



LAWKEEPER

전문 변호사 수준의 신뢰할 수 있는 AI 법률 상담



<https://lawkeeper.net/>

팀원 서동현, 유동현, 김준희, 노승욱



CONTENTS

Chapter 01

목적 및 전체 프로세스

Chapter 02

데이터 수집 및 전처리

Chapter 03

임베딩 전략 및 RAG 기반 검색

Chapter 04

RAG 기반 구현 코드 및 프롬프트 구성

Chapter 05

홈페이지 기능 설명 및 배포 방법

Chapter 06

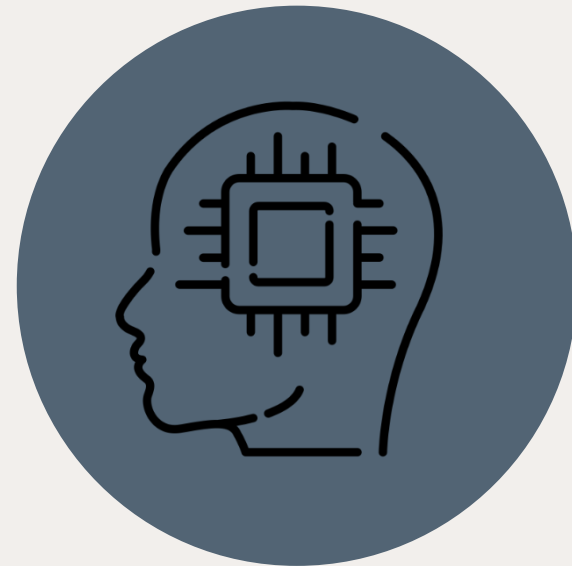
자사 활용 예시

01 프로젝트 목적



AI agent

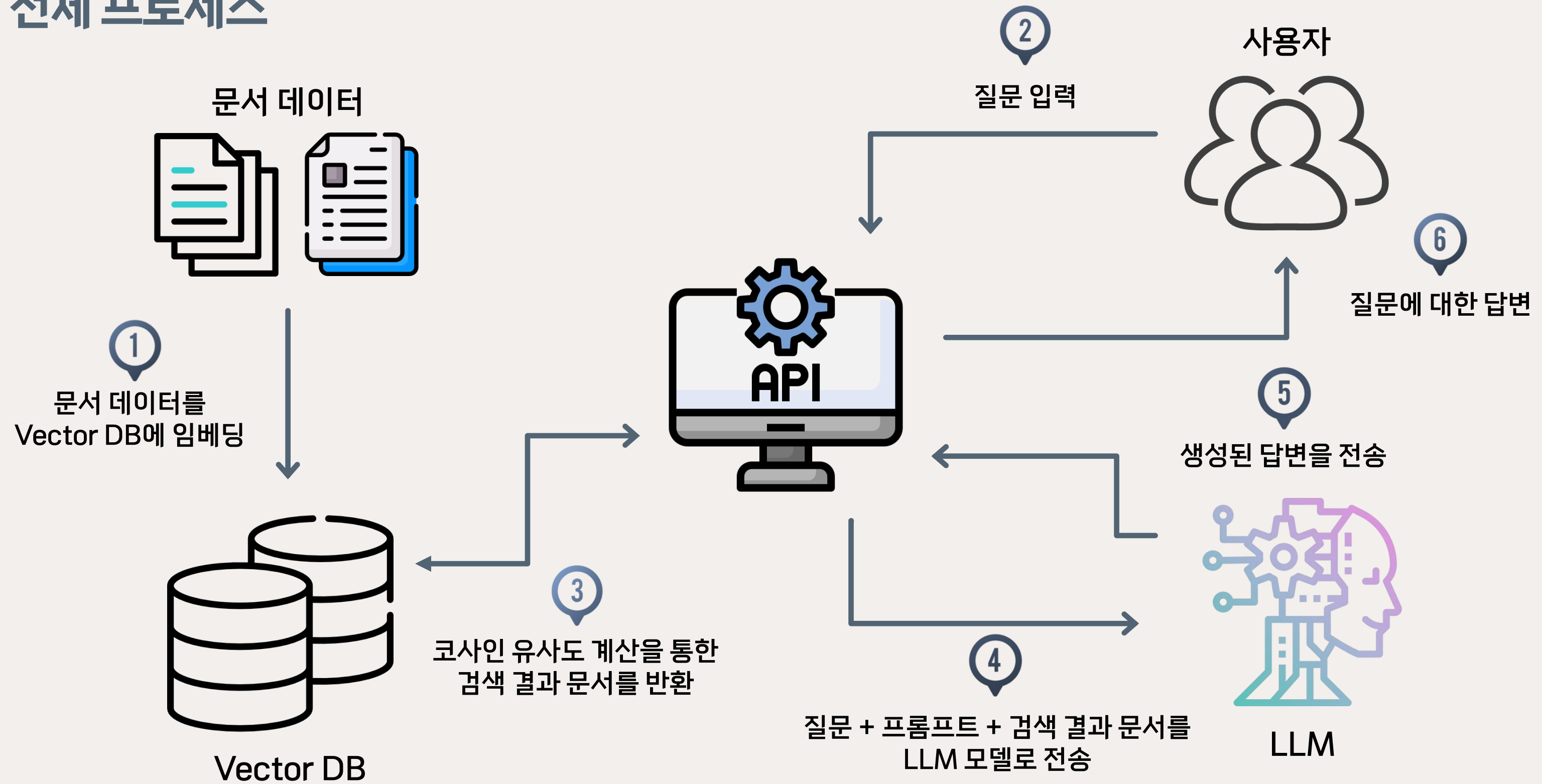
일반적인 LLM 모델에
내가 가진 문서나 데이터를 결합해서
한 분야에 특화된 지능형
행동 시스템을 구성하는 것임



