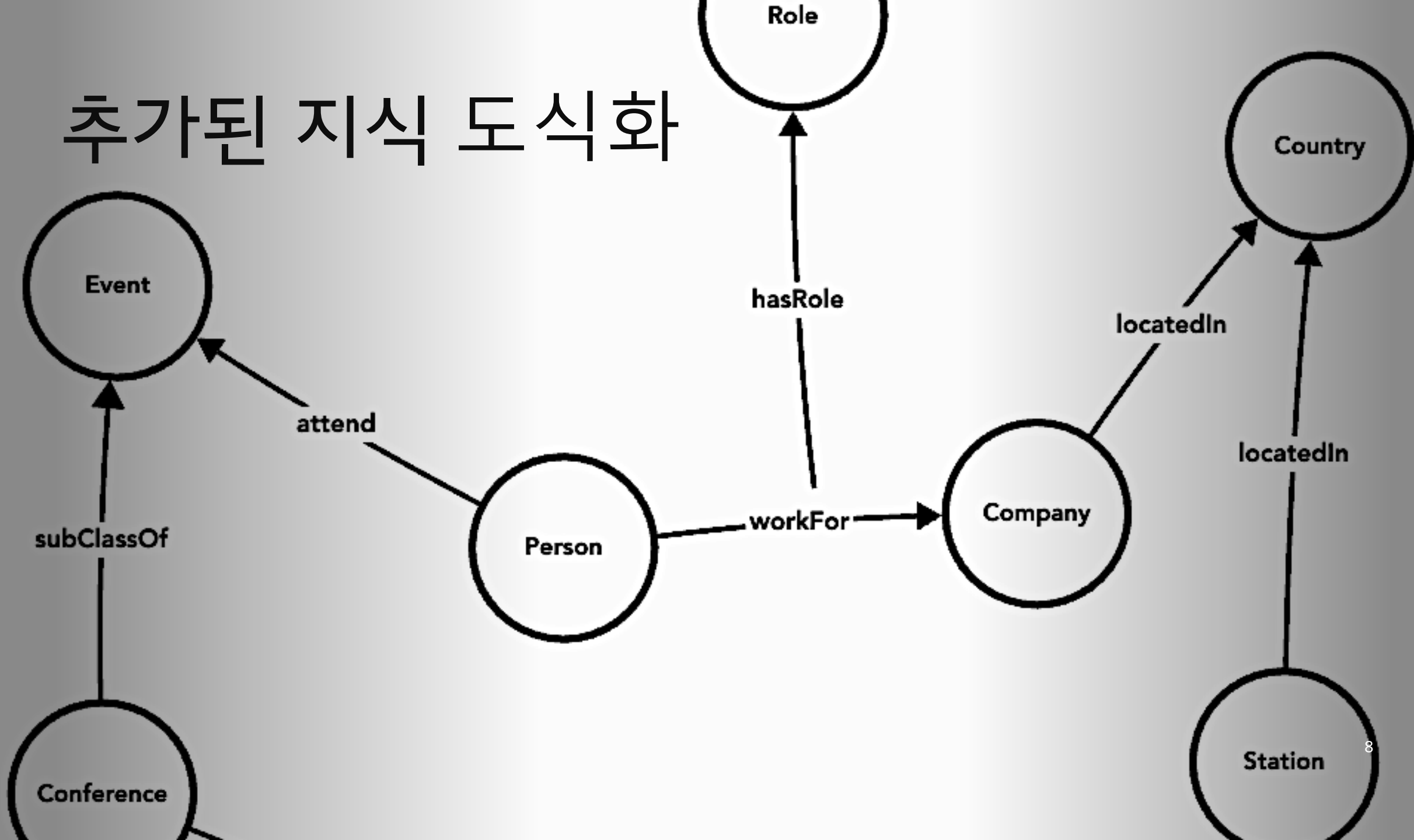
URI



```
kg:Company a owl:Class ;  
  rdfs:label "Company" ;  
  rdfs:comment "회사" .  
kg:Person a owl:Class ;  
  rdfs:label "Person" ;  
  rdfs:comment "사람" .  
kg:workAt a owl:ObjectProperty ;  
  rdfs:label "workFor" ;  
  so:domainIncludes kg:Person ;  
  so:rangeIncludes kg:Company .
```

<..생략..>

추가된 지식 도식화



데이터를 결합하기

joyhong은 KB에서 엔지니어로 일한다.

