

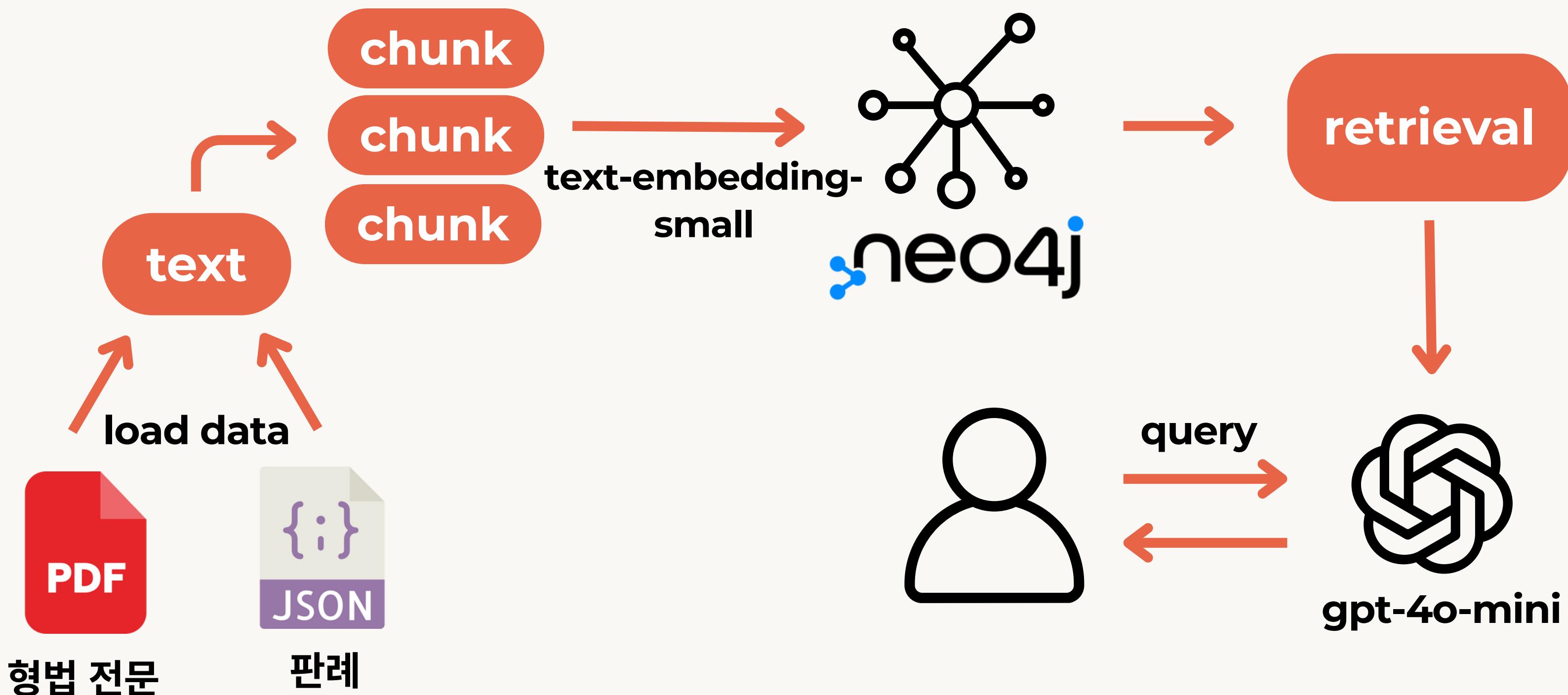
# assignment PT.

강병수

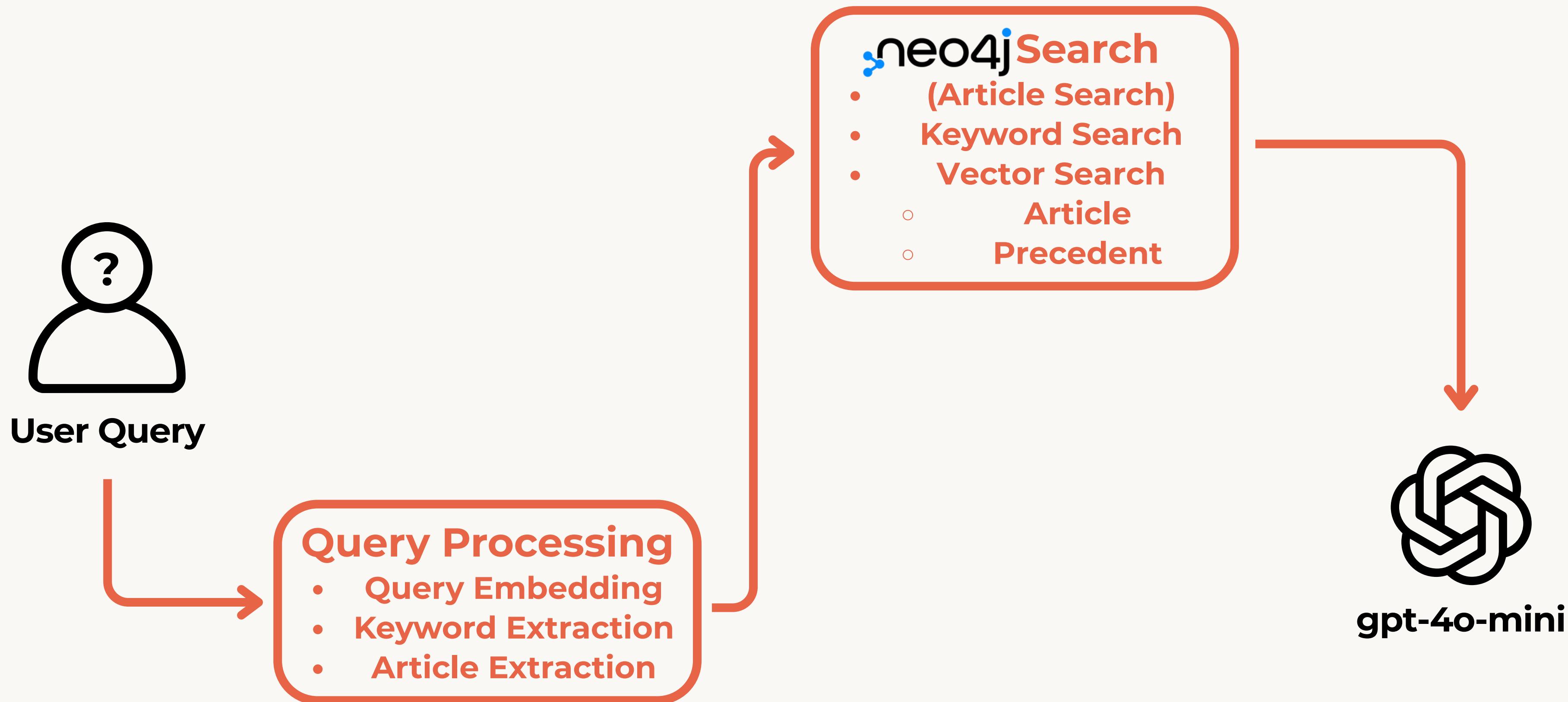
# Assignment remind.

- KMMLU Criminal-law 카테고리 평가를 위한 RAG 에이전트
- RAG에 사용될 데이터의 raw data와 정제하는 과정 코드에 포함
- 구축과 평가 모두 합쳐서 1시간 이내로 실행 가능
- gpt-4o-mini, text-embedding-small 만 사용

# Agent pipeline.

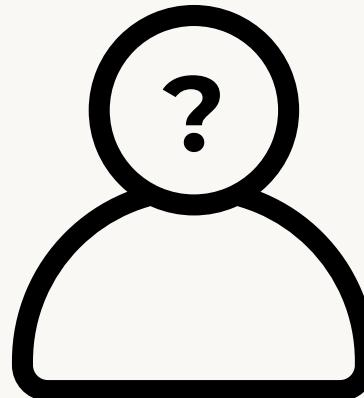


# How retrieval works.



# How retrieval works.

친구가 제 사진을 찍었는데  
초상권 침해 아닌가요?