프로젝트 목적

법률적인 문제가 생기거나
관련 궁금증이 있을 때,
누구나 쉽게 상담 받을 수 있도록
하고자 만들었음

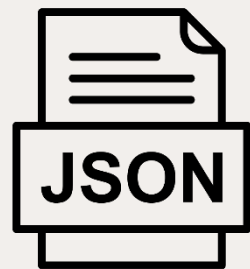
01 전체 프로세스



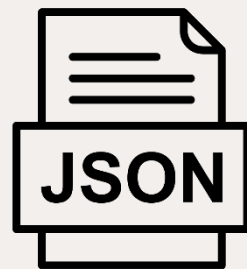
02 데이터 수집



법률/규정 상황별 판례 데이터
법률/규정 판결서 약관 데이터



대법원 판례 데이터
법률 상담 데이터



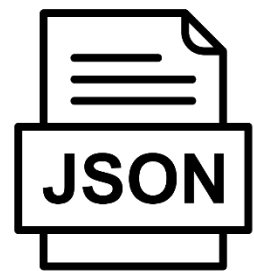
대학교 법학과 법전 데이터



법률 상담 데이터
사이버 상담 데이터



02 데이터 전처리



판례, 약관 등
json 형식의 데이터

A

- 1) 원본 JSON 파일 로드
- 2) 판결문 등 주요 내용을 본문으로 그 외 정보는 메타데이터로 분리
- 3) 조문·판례·약관 단위로 별도 JSON 파일로 저장



크롤링 기반 Q&A
CSV 형식의 데이터

B

- 1) 원본 CSV 파일 로드
- 2) 질문 & 답변을 본문으로 그 외 정보는 메타데이터로 분리
- 3) 각 행의 데이터를 별도 JSON 파일로 저장



법령 및 조문
PDF 형식의 데이터

C

- 1) 원본 PDF 파일 로드
- 2) 글꼴과 글자 크기를 활용한 조문 텍스트 추출
- 3) 조문 내용은 본문으로 그 외 정보는 메타데이터로 분리
- 4) 법령별 별도 JSON 파일로 저장