그리고 GUG 2번째 세미나에 발표자로
참여한다.

KB는 회사이고, 강남역은 전철역이다.

KB와 강남역은 SouthKorea에 위치한다.

kg:joyhong a kg:Person ;
kg:attend {kg:hasRole kg:presenter } kg:gug2nd ;
kg:workFor {kg:hasRole kg:engineer} kg:KB .

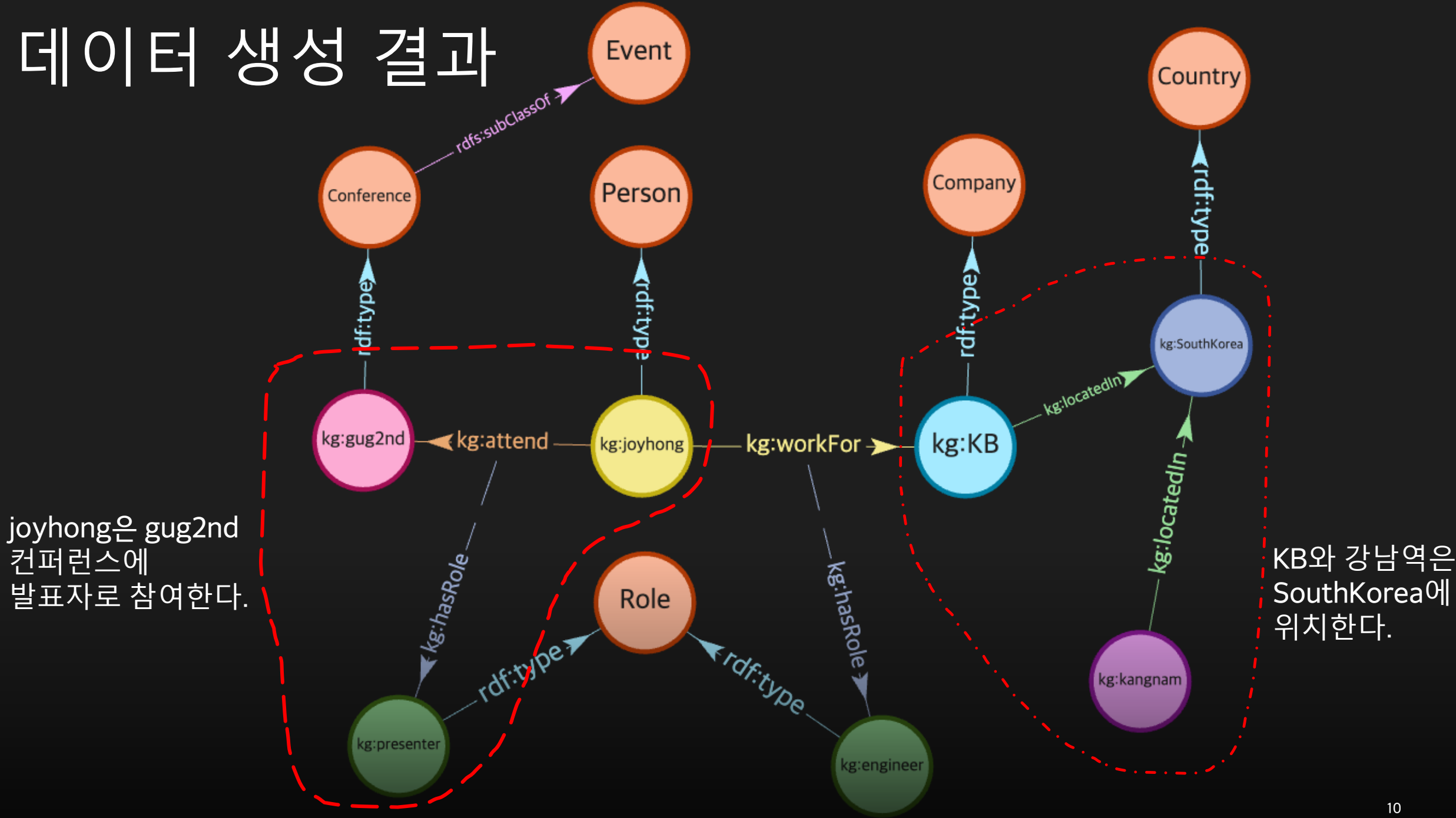
kg:gug2nd a kg:Conference .