User Query

neo4jSearch  
(Article Search)  
Keyword Search  
Vector Search  
Article  
Precedent

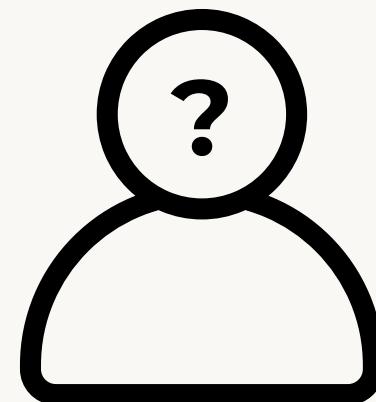
Query Processing  
• Query Embedding  
• Keyword Extraction  
• Article Extraction



gpt-4o-mini

# How retrieval works.

친구가 제 사진을 찍었는데  
초상권 침해 아닌가요?



User Query

- [0.023, -0.046, ..., -0.002]
- ['친구', '사진', '초상권침해']

if (query에 조항 번호 있으면)  
→ 제32조

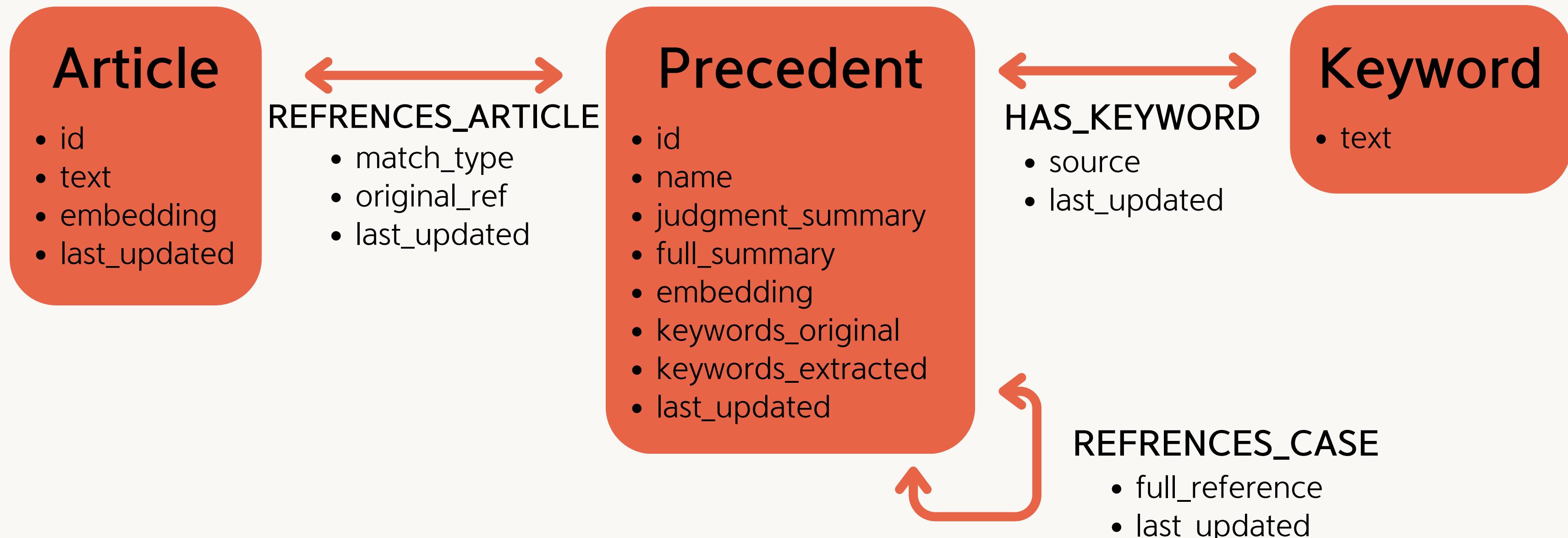
neo4jSearch  
(Article Search)  
Keyword Search  
Vector Search  
Article  
Precedent

Query Processing  
• Query Embedding  
• Keyword Extraction  
• (Article Extraction)



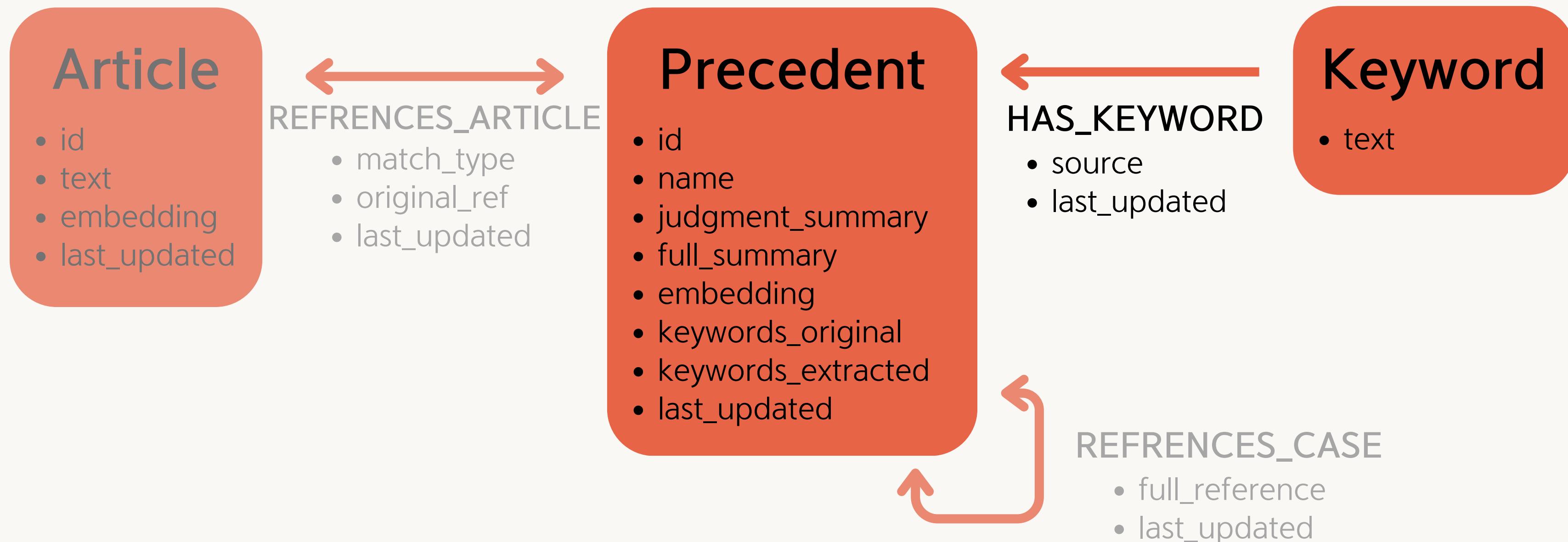
gpt-4o-mini

# Nodes & Relations.



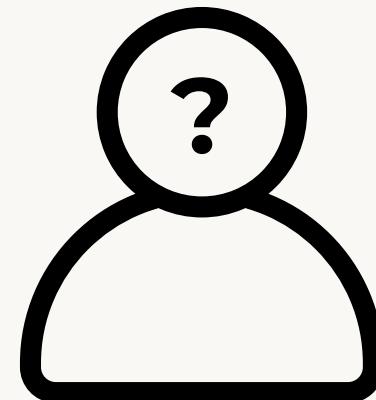
# Nodes & Relations.

[‘사진’, ‘초상권침해’]



# How retrieval works.

친구가 제 사진을 찍었는데  
초상권 침해 아닌가요?