03 임베딩 전략 및 RAG 기반 검색



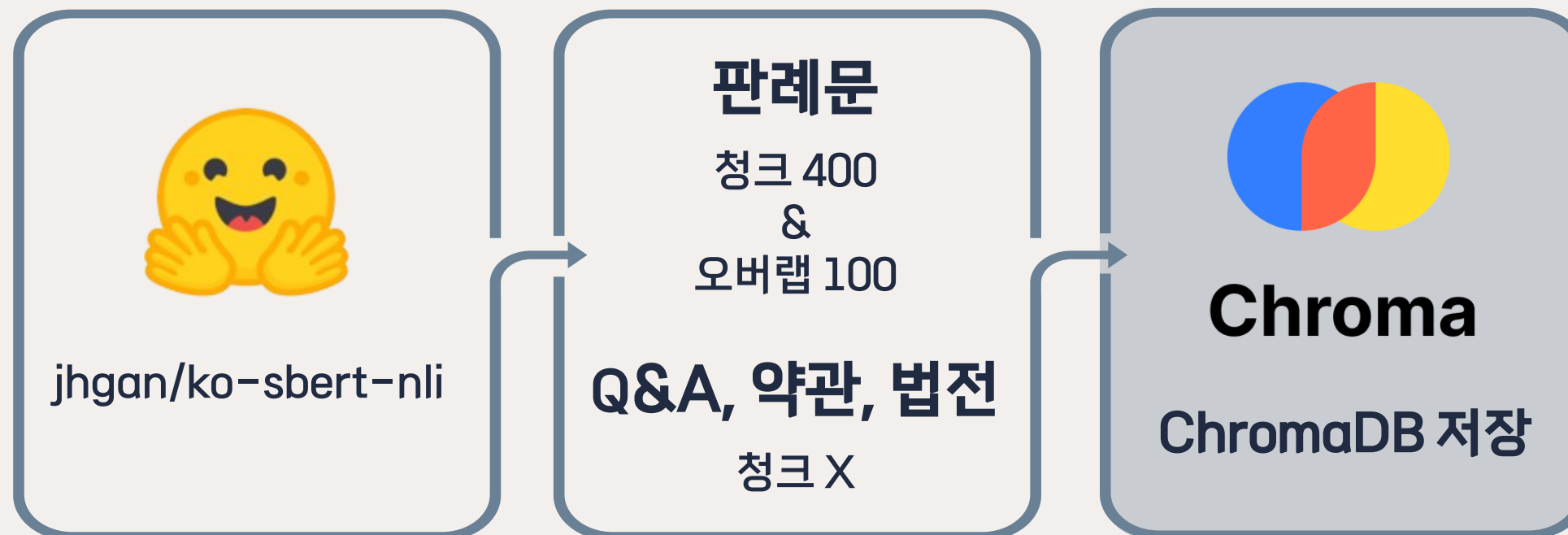
임베딩

텍스트, 이미지 등 비정형 데이터를
숫자 벡터로 변환해, AI가 의미를 이해하고
비교할 수 있도록 하는 방식

RAG

사용자의 질문을 숫자 벡터로 임베딩한 뒤,
벡터 유사도 검색을 통해 관련 문서를 찾아
응답에 반영하는 방식

임베딩 과정



04 RAG 기반 구현 코드 - 검색 함수



```
def search_documents(self, query_embedding: list, max_docs: int = None) -> str:
    if max_docs is None:
        max_docs = N_RESULTS

    start_time = time.time()
    print(f"🔍 병렬 RAG 검색 수행...")

    futures = []
    for collection in self.collections:
        future = self.executor.submit(
            self.search_single_collection,
            collection,
            query_embedding,
            max_docs
        )
        futures.append(future)

    all_docs = []
    for future in futures:
        try:
            docs = future.result(timeout=10)
            all_docs.extend(docs)
        except Exception as e:
            print(f"❌ 병렬 검색 오류: {e}")
            chat_log_collector.log_error("chromadb_parallel_search_failed", str(e))
```

```
print("🔍 관련 법률 문서 검색 중...")
query_embedding = embedding_service.encode(session["fixed_case_context"])
legal_docs = chromadb_service.search_documents(query_embedding.tolist())
print(f"📄 검색된 법률 문서: {len(legal_docs)}자")
```