kg:KB a kg:Company;
kg:locatedIn kg:SouthKorea .

kg:kangnam a kg:Station ;
kg:locatedIn kg:SouthKorea .

kg:SouthKorea a kg:Country .

데이터 생성 결과





월리를 찾아라? 데이터를 찾아라!

SPARQL을 통한 데이터 찾기

RDF Query Language

```
SELECT 찾을변수  
FROM 대상 그래프  
WHERE {  
    조건식  
}
```

```
PREFIX kg: <http://joyhong.tistory.com/gug/>  
SELECT *  
WHERE {  
    ?s kg:stayAt ?o.  
    VALUES ?s { kg:joyhong }  
}
```

joyhong은 현재 어디에 머무르고 있는가?



결과 없음

새로운 데이터 추가

joyhong은 KB에서 엔지니어로 일한다.

그리고 GUG 2번째 세미나에 발표자로
참여한다.

KB는 회사이고, 강남역은 전철역이다.

KB와 강남역은 SouthKorea에 위치한다.

GUG 2번째 세미나는 강남역 인근에서
개최된다.

```
kg:joyhong a kg:Person ;  
  kg:attend {kg:hasRole kg:presenter } kg:gug2nd ;  
  kg:workFor {kg:hasRole kg:engineer} kg:KB .
```

```
kg:gug2nd a kg:Conference .
```

```
kg:KB a kg:Company;  
  kg:locatedIn kg:SouthKorea .
```

```
kg:kangnam a kg:Station ;  
  kg:locatedIn kg:SouthKorea .
```

```
kg:SouthKorea a kg:Country .
```

```
kg:gug2nd kg:heldNear kg:kangnam .
```


SPARQL을 통한 데이터 찾기

joyhong은 현재 어디에 머무르고 있는가?

RULE rule:StayAtRule

```
IF {  
  ?x a kg:Person ;  
    kg:attend ?conference .  
  ?conference kg:heldNear ?place .  
  ?place kg:locatedIn ?country .  
}  
THEN {  
  ?x kg:stayAt ?country .  
}
```

추론규칙

```
PREFIX kg: <http://joyhong.tistory.com/gug/>  
SELECT *  
WHERE {  
  ?s kg:stayAt ?o.  
  VALUES ?s { kg:joyhong }  
}
```



kg:joyhong kg:stayAt kg:SouthKorea .

Knowledge Graph

- 현실 세계의 지식을 표현, 구조화에 활용
- 개체(Entity)와 개체 간의 관계(Relationship)를 표현하여 지식의 의미적인 관점을 제공
- 지식 그래프는 데이터 검색, 질의 응답, 추론 등 다양한 지식 기반 작업에 활용
- 지식 그래프는 도메인 전문 지식을 표현하고 쿼리할 수 있으며, 질문에 대한 정확하고 구조화된 답변을 제공할 수 있음
- 지식 그래프는 대규모 지식 베이스를 구축하고, 데이터 통합, 정보 검색, 추천 시스템, 자연어 이해 등 다양한 응용 분야에서 활용