User Query

- [0.023, -0.046, ..., -0.002]
- ['친구', '사진', '초상권침해']
- if (query에 조항 번호 있으면)  
→ 제32조

**Query Processing**

- Query Embedding
- Keyword Extraction
- (Article Extraction)

neo4j **Search**  
(Article Search)  
Keyword Search  
Vector Search  
Article  
Precedent

- (제32조 검색)
- ['사진', '초상권침해']
- [0.023, -0.046, ..., -0.002]
  - 명예훼손, 모욕죄, ...
  - 예시 판례 ...



gpt-4o-mini

# Agent methods.

## 초기 설정 및 데이터 로드

- `load_config_and_initialize()`
- `load_articles_from_pdf()`
- `load_precedents_from_json()`

## 그래프 데이터베이스 구축

- `connect_neo4j()`
- `setup_neo4j_constraints_and_indexes()`
- `create_graph_nodes_and_relationships()`

## 정보 검색 및 컨텍스트 구성

- `retrieve_context_from_graph()`
- `build_optimized_context()`

## 배치 처리 및 평가

- `prepare_batch_requests()`
- `run_batch_job()`
- `extract_choice_from_response()`
- `process_batch_results_and_evaluate()`
- `analyze_results()`

# Data preview.

## (형법)

데이터 출처:  
[법제처 국가법령정보센터](#)

### 제1편 총칙

#### 제1장 형법의 적용범위

- 제1조(범죄의 성립과 처벌) ① 범죄의 성립과 처벌은 행위 시의 법률에 의한다.  
② 범죄 후 법률의 변경에 의하여 그 행위가 범죄를 구성하지 아니하거나 형이 구법보다 경한 때에는 신법에 의한다.  
③ 재판확정 후 법률의 변경에 의하여 그 행위가 범죄를 구성하지 아니하는 때에는 형의 집행을 면제한다.

제2조(국내법) 본법은 대한민국영역 내에서 죄를 범한 내국인과 외국인에게 적용한다.

제3조(내국인의 국외범) 본법은 대한민국영역 외에서 죄를 범한 내국인에게 적용한다.

제4조(국외에 있는 내국선박 등에서 외국인이 범한 죄) 본법은 대한민국영역 외에 있는 대한민국의 선박 또는 항공기 내에서 죄를 범한 외국인에게 적용한다.

제5조(외국인의 국외범) 본법은 대한민국영역 외에서 다음에 기재한 죄를 범한 외국인에게 적용한다.

1. 내란의 죄
2. 외환의 죄
3. 국기에 관한 죄
4. 통화에 관한 죄
5. 유가증권, 우표와 인지에 관한 죄
6. 문서에 관한 죄 중 제225조 내지 제230조
7. 인장에 관한 죄 중 제238조

제6조(대한민국과 대한민국 국민에 대한 국외범) 본법은 대한민국영역 외에서 대한민국 또는 대한민국 국민에 대하여 전조에 기재한 이외의 죄를 범한 외국인에게 적용한다. 단 행위지의 법률에 의하여 범죄를 구성하지 아니하거나 소추 또는 형의 집행을 면제할 경우에는 예외로 한다.

# Data preview.

## (판례)

데이터 출처:

[AI Hub -](#)

법률/규정 텍스트 분석 데이터

```
"info": {  
    "id": "51325797",  
    "dataType": "판결문",  
    "caseNm": "국가보안법위반·간첩방조",  
    "caseTitle": "대법원 1979. 10. 10., 선고, 75도1003, 판결",  
    "courtType": "판례(대법원)",  
    "courtNm": "대법원",  
    "judmnAdjuDe": "1979-10-10",  
    "caseNoID": "75도1003",  
    "caseNo": "75도1003"  
},  
"dgmn": "간첩을 숨겨준 행위가 간첩방조가 되는지 여부",  
"dgmnInfo": [  
    {  
        "question": "간첩을 숨겨준 행위가 간첩방조가 되는가?",  
        "answer": "부정"  
    }  
],  
"Summary": [  
    {  
        "summ_ctxt": "간첩을 숨겨준 사실이 있다 하더라도 그 간첩이 군사기밀을 ...생략... 아니한다.",  
        "summ_pass": "간첩의 범행을 용이하게 하려는 의사가 있다고 볼 수 없으면 간첩방조죄는 성립되지 아니한다."  
    }  
],  
"keyword_tagg": [  
    {"id": 1, "keyword": "국가보안법위반"},  
    {"id": 2, "keyword": "간첩방조"}  
],  
"Reference_info": {  
    "reference_rules": "형법 제98조",  
    "reference_court_case": "대법원 1967.1.31. 선고 66도1616 판결"  
},  
"Class_info": {  
    "class_name": "형사A(생활형)",  
    "instance_name": "국가보안법위반"  
}
```

# Why graphRAG ?

```
“Refrence_info”: {  
    “refrence_rules”: “형법 제98조”,  
    “refrence_court_case”: “대법원 2001.6.20. 선고 ~~ 판결”  
}
```

## 제98조(간첩)

1. 적국을 위하여 간첩하거나 ~~ 7년 이상의 징역에 처한다.
2. 군사상의 기밀을 적국에 누설한 자도 전항의 형과 같다.

# Problems & solutions.

- 일반적인 embedding 기반 RAG 방식의 retrieval 성능 문제
- 초기 GraphRAG 구현 시 relation 설정 미흡
- 모델 파라미터 튜닝 미흡 (context length, batch size, temperature)

# Problems & solutions.

- 일반적인 embedding 기반 RAG 방식의 retrieval 성능 문제  
→ GraphRAG를 도입하여 관계 기반 탐색을 가능하게 함
- 초기 GraphRAG 구현 시 relation 설정 미흡
- 모델 파라미터 튜닝 미흡 (context length, batch size, temperature)

# Embedding based Vector Retrieval.

query = “정당방위의 요건은 무엇인가?”

--- Top Results ---

1. 유형: Precedent, ID: 92도2540, Score: 0.7387

이름: 살인

키워드: ['타인의 법익', '상당한 이유']

참조 법조항: ['제10조(심신장애자)', '제21조(정당방위)', '제308조(사자의 명예훼손)', '제308조', '제10조 (폐지되는 법률등)']

미리보기: 정당방위의 성립요건으로서의 방어행위에는 순수한 수비적 방어뿐 아니라 적극적 반격을 포함하는 반격방어의 형태도 포함됨은 소론과 같다고 하겠으나, 그 방어행위는 자기 또는 타인의 법익침해를 방위하기 위한 행위로서 상당한 이유가 있어야 하는 것인데, 피고인들의 판시 행위가 위에서 본 바와 같이 그 상당성을 결여한 것인 이상 정당방위행위로 평가될 수는 없는 것이므로, 원심이 피고인들의 이 사건 범행이 현재의 부당한 침해를 방위할 의사로 행해졌다기보다는 공격의 의사로 행하여졌다고 인정한 것이 적절하지 못하다고 하더라도, 정당방위행위가 되지 ...

-----  
2. 유형: Precedent, ID: 2009도2114, Score: 0.6939

이름: 특수공무집행방해[변경된죄명: 폭력행위등처벌에관한법률위반(공동폭행)]

키워드: ['사회상규에 위배', '소극적인 방어행위']

참조 법조항: ['제6조(대한민국과 대한민국 국민에 대한 국외범)', '제20조(정당행위)', '제21조(정당방위)', '제136조(공무집행방해)', '제136조', '제260조(폭행, 존속 폭행)', '제260조', '제6조 (경합법에 대한 신법의 적용례)', '제6조']

미리보기: 비록 경찰관들의 위법한 상경 제지 행위에 대항하기 위하여 한 것이라 하더라도, 피고인들이 다른 시위참가자들과 공동하여 위와 같이 경찰관들을 때리고 진압방패와 채증장비를 빼앗는 등의 폭행행위를 한 것은 소극적인 방어행위를 넘어서 공격의 의사를 포함하여 이루어진 것으로서 그 수단과 방법에 있어서 상당성이 인정된다고 보기 어려우며 긴급하고 불가피한 수단이었다고 볼 수도 없으므로, 이를 사회상규에 위배되지 아니하는 정당행위나 현재의 부당한 침해를 방어하기 위한 정당방위에 해당한다고 볼 수 없다.

그럼에도 불구하고, 원심은 그 판시와 같은 ...

-----  
3. 유형: Precedent, ID: 83누383, Score: 0.6789

이름: 영업정지처분무효확인

키워드: ['법령 위반행위']

참조 법조항: ['제1조(범죄의 성립과 처벌)', '제34조(간접정범, 특수한 교사, 방조에 대한 형의 가중)', '제37조(경합법)', '제38조(경합법과 처벌례)', '제34조', '제1조 (구형법 기타 법령과 형의 경증)', '제1조 (시행일)']

미리보기: 정당한 절차에 의하지 않고 구두에 의한 하도급계약을 체결하여 공사를 시작한 때에 건설업법 제34조 제3항의 위반행위를 범한 것이 되니 그 위반행위를 이유로 한 행정상의 제재처분(행위당시에는 필요적 취소사유)을 하려면 그 위반행위 이후 법령의 변경에 의하여 처분의 종류를 달리(영업정지 사유로) 규정하였다 하더라도 그 법률적용에 관한특별규정이 없다면 위반행위 당시에 시행되던 법령을 근거로 처분을 하여야 마땅하다....

# Embedding based Vector Retrieval.

query = “정당방위의 요건은 무엇인가?”

--- Top Results ---

1. 유형: Precedent, ID: 92도2540, Score: 0.7387

이름: 살인

키워드: ['타인의 법익', '상당한 이유']

참조 법조항: ['제10조(심신장애자)', '제21조(정당방위)', '제308조(사자의 명예훼손)', '제308조', '제10조 (폐지되는 법률등)']

미리보기: 정당방위의 성립요건으로서의 방어행위에는 순수한 수비적 방어뿐 아니라 적극적 반격을 포함하는 반격방어의 형태도 포함됨은 소론과 같다고 하겠으나, 그 방어행위는 자기 또는 타인의 법익침해를 방위하기 위한 행위로서 상당한 이유가 있어야 하는 것인데, 피고인들의 판시 행위가 위에서 본 바와 같이 그 상당성을 결여한 것인 이상 정당방위행위로 평가될 수는 없는 것이므로, 원심이 피고인들의 이 사건 범행이 현재의 부당한 침해를 방위할 의사로 행해졌다기보다는 공격의 의사로 행하여졌다고 인정한 것이 적절하지 못하다고 하더라도, 정당방위행위가 되지 ...

-----  
2. 유형: Precedent, ID: 2009도2114, Score: 0.6939

이름: 특수공무집행방해[변경된죄명: 폭력행위등처벌에관한법률위반(공동폭행)]

키워드: ['사회상규에 위배', '소극적인 방어행위']

참조 법조항: ['제6조(대한민국과 대한민국 국민에 대한 국외범)', '제20조(정당행위)', '제21조(정당방위)', '제136조(공무집행방해)', '제136조', '제260조(폭행, 존속 폭행)', '제260조', '제6조 (경합법에 대한 신법의 적용례)', '제6조']

미리보기: 비록 경찰관들의 위법한 상경 제지 행위에 대항하기 위하여 한 것이라 하더라도, 피고인들이 다른 시위참가자들과 공동하여 위와 같이 경찰관들을 때리고 진압방패와 채증장비를 빼앗는 등의 폭행행위를 한 것은 소극적인 방어행위를 넘어서 공격의 의사를 포함하여 이루어진 것으로서 그 수단과 방법에 있어서 상당성이 인정된다고 보기 어려우며 긴급하고 불가피한 수단이었다고 볼 수도 없으므로, 이를 사회상규에 위배되지 아니하는 정당행위나 현재의 부당한 침해를 방어하기 위한 정당방위에 해당한다고 볼 수 없다.

그럼에도 불구하고, 원심은 그 판시와 같은 ...

-----  
3. 유형: Precedent, ID: 83누383, Score: 0.6789

이름: 영업정지처분무효확인

키워드: ['법령 위반행위']

참조 법조항: ['제1조(범죄의 성립과 처벌)', '제34조(간접정범, 특수한 교사, 방조에 대한 형의 가중)', '제37조(경합법)', '제38조(경합법과 처벌례)', '제34조', '제1조 (구형법 기타 법령과 형의 경증)', '제1조 (시행일)']

미리보기: 정당한 절차에 의하지 않고 구두에 의한 하도급계약을 체결하여 공사를 시작한 때에 건설업법 제34조 제3항의 위반행위를 범한 것이 되니 그 위반행위를 이유로 한 행정상의 제재처분(행위당시에는 필요적 취소사유)을 하려면 그 위반행위 이후 법령의 변경에 의하여 처분의 종류를 달리(영업정지 사유로) 규정하였다 하더라도 그 법률적용에 관한특별한 규정이 없다면 위반행위 당시에 시행되던 법령을 근거로 처분을 하여야 마땅하다....

# Hybrid Retrieval.

query = “정당방위의 요건은 무엇인가?”

--- Top Results ---

1. 유형: Precedent, ID: 92도2540, 스코어: 0.7910

이름: 살인

키워드: ['타인의 법익', '상당한 이유']

참조 법조항: ['제10조(심신장애자)', '제21조(정당방위)', '제308조(사자의 명예훼손)', '제308조', '제10조 (폐지되는 법률등)']

미리보기: 정당방위의 성립요건으로서의 방어행위에는 순수한 수비적 방어뿐 아니라 적극적 반격을 포함하는 반격방어의 형태도 포함됨은 소론과 같다고 하겠으나, 그 방어행위는 자기 또는 타인의 법익침해를 방위하기 위한 행위로서 상당한 이유가 있어야 하는 것인데, 피고인들의 판시 행위가 위에서 본 바와 같이 그 상당성을 결여한 것인 이상 정당방위행위로 평가될 수는 없는 것이므로, 원심이 피고인들의 이 사건 범행이 현재의 부당한 침해를 방위할 의사로 행해졌다기보다는 공격의 의사로 행하여졌다고 인정한 것이 적절하지 못하다고 하더라도, 정당방위행위가 되지 ...

-----  
2. 유형: Precedent, ID: 2003도4934, 스코어: 0.7295

이름: 명예훼손(일부 인정된 죄명 : 모욕)·폭행

키워드: ['사회상규에 위배되지 아니하는 행위', '행위의 수단이나 방법의 상당성']

참조 법조항: ['제20조(정당행위)', '제21조(정당방위)', '제232조(자격모용에 의한 사문서의 작성)', '제232조의2(사전자기록위작·변작)', '제260조(폭행, 존속폭행)', '제260조', '제307조(명예훼손)', '제307조', '제311조(모욕)', '제312조(고소와 피해자의 의사)', '제311조', '제314조(업무방해)', '제316조(비밀침해)']

미리보기: 형법 제20조 소정의 ‘사회상규에 위배되지 아니하는 행위’라 함은 법질서 전체의 정신이나 그 배후에 놓여 있는 사회윤리 내지 사회통념에 비추어 용인될 수 있는 행위를 말하고, 어떠한 행위가 사회상규에 위배되지 아니하는 정당한 행위로서 위법성이 조각되는 것인지는 구체적인 사정 아래서 합목적적, 합리적으로 고찰하여 개별적으로 판단되어야 하므로, 이와 같은 정당행위가 인정되려면, 첫째 그 행위의 동기나 목적의 정당성, 둘째 행위의 수단이나 방법의 상당성, 셋째 보호이익과 침해이익의 법익 균형성, 넷째 긴급성, 다섯째 그 행위 이외의 다른...  
-----

3. 유형: Precedent, ID: 83도3124, 스코어: 0.7163

이름: 출판물에의한명예훼손

키워드: ['불특정 다수인']

참조 법조항: ['제20조(정당행위)', '제307조(명예훼손)', '제309조(출판물 등에 의한 명예훼손)', '제307조', '제310조(위법성의 조각)', '제309조']

미리보기: 명예훼손죄의 요건인 공연성은 불특정 또는 다수인이 인식할 수 있는 상태를 말하므로, 원심확정과 같이 피고인들이 이 사건 출판물 15부를 피고인들이 소속된 교회의 교인 15인에게 배부한 이상 공연성의 요건은 충족된 것이라고 보겠으며 배부받은 사람중 일부가 소론과 같이 위 출판물작성에 가담한 사람들이라고 하여도 결론에 아무런 소장이 없으니, 명예훼손죄의 성립을 인정한 원심조치에 소론과 같이 공연성에 관한 법리를 오해한 위법이 없다. 또 위 출판물에 게재된 내용이 허위사실인 이상 당연히 위법성이 인정되는 것이고 사회통념상 위법성이 인정되...

# Hybrid Retrieval.

query = “정당방위의 요건은 무엇인가?”