LARGE LANGUAGE MODELS



현실 세계의 지식을
표현하고 구조화

방대한 데이터로 훈련하여
단어의 순서를 예측

KNOWLEDGE
GRAPH와

LARGE
LANGUAGE
MODELS

LLM과 KG

송강호와 함께 출연한 배우는

송강호와 함께 출연한 배우는

2021년

이전까지

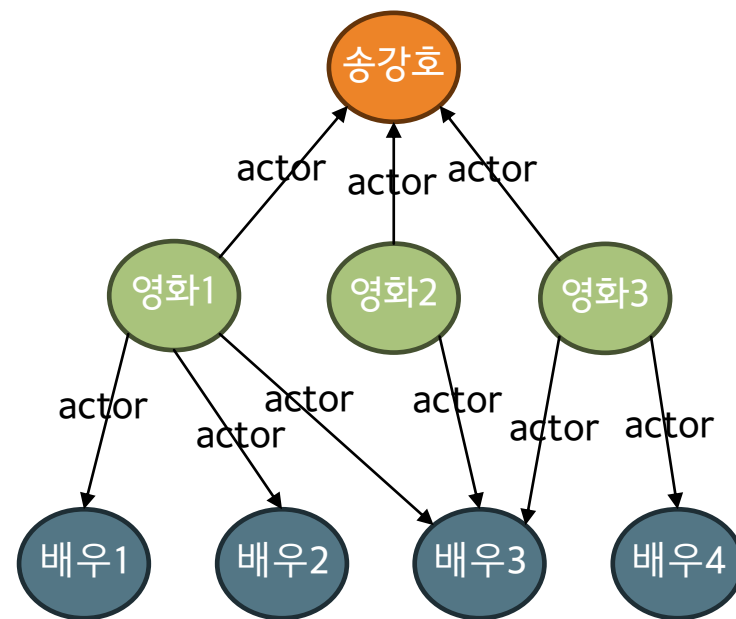
정보에

따르면

송강호와

공동

LLM



Knowledge Graph



송강호와 공동 출연한 배우는?



송강호와 공동으로 출연한 다른 배우들은 많이 있습니다. 몇몇 유명한 배우들 중 일부는 다음과 같습니다.

1. 이선균 (Lee Sun-kyun): 영화 '기생충'에서 함께 출연했습니다.
2. 조여정 (Cho Yeo-jeong): 영화 '기생충'에서 함께 출연했습니다.
3. 박소담 (Park So-dam): 영화 '기생충'에서 함께 출연했습니다.
4. 최민식 (Choi Min-sik): 영화 '쇼트 터미널'에서 함께 출연했습니다.
5. 김하늘 (Kim Hye-na): 영화 '살인의 추억'에서 함께 출연했습니다.
6. 김영애 (Kim Young-ae): 영화 '살인의 추억'에서 함께 출연했습니다.
7. 공지태 (Kong Hyo-jin): 영화 '트림'에서 함께 출연했습니다.

이 외에도 송강호는 다양한 배우들과 함께 작품에 출연하였으며, 한국 영화 산업에서 많은 배우들과 협업하였습니다.



송강호와 공동 출연한 배우는?



2021년 9월 이전까지의 정보에 따르면, 송강호가 공동 출연한 배우는 다
지에 따라 다를 수 있습니다. 송강호는 한국의 유명한 배우로서 다양한 직
동 출연한 배우는 그의 다양한 작품에 따라 다를 것입니다.

송강호의 최근 작품이나 공동 출연한 배우에 대한 정보는 제가 알고 있는
보를 확인하려면 인터넷 검색 엔진이나 영화 및 드라마 관련 웹사이트를

LLM의 결과 (시점이 변경되면 결과도 변동)

KG의 결과

- SPARQL 을 작성해야 하는 어려움이 존재

▶ Run

Show Plan

Reasoning

Movie

Q

```
1 select ?mo (SAMPLE(?movie) as ?movie) ?o (SAMPLE(?actor) as ?actor)
2 where {
3     ?mo ont:actor <http://localhost/resource/person_10037018> , ?o ; rdfs:label ?movie.
4     ?o rdfs:label ?actor .
5     FILTER(LANG(?actor) != 'en' && LANG(?movie) != 'en')
6 } group by ?mo ?o
```