--- Top Results ---

1. 유형: Precedent, ID: 92도2540, 스코어: 0.7910

이름: 살인

키워드: ['타인의 법익', '상당한 이유']

참조 법조항: ['제10조(심신장애자)', '제21조(정당방위)', '제308조(사자의 명예훼손)', '제308조', '제10조 (폐지되는 법률등)']

미리보기: 정당방위의 성립요건으로서의 방어행위에는 순수한 수비적 방어뿐 아니라 적극적 반격을 포함하는 반격방어의 형태도 포함됨은 소론과 같다고 하겠으나, 그 방어행위는 자기 또는 타인의 법익침해를 방위하기 위한 행위로서 상당한 이유가 있어야 하는 것인데, 피고인들의 판시 행위가 위에서 본 바와 같이 그 상당성을 결여한 것인 이상 정당방위행위로 평가될 수는 없는 것이므로, 원심이 피고인들의 이 사건 범행이 현재의 부당한 침해를 방위할 의사로 행해졌다기보다는 공격의 의사로 행하여졌다고 인정한 것이 적절하지 못하다고 하더라도, 정당방위행위가 되지 ...

-----  
2. 유형: Precedent, ID: 2003도4934, 스코어: 0.7295

이름: 명예훼손(일부 인정된 죄명 : 모욕)·폭행

키워드: ['사회상규에 위배되지 아니하는 행위', '행위의 수단이나 방법의 상당성']

참조 법조항: ['제20조(정당행위)', '제21조(정당방위)', '제232조(자격모용에 의한 사문서의 작성)', '제232조의2(사전자기록위작·변작)', '제260조(폭행, 존속폭행)', '제260조', '제307조(명예훼손)', '제307조', '제311조(모욕)', '제312조(고소와 피해자의 의사)', '제311조', '제314조(업무방해)', '제316조(비밀침해)']

미리보기: 형법 제20조 소정의 ‘사회상규에 위배되지 아니하는 행위’라 함은 법질서 전체의 정신이나 그 배후에 놓여 있는 사회윤리 내지 사회통념에 비추어 용인될 수 있는 행위를 말하고, 어떠한 행위가 사회상규에 위배되지 아니하는 정당한 행위로서 위법성이 조각되는 것인지는 구체적인 사정 아래서 합목적적, 합리적으로 고찰하여 개별적으로 판단되어야 하므로, 이와 같은 정당행위가 인정되려면, 첫째 그 행위의 동기나 목적의 정당성, 둘째 행위의 수단이나 방법의 상당성, 셋째 보호이익과 침해이익의 법익 균형성, 넷째 긴급성, 다섯째 그 행위 이외의 다른...  
-----

3. 유형: Precedent, ID: 83도3124, 스코어: 0.7163

이름: 출판물에의한명예훼손

키워드: ['불특정 다수인']

참조 법조항: ['제20조(정당행위)', '제307조(명예훼손)', '제309조(출판물 등에 의한 명예훼손)', '제307조', '제310조(위법성의 조각)', '제309조']

미리보기: 명예훼손죄의 요건인 공연성은 불특정 또는 다수인이 인식할 수 있는 상태를 말하므로, 원심확정과 같이 피고인들이 이 사건 출판물 15부를 피고인들이 소속된 교회의 교인 15인에게 배부한 이상 공연성의 요건은 충족된 것이라고 보겠으며 배부받은 사람중 일부가 소론과 같이 위 출판물작성에 가담한 사람들이라고 하여도 결론에 아무런 소장이 없으니, 명예훼손죄의 성립을 인정한 원심조치에 소론과 같이 공연성에 관한 법리를 오해한 위법이 없다. 또 위 출판물에 게재된 내용이 허위사실인 이상 당연히 위법성이 인정되는 것이고 사회통념상 위법성이 인정되...

# Multi Step Retrieval.

query = “정당방위의 요건은 무엇인가?”

--- Top Results ---

1. 유형: Precedent, ID: 92도2540, 스코어: 0.7386

이름: 살인

키워드: ['타인의 법익', '상당한 이유']

참조 법조항: ['제10조(심신장애자)', '제21조(정당방위)', '제308조(사자의 명예훼손)', '제308조', '제10조 (폐지되는 법률등)']

미리보기: 정당방위의 성립요건으로서의 방어행위에는 순수한 수비적 방어뿐 아니라 적극적 반격을 포함하는 반격방어의 형태도 포함됨은 소론과 같다고 하겠으나, 그 방어행위는 자기 또는 타인의 법익침해를 방위하기 위한 행위로서 상당한 이유가 있어야 하는 것인데, 피고인들의 판시 행위가 위에서 본 바와 같이 그 상당성을 결여한 것인 이상 정당방위행위로 평가될 수는 없는 것이므로, 원심이 피고인들의 이 사건 범행이 현재의 부당한 침해를 방위할 의사로 행해졌다기보다는 공격의 의사로 행하여졌다고 인정한 것이 적절하지 못하다고 하더라도, 정당방위행위가 되지 ...

-----  
2. 유형: Precedent, ID: 2009도2114, 스코어: 0.6940

이름: 특수공무집행방해[변경된죄명: 폭력행위등처벌에관한법률위반(공동폭행)]

키워드: ['사회상규에 위배', '소극적인 방어행위']

참조 법조항: ['제6조(대한민국과 대한민국 국민에 대한 국외법)', '제20조(정당행위)', '제21조(정당방위)', '제136조(공무집행방해)', '제136조', '제260조(폭행, 존속 폭행)', '제260조', '제6조 (경합법에 대한 신법의 적용례)', '제6조']

미리보기: 비록 경찰관들의 위법한 상경 제지 행위에 대항하기 위하여 한 것이라 하더라도, 피고인들이 다른 시위참가자들과 공동하여 위와 같이 경찰관들을 때리고 진압방패와 채증장비를 빼앗는 등의 폭행행위를 한 것은 소극적인 방어행위를 넘어서 공격의 의사를 포함하여 이루어진 것으로서 그 수단과 방법에 있어서 상당성이 인정된다고 보기 어려우며 긴급하고 불가피한 수단이었다고 볼 수도 없으므로, 이를 사회상규에 위배되지 아니하는 정당행위나 현재의 부당한 침해를 방어하기 위한 정당방위에 해당한다고 볼 수 없다.

그럼에도 불구하고, 원심은 그 판시와 같은 ...

-----  
3. 유형: Precedent, ID: 83누383, 스코어: 0.6791

이름: 영업정지처분무효확인

키워드: ['법령 위반행위']

참조 법조항: ['제1조(범죄의 성립과 처벌)', '제34조(간접정범, 특수한 교사, 방조에 대한 형의 가중)', '제37조(경합법)', '제38조(경합법과 처벌례)', '제34조', '제1조 (구형법 기타 법령과 형의 경증)', '제1조 (시행일)']

미리보기: 정당한 절차에 의하지 않고 구두에 의한 하도급계약을 체결하여 공사를 시작한 때에 건설업법 제34조 제3항의 위반행위를 범한 것이 되니 그 위반행위를 이유로 한 행정상의 제재처분(행위당시에는 필요적 취소사유)을 하려면 그 위반행위 이후 법령의 변경에 의하여 처분의 종류를 달리(영업정지 사유로) 규정하였다 하더라도 그 법률적용에 관한특별한 규정이 없다면 위반행위 당시에 시행되던 법령을 근거로 처분을 하여야 마땅하다.

# Multi Step Retrieval.

query = “정당방위의 요건은 무엇인가?”

--- Top Results ---

1. 유형: Precedent, ID: 92도2540, 스코어: 0.7386

이름: 살인

키워드: ['타인의 법익', '상당한 이유']

참조 법조항: ['제10조(심신장애자)', '제21조(정당방위)', '제308조(사자의 명예훼손)', '제308조', '제10조 (폐지되는 법률등)']

미리보기: 정당방위의 성립요건으로서의 방어행위에는 순수한 수비적 방어뿐 아니라 적극적 반격을 포함하는 반격방어의 형태도 포함됨은 소론과 같다고 하겠으나, 그 방어행위는 자기 또는 타인의 법익침해를 방위하기 위한 행위로서 상당한 이유가 있어야 하는 것인데, 피고인들의 판시 행위가 위에서 본 바와 같이 그 상당성을 결여한 것인 이상 정당방위행위로 평가될 수는 없는 것이므로, 원심이 피고인들의 이 사건 범행이 현재의 부당한 침해를 방위할 의사로 행해졌다기보다는 공격의 의사로 행하여졌다고 인정한 것이 적절하지 못하다고 하더라도, 정당방위행위가 되지 ...

-----  
2. 유형: Precedent, ID: 2009도2114, 스코어: 0.6940

이름: 특수공무집행방해[변경된죄명: 폭력행위등처벌에관한법률위반(공동폭행)]

키워드: ['사회상규에 위배', '소극적인 방어행위']

참조 법조항: ['제6조(대한민국과 대한민국 국민에 대한 국외법)', '제20조(정당행위)', '제21조(정당방위)', '제136조(공무집행방해)', '제136조', '제260조(폭행, 존속 폭행)', '제260조', '제6조 (경합법에 대한 신법의 적용례)', '제6조']

미리보기: 비록 경찰관들의 위법한 상경 제지 행위에 대항하기 위하여 한 것이라 하더라도, 피고인들이 다른 시위참가자들과 공동하여 위와 같이 경찰관들을 때리고 진압방패와 채증장비를 빼앗는 등의 폭행행위를 한 것은 소극적인 방어행위를 넘어서 공격의 의사를 포함하여 이루어진 것으로서 그 수단과 방법에 있어서 상당성이 인정된다고 보기 어려우며 긴급하고 불가피한 수단이었다고 볼 수도 없으므로, 이를 사회상규에 위배되지 아니하는 정당행위나 현재의 부당한 침해를 방어하기 위한 정당방위에 해당한다고 볼 수 없다.

그럼에도 불구하고, 원심은 그 판시와 같은 ...

-----  
3. 유형: Precedent, ID: 83누383, 스코어: 0.6791

이름: 영업정지처분무효확인

키워드: ['법령 위반행위']

참조 법조항: ['제1조(범죄의 성립과 처벌)', '제34조(간접정범, 특수한 교사, 방조에 대한 형의 가중)', '제37조(경합법)', '제38조(경합법과 처벌례)', '제34조', '제1조 (구형법 기타 법령과 형의 경증)', '제1조 (시행일)']

미리보기: 정당한 절차에 의하지 않고 구두에 의한 하도급계약을 체결하여 공사를 시작한 때에 건설업법 제34조 제3항의 위반행위를 범한 것이 되니 그 위반행위를 이유로 한 행정상의 제재처분(행위당시에는 필요적 취소사유)을 하려면 그 위반행위 이후 법령의 변경에 의하여 처분의 종류를 달리(영업정지 사유로) 규정하였다 하더라도 그 법률적용에 관한특별한 규정이 없다면 위반행위 당시에 시행되던 법령을 근거로 처분을 하여야 마땅하다.

# Graph based Retrieval.

query = “정당방위의 요건은 무엇인가?”

--- Top Results ---

1. 유형: Article, ID: 제21조(정당방위), 스코어: 0.7724

관련 판례 수: 7

관련 키워드: ['공소시효', '정지', '연장', '배제', '특례조항', '소급적용', '경과규정']

미리보기: 제21조(정당방위) ①자기 또는 타인의 법익에 대한 현재의 부당한 침해를 방위하기 위한 행위는 상당한 이유가 있는 때에는 벌하지 아니한다.

②방위행위가 그 정도를 초과한 때에는 정황에 의하여 그 형을 감경 또는 면제할 수 있다.

③전항의 경우에 그 행위가 야간 기타 불안스러운 상태하에서 공포, 경악, 흥분 또는 당황으로 인한 때에는 벌하지 아니한다.

-----  
2. 유형: Article, ID: 제21조(정당방위), 스코어: 0.7624

관련 판례 수: 6

관련 키워드: ['성범죄', '선고형', '경합범', '성폭력처벌법', '개정', '신상정보 등록기간']

미리보기: 제21조(정당방위) ①자기 또는 타인의 법익에 대한 현재의 부당한 침해를 방위하기 위한 행위는 상당한 이유가 있는 때에는 벌하지 아니한다.

②방위행위가 그 정도를 초과한 때에는 정황에 의하여 그 형을 감경 또는 면제할 수 있다.

③전항의 경우에 그 행위가 야간 기타 불안스러운 상태하에서 공포, 경악, 흥분 또는 당황으로 인한 때에는 벌하지 아니한다.

-----  
3. 유형: Precedent, ID: 92도2540, 스코어: 0.7387

이름: 살인

키워드: ['타인의 법익', '상당한 이유']

참조 법조항: ['제10조(심신장애자)', '제21조(정당방위)', '제308조(사자의 명예훼손)', '제308조', '제10조 (폐지되는 법률등)']

미리보기: 정당방위의 성립요건으로서의 방어행위에는 순수한 수비적 방어뿐 아니라 적극적 반격을 포함하는 반격방어의 형태도 포함됨은 소론과 같다고 하겠으나, 그 방어행위는 자기 또는 타인의 법익침해를 방위하기 위한 행위로서 상당한 이유가 있어야 하는 것인데, 피고인들의 판시 행위가 위에서 본 바와 같이 그 상당성을 결여한 것인 이상 정당방위행위로 평가될 수는 없는 것이므로, 원심이 피고인들의 이 사건 범행이 현재의 부당한 침해를 방위할 의사로 행해졌다기보다는 공격의 의사로 행하여졌다고 인정한 것이 적절하지 못하다고 하더라도, 정당방위행위가 되지 ...

# Graph based Retrieval.

query = “정당방위의 요건은 무엇인가?”

--- Top Results ---

1. 유형: Article, ID: 제21조(정당방위), 스코어: 0.7724

관련 판례 수: 7

관련 키워드: ['공소시효', '정지', '연장', '배제', '특례조항', '소급적용', '경과규정']

미리보기: 제21조(정당방위) ①자기 또는 타인의 법익에 대한 현재의 부당한 침해를 방위하기 위한 행위는 상당한 이유가 있는 때에는 벌하지 아니한다.

②방위행위가 그 정도를 초과한 때에는 정황에 의하여 그 형을 감경 또는 면제할 수 있다.

③전항의 경우에 그 행위가 야간 기타 불안스러운 상태하에서 공포, 경악, 흥분 또는 당황으로 인한 때에는 벌하지 아니한다.

-----  
2. 유형: Article, ID: 제21조(정당방위), 스코어: 0.7624

관련 판례 수: 6

관련 키워드: ['성범죄', '선고형', '경합범', '성폭력처벌법', '개정', '신상정보 등록기간']

미리보기: 제21조(정당방위) ①자기 또는 타인의 법익에 대한 현재의 부당한 침해를 방위하기 위한 행위는 상당한 이유가 있는 때에는 벌하지 아니한다.

②방위행위가 그 정도를 초과한 때에는 정황에 의하여 그 형을 감경 또는 면제할 수 있다.

③전항의 경우에 그 행위가 야간 기타 불안스러운 상태하에서 공포, 경악, 흥분 또는 당황으로 인한 때에는 벌하지 아니한다.

-----  
3. 유형: Precedent, ID: 92도2540, 스코어: 0.7387

이름: 살인

키워드: ['타인의 법익', '상당한 이유']

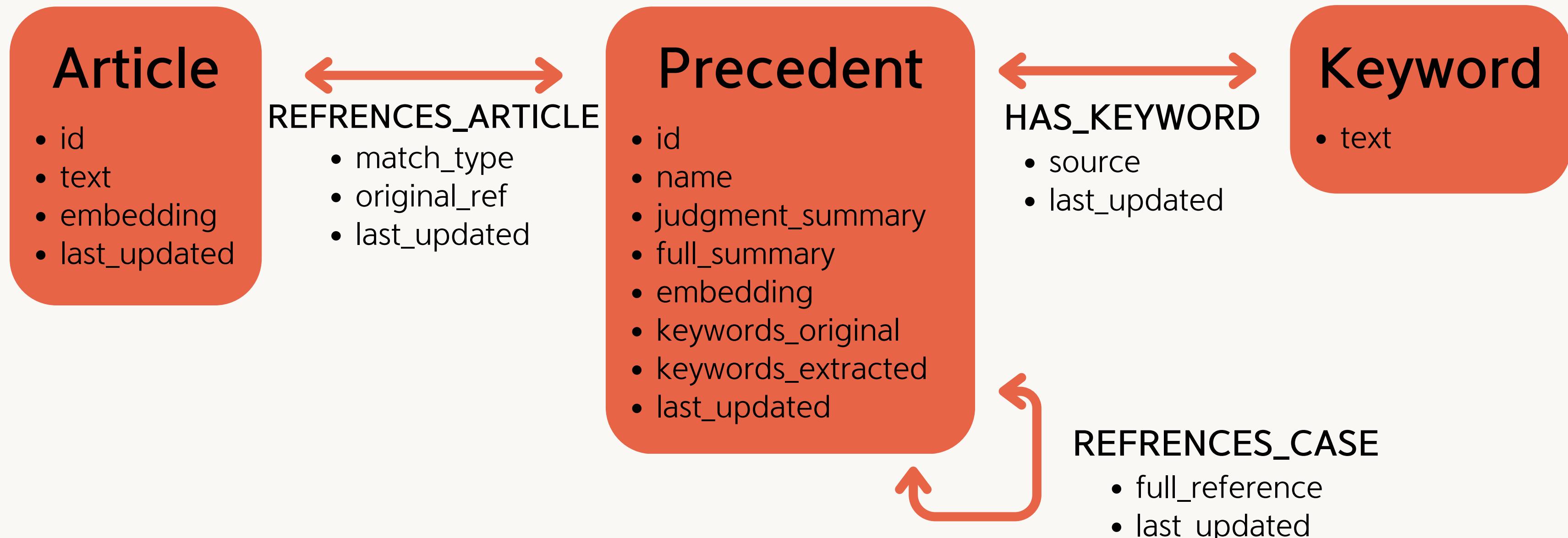
참조 법조항: ['제10조(심신장애자)', '제21조(정당방위)', '제308조(사자의 명예훼손)', '제308조', '제10조 (폐지되는 법률등)']