Run to File

Text

Charts

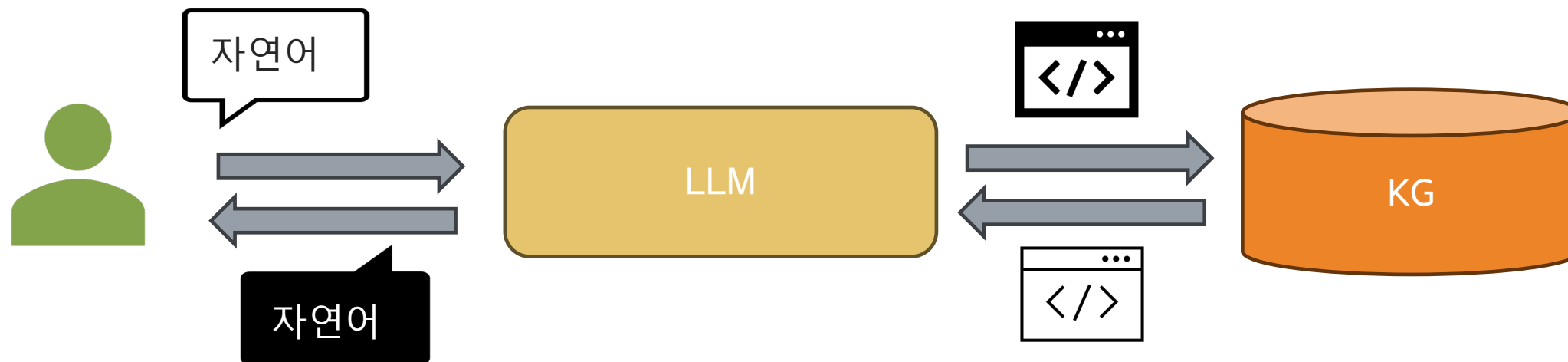
Visualize

1,000 Results, 655 ms

mo	movie	o
http://localhost/resource/movie_19970048	"초록물고기"	http://localhost/resource/person_20143709
http://localhost/resource/movie_20070072	"우아한 세계"	http://localhost/resource/person_20143709
http://localhost/resource/movie_20110249	"하울링"	http://localhost/resource/person_20187990
http://localhost/resource/movie_20110301	"푸른소금"	http://localhost/resource/person_20187990
http://localhost/resource/movie_19970022	"넘버 3"	http://localhost/resource/person_10029970
http://localhost/resource/movie_19960051	"돼지가 우물에 빠진 날"	http://localhost/resource/person_10029970
http://localhost/resource/movie_20060151	"괴물"	http://localhost/resource/person_10002379
http://localhost/resource/movie_20030046	"살인의 추억"	http://localhost/resource/person_10002379
http://localhost/resource/movie_19960051	"돼지가 우물에 빠진 날"	http://localhost/resource/person_20138079

LLM과 KG

구분	LLM	KG
장점	✓ 다양한 문맥, 상황에 맞는 자연스러운 결과(자연어)를 제공	✓ 도메인 전문 지식을 표현 ✓ 정확하고 구조화된 정보를 제공
재밋는 점	✓ 종종 자신만의 세계를 자랑함	✓ “질의언어”라는 보이는 장벽이 기다리고 있음



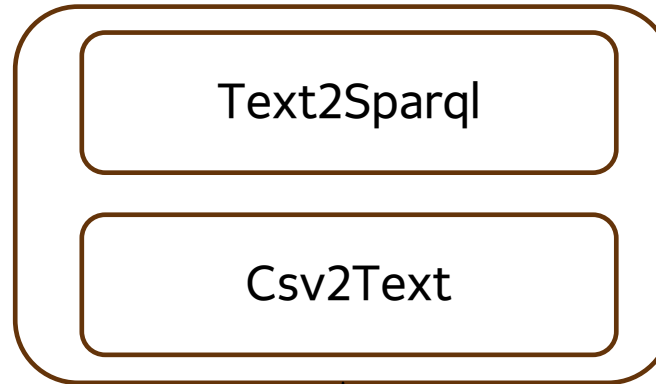
KG를 활용한 대화 프로세스



질문

결과

TEXT



Text2Sparql

Csv2Text

Knowledge
Graph

LLM

```
def run(question):
    question = prev_prompt + question + " sparql:"
    response = openai.Completion.create(
        model="text-davinci-003",
        temperature= 0.3,
        max_tokens= 800,
        top_p= 1,
        frequency_penalty= 0,
        presence_penalty= 0,
        prompt= question
    )
    return response['choices'][0]['text'].replace("\'", "\'")[1:]
```