미리보기: 정당방위의 성립요건으로서의 방어행위에는 순수한 수비적 방어뿐 아니라 적극적 반격을 포함하는 반격방어의 형태도 포함됨은 소론과 같다고 하겠으나, 그 방어행위는 자기 또는 타인의 법익침해를 방위하기 위한 행위로서 상당한 이유가 있어야 하는 것인데, 피고인들의 판시 행위가 위에서 본 바와 같이 그 상당성을 결여한 것인 이상 정당방위행위로 평가될 수는 없는 것이므로, 원심이 피고인들의 이 사건 범행이 현재의 부당한 침해를 방위할 의사로 행해졌다기보다는 공격의 의사로 행하여졌다고 인정한 것이 적절하지 못하다고 하더라도, 정당방위행위가 되지 ...

# Problems & solutions.

- 일반적인 embedding 기반 RAG 방식의 retrieval 성능 문제  
→ GraphRAG를 도입하여 관계 기반 탐색을 가능하게 함
- 초기 GraphRAG 구현 시 relation 설정 미흡  
→ Node & Relation 구현
- 모델 파라미터 튜닝 미흡 (context length, batch size, temperature)

# Nodes & Relations.



# Problems & solutions.

- 일반적인 embedding 기반 RAG 방식의 retrieval 성능 문제  
→ GraphRAG를 도입하여 관계 기반 탐색을 가능하게 함
- 초기 GraphRAG 구현 시 relation 설정 미흡  
→ Node & Relation 구현
- 모델 파라미터 튜닝 미흡 (context length, batch size, temperature)  
→ Bayesian optimization (ongoing)

# Bayesian optimization.

$(x - 2)^2$ 에서 최적의 x값 찾기

## 1. 무작위 평가

$f(0) = 4, f(5) = 9, \dots$

## 2. 대체 모델 학습

1을 바탕으로 대체 모델이 함수를 예측함

## 3. 획득 함수로 다음 지점 선택

## 4. 모델 평가

```
import numpy as np
from skopt import gp_minimize
from skopt.space import Real

# 최적화할 함수 정의
def objective(x):
    return (x[0] - 2) ** 2 # x는 리스트이므로 x[0]를 사용

# 하이퍼파라미터 범위 설정
space = [Real(-10, 10, name='x')] # x는 -10에서 10 사이의 실수

# 베이지안 최적화 실행
res = gp_minimize(
    objective, # 최적화할 함수
    space, # 하이퍼파라미터 범위
    n_calls=10, # 총 10번 시도
    random_state=0 # 결과 재현성을 위해
)

# 결과 출력
print(f"최적의 x 값: {res.x[0]}") # 최적의 하이퍼파라미터
print(f"최적의 함수 값: {res.fun}") # 그때의 함수 값

# 출력 결과
# 최적의 x 값: 1.8568923645003679
# 최적의 함수 값: 0.020479795338295567
```

# Problems & solutions.



# Probler

On the data efficiency question, humans, for whatever other flaws we have about learning things, we seem unbelievably data efficient.  
How far away is our very algorithm currently from human level data?



# Problems & solutions.

Really hard to measure apples to apples.  
i think just like vibes,  
by in language  
astronomically far away



# Improvement plans.

- 법 domain 지식 이해
- 더 많은 데이터 (ex: 기출문제집 해설)
- RAFT: Adapting LM to Domain Specific RAG  
<https://arxiv.org/pdf/2403.10131>
- NodeRAG  
<https://arxiv.org/pdf/2504.11544>

# Improvement plans.

- 법 domain 지식 이해
- 더 많은 데이터 (ex: 기출문제집 해설)
- RAFT: Adapting LM to Domain Specific RAG  
<https://arxiv.org/pdf/2403.10131>
- NodeRAG  
<https://arxiv.org/pdf/2504.11544>

# Improvement plans.

## 법 domain 지식 이해

## 더 많은 데이터 (e.g., 기출문제집 해설)

## RAFT: Adapting LM to Domain Spec

<https://arxiv.org/pdf/2403.10131.pdf>

## NodeRAG

<https://arxiv.org/pdf/2504.11544.pdf>

업무상 과실치사상죄에 대한 다음의 설명 중 옳지 않은 것은 모두 몇 개인가? (다툼이 있는 경우 판례에 의함)

16. 경간부

- ㉠ 수술 도중에 수술용 메스가 부러지자 담당의사가 부러진 메스조각을 찾아 제거하려고 노력을 다하였으나 찾지 못하자 메스조각의 정확한 위치와 이동상황을 파악한 후 재수술을 생각으로 수술부위를 봉합한 경우에 담당의사의 업무상 과실을 인정할 수 없다.
- ㉡ 야간 당직간호사가 담당 환자의 심근경색증상을 당직의사에게 제대로 보고하지 않아 당직의사가 필요한 조치를 취하지 못한 채 환자가 사망한 경우, 당직간호사에게 업무상 과실을 인정할 수 없다.
- ㉢ 내과의사가 신경과 전문의에 대한 협의진료 결과와 환자에 대한 진료경과 등을 신뢰하여 뇌혈관계통 질환의 가능성을 염두에 두지 않고 내과 영역의 진료행위를 계속하다가 환자의 뇌지주막하출혈을 발견하지 못하여 식물인간 상태에 이르게 한 경우, 내과의사의 업무상 과실이 인정된다.
- ㉣ 교사가 징계목적으로 학생들의 손바닥을 때리기 위해 회초리를 들어 올리는 순간 이를 구경하기 위해 옆으로 고개를 돌려 일어나는 다른 학생의 눈을 찔러 그로 하여금 우안 실명의 상해를 입게 한 경우, 업무상 과실치상죄에 해당한다.
- ㉤ 건설회사가 건설공사 중 타워크레인의 설치작업을 전문업자에게 도급주어 타워크레인 설치작업을 하던 중 발생한 사고에 대하여 건설회사의 현장대리인에게 업무상 과실이 인정된다.

① 1개

② 2개

③ 3개

④ 4개

- ㉠ ○ ⇨ 대판 1999.12.10, 99도3711 ; 부러진 메스조각 사건
- ㉡ × ⇨ 병원의 야간당직 운영체계상 당직간호사에게 환자의 사망을 예견하거나 회피하지 못한 업무상 과실이 있고, 당직의사에게는 업무상 과실을 인정하기 어렵다(대판 2007.9.20, 2006도294).
- ㉢ × ⇨ 내과의사가 신경과 전문의에 대한 협의진료 결과 피해자의 증세와 관련하여 신경과 영역에서 이상이 없다는 회신을 받았고, 그 회신 전후의 진료 경과에 비추어 그 회신 내용에 의문을 품을 만한 사정이 있다고 보이지 않자 그 회신을 신뢰하여 뇌혈관계통 질환의 가능성을 염두에 두지 않고 내과 영역의 진료행위를 계속하다가 피해자의 증세가 호전되기에 이르자 퇴원하도록 조치한 경우, 피해자의 지주막하출혈을 발견하지 못한 데 대하여 내과의사의 업무상 과실은 부정된다(대판 2003.1.10, 2001도3292).
- ㉣ × ⇨ 직접 징계당하는 학생의 옆에 있는 다른 학생이 징계당하는 것을 구경하기 위하여 고개를 돌려 뒤에서 다가선다든가 옆자리에서 일어나는 것까지 예견할 수는 없다고 할 것이다(대판 1985.7.9, 84도822).
- ㉤ × ⇨ 건설회사의 현장대리인에게 이 사건 타워크레인의 설치작업을 관리하고 통제할 실질적인 지휘·감독 권한이 있었던 것으로는 보이지 아니하므로 업무상 과실치사상의 죄책을 물을 수 없다(대판 2005.9.9, 2005도3108 ; 현장대리인 사건).