executed in 7ms, finished 23:53:38 2023-09-19

▶

▶

```
x = run("송강호와 함께 출연한 배우는")
```

자연어 질문

executed in 5.53s, finished 23:53:47 2023-09-19

▶

x

SPARQL 생성

executed in 5ms, finished 23:53:48 2023-09-19

▶

```
'SELECT (SAMPLE(?a2label) as ?a2label) (count(?actor2) as ?cnt) (GROUP_CONCAT(?movie; SEPARATOR="|") as ?movie)
where {
    ?m ont:actor ?actor1.
    ?m ont:actor ?actor2.
    ?m rdfs:label ?movie.
    filter(LANG(?movie)!="en")
    FILTER(?actor1!=?actor2)
    ?m ont:dateOfOpen ?open.
    BIND(YEAR(?open) as ?year)
    ?actor1 rdfs:label ?a1label.
    filter(LANG(?a1label)!="en")
    ?actor2 rdfs:label ?a2label.
    filter(LANG(?a2label)!="en")
    filter(?a1label="송강호")}
group by ?actor1 ?actor2 ORDER BY DESC(?cnt)'
```

SPARQL 실행 결과

```

1 SELECT (SAMPLE(?a2label) as ?a2label)
2 (count(?actor2) as ?cnt)
3 (GROUP_CONCAT(?movie; SEPARATOR="|") as ?movie)
4 where {
5     ?m ont:actor ?actor1.
6     ?m ont:actor ?actor2.
7     ?m rdfs:label ?movie.
8     filter(LANG(?movie)!="en")
9     FILTER(?actor1!=?actor2)
10    ?m ont:dateOfOpen ?open.
11    BIND(YEAR(?open) as ?year)
12    ?actor1 rdfs:label ?a1label.
13    filter(LANG(?a1label)!="en")
14    ?actor2 rdfs:label ?a2label.
15    filter(LANG(?a2label)!="en")
16    filter(?a1label="송강호")
17 }
18 group by ?actor1 ?actor2 ORDER BY DESC(?cnt)
19

```

Run to File

Text

Charts

Visualize

977 Results, 809 ms

a2label	cnt	movie
"박진우"	7	"우아한 세계 살인의 추억 괴물 관상 좋은 놈, 나쁜 놈, 이...
"오달수"	7	"친절한 금자씨 괴물 박쥐 우아한 세계 변호인 효자동 이...
"조덕제"	6	"괴물 좋은 놈, 나쁜 놈, 이상한 놈 살인의 추억 조용한 가...
"이대연"	6	"YMCA야구단 공동경비구역 JSA 우아한 세계 친절한 ...
"최덕문"	5	"사도 나랏말싸미 푸른소금 YMCA야구단 남극일기"
"최교식"	4	"효자동 이발사 괴물 살인의 추억 관상"
"윤제문"	4	"괴물 좋은 놈, 나쁜 놈, 이상한 놈 남극일기 우아한 세계"
"이병헌"	4	"좋은 놈, 나쁜 놈, 이상한 놈 밀정 비상선언 공동경비구...
"최민식"	4	"조용한 가족 변호인 익선동 친절한 금자씨"

SPARQL 질의 결과

적용 결과

Chatbot with Knowledge Graph

Enter your text

송강호와 공동 출연한 배우는?

클리어

제출하기

output

송강호와 공동 출연한 배우는 박진우 오달수 조덕제 이대연 최덕문 최교식 윤제문 이병헌 최민식 배두나 입니다.

플래그

마무리

- 지식그래프는 데이터와 지식을 결합하여 데이터를 구조화시킨 지식베이스
- LLM은 단어를 순서를 예측하는데 초점을 맞추고 있는 언어 모델
- LLM 기술을 사용하는 ChatGPT와 지식그래프를 융합하여 활용
 - 지식그래프를 활용하여 지식 기반의 질문에 응답
 - 대규모로 학습한 일반적 지식을 통해 지식그래프 확장





감사합니다

Blog : <https://joyhong.tistory.com/>

Mail : su4620@gmail.com