# RAFT.

## Adapting LM to Domain Specific RAG

LLM 을 특정 domain에 맞게  
fine-tuning 하는 방법론

- **Question**
- **Golden document**
- **Distractor document**

→ Answer (chain-of-thought)

arXiv:2403.10131v2 [cs.CL] 5 Jun 2024

### RAFT: Adapting Language Model to Domain Specific RAG

Tianjun Zhang \*  
Department of Computer Science  
UC Berkeley  
Berkeley, CA 94720, USA  
[{tianjunz}@berkeley.edu](mailto:{tianjunz}@berkeley.edu)

Shishir G. Patil, Naman Jain, Sheng Shen  
Department of Computer Science  
UC Berkeley  
Berkeley, CA 94720, USA  
[{shishirpatil,naman\\_jain,sheng.s}@berkeley.edu](mailto:{shishirpatil,naman_jain,sheng.s}@berkeley.edu)

Matei Zaharia, Ion Stoica, Joseph E. Gonzalez  
Department of Computer Science  
UC Berkeley  
Berkeley, CA 94720, USA  
[{matei,istoica,jegonzal}@berkeley.edu](mailto:{matei,istoica,jegonzal}@berkeley.edu)

#### Abstract

Pretraining Large Language Models (LLMs) on large corpora of textual data is now a standard paradigm. When using these LLMs for many downstream applications, it is common to additionally incorporate new information into the pretrained model either through RAG-based-prompts, or finetuning. However, the best methodology to incorporate information remains an open question. In this paper, we present Retrieval Augmented Fine Tuning (RAFT), a training recipe which improves the model's ability to answer questions in "open-book" in-domain settings. In training RAFT, given a question, and a set of retrieved documents, we train the model to ignore those documents that don't help in answering the question, which we call, distractor documents. RAFT accomplishes this by citing verbatim the right sequence from the relevant document to help answer the question. This coupled with RAFT's chain-of-thought-style response helps improve the model's ability to reason. In domain specific RAG, RAFT consistently improves the model's performance across PubMed, HotpotQA, and Gorilla datasets, presenting a post-training recipe to improve pre-trained LLMs to in-domain RAG.

#### 1 Introduction

Trained on vast quantities of public data, Large Language Models LLMs have achieved significant advances in a wide range of general knowledge reasoning tasks Brown et al. (2020); Wei et al. (2022). However, increasingly LLMs are being employed in specialized domains to support tasks ranging from code completion for specific software frameworks to question answering on specific document collections (e.g., legal or medical documents). In these settings, general knowledge reasoning is less critical and instead the primary goal is to maximize accuracy based on a given set of documents. Indeed, adapting LLMs to the specialized domains (e.g., recent news, enterprise private documents, or program resources constructed after the training cutoff) is essential to many emerging applications (Vu et al. 2023; Lazaridou et al. 2022) and is the focus of this work.

This paper studies the following question – *How do we adapt pre-trained LLMs for Retrieval Augmented Generation (RAG) in specialized domains?*

When it comes to adapting LLMs to specialized domains, we consider the following two candidates: in-context learning through Retrieval-Augmented Generation (RAG) and supervised fine-tuning. RAG based methods allow the LLM to reference the documents when

\*Corresponding author, personal website: [tianjunz.github.io](https://tianjunz.github.io)

# NodeRAG.

## Structuring Graph-based RAG with Heterogenous Nodes

### Heterograph

- Entity (N)
- Relationship (R)
- Semantic Unit (S)
- Attribute (A)
- High-Level Element (H)
- High-Level Overview (O)
- Text (T)

arXiv:2504.11544v1 [cs.AI] 15 Apr 2025

### NodeRAG: Structuring Graph-based RAG with Heterogeneous Nodes

Tianyang Xu<sup>1</sup>, Haojie Zheng<sup>2</sup>, Chengze Li<sup>1</sup>, Haoxiang Chen<sup>1</sup>  
Yixin Liu<sup>3</sup>, Ruoxi Chen, Lichao Sun<sup>3</sup>

<sup>1</sup> Columbia University, <sup>2</sup> University of Pennsylvania, <sup>3</sup> Lehigh University  
tx2240@columbia.edu, haojiez@seas.upenn.edu

#### Abstract

Retrieval-augmented generation (RAG) empowers large language models to access external and private corpus, enabling factually consistent responses in specific domains. By exploiting the inherent structure of the corpus, graph-based RAG methods further enrich this process by building a knowledge graph index and leveraging the structural nature of graphs. However, current graph-based RAG approaches seldom prioritize the design of graph structures. Inadequately designed graph not only impede the seamless integration of diverse graph algorithms but also result in workflow inconsistencies and degraded performance. To further unleash the potential of graph for RAG, we propose NodeRAG, a graph-centric framework introducing heterogeneous graph structures that enable the seamless and holistic integration of graph-based methodologies into the RAG workflow. By aligning closely with the capabilities of LLMs, this framework ensures a fully cohesive and efficient end-to-end process. Through extensive experiments, we demonstrate that NodeRAG exhibits performance advantages over previous methods, including GraphRAG and LightRAG, not only in indexing time, query time, and storage efficiency but also in delivering superior question-answering performance on multi-hop benchmarks and open-ended head-to-head evaluations with minimal retrieval tokens. Our GitHub repository could be seen at this [link](#).

#### 1 Introduction

Retrieval-augmented generation (RAG) has emerged as a solution to the challenges posed by the rapid evolution of real-world knowledge domains (Fan et al., 2024), coupling large language models (LLMs) with an external retrieval mechanism to ensure the generation of factually consistent and contextually relevant information (Tonmoy et al., 2024; Shrestha et al., 2024; Liu

et al., 2024). Despite recent progress, current RAG methods face notable shortcomings in handling multi-hop reasoning (Luo et al., 2023; Wang et al., 2024b) and summary-level queries (Han et al., 2024a; Wen et al., 2023) due to their insufficient utilization of data structures and lack of high-level understanding of the text corpus. Graph-based RAG methods (Tian et al., 2024; Park et al., 2023) have been proposed to enhance retrieval and question-answering performance, specifically addressing the two main challenges faced by traditional RAG approaches. Leveraging LLMs to decompose raw data into graph structures (Jiménez Gutiérrez et al., 2024; He et al., 2024) for utilizing structural information, as well as employing LLMs for summary-based enhancements (Edge et al., 2024; Guo et al., 2024) to derive insights beyond the original text, has gradually become mainstream approaches.

However, previous Graph-based RAG works (Trajanoska et al., 2023; Jiménez Gutiérrez et al., 2024) have rarely considered the critical role of graph structures, i.e., what forms of graph better support RAG. Among existing approaches, knowledge graphs (Sanmartin, 2024; Wang et al., 2024b) extract triples, with the graph containing only structural information, yet retrieval context remains confined to text chunks, which often lack semantic coherence and include unrelated information. While current methods attempt to incorporate more information into the graph and extract deeper insights, they suffer from inefficiencies and inconsistencies due to inadequately designed structures. For instance, as illustrated in Figure 1, GraphRAG (Edge et al., 2024), adopt a tightly coupled entity-event homogeneous structure, hindering the integration of original context and summary information into the graph. This results in inconsistencies in retrieval methods (separating local and global retrieval) and leads to coarse-grained retrieval, where retrieving an entity indiscriminately includes all

# **Thank you.**

**정말 감사합니다 !!**