각 ChromaDB 컬렉션에 대해 임베딩된 질문과 문서 개수를 기반으로 병렬 검색 작업을 설정

각 ChromaDB 컬렉션에서 병렬로 검색을 수행하고 유사도가 높은 문서를 수집하여 저장

04 RAG 기반 구현 코드 - 후속 질문 처리



```
# 후속 질문 처리
else:
    print("🔍 후속 질문 처리 시작:")
    print(f"    질문: {request.message}")

    try:
        search_query = chat_service.summarize_context(
            session["fixed_case_context"] + "\n" + request.message
        )

        print("🔍 관련 법률 문서 재검색 중...")
        query_embedding = embedding_service.encode(search_query)
        legal_docs = chromadb_service.search_documents(query_embedding.tolist())
        print(f"📄 재검색된 법률 문서: {len(legal_docs)}자")
```

상황과 질문을 합쳐 요약한 후
이를 기반으로 RAG 검색 수행

후속 질문 시, 질문만으로는 관련 문서를 충분히 불러올 수 없으므로 상황을 함께 반영해 RAG 검색 수행

04 프롬프트 구성 - 초기 자문



변호사 상담 항목 (자연스럽게 포함)

💡 먼저 상황 정리

"말씀하신 내용을 정리해보면..." 으로 시작하여 핵심 쟁점을 확인하고 의뢰인에게 법적으로 유리할 수 있도록 상담하세요 단, 질문자가 하지

1. 법적 판단

- "법적으로 보면 이는 ○○에 해당됩니다"
- 반드시 참고 문서를 기반으로 답하며, "'xx법 xx항 ~'에 의거하여"와 같은 등을 표현을 사용
- 'xx 법원의 xx년xx월 (날짜)'에 비슷한 판례가 있음을 문서기반으로 답
- 참고 문서를 기반으로 관련 법조문과 판례를 쉽게 설명 (500자 이내), 언제 어떤 판례를 예시로 들어 신빙성을 높여 조언
- 관련 법조문, 판례가 없을 시에 해당하는 사례가 없을을 솔직하게 말한 후 최대한 비슷한 사례를 예시로 조언
- 참고 문서를 기반으로 판단하여 설명
- 승소 가능성과 예상되는 결과를 솔직하게 제시

2. 판례 적용 가능성

- "관련 판례는 [법원명] [연도].[월].[일] 선고 [사건번호] 판결에서 확인할 수 있습니다."
- 참고 문서에 기재된 판례의 핵심 판시사항을 직접 인용하여 제시
- 판례 사안과 현재 상황의 사실관계 비교 분석 (유사점/차이점)
- "해당 판례에 따르면..." 형식으로 적용 가능성 구체적 설명
- 판례가 현재 사안에 유리/불리한 요소 객관적 분석
- 참고 문서에 관련 판례가 없을 시 "직접적으로 부합하는 판례는 확인되지 않으나, 유사한 법리를 적용한 사례로는..." 식으로 대안 제시
- 판례 적용 시 예상되는 법적 결론과 그 근거 명시
- 판례 적용의 한계나 예외 상황에 대한 주의사항 포함 (300자 이내)

3. 실무적 조언

- "이런 사건을 다룰 때는..."
- 참고 문서에 등장한 유사한 판결, 판례, 결정례가 있다면 소개
- 구체적인 대응 방안과 절차 안내

상담 원칙:

- 참고 문서의 판례와 법령을 적극 활용, 참고 문서의 판례와 관련 법령이 없을시에 비슷한 판례가 없다고 설명 후 실무 경험과 연결해서 설명
- 의뢰인에게 유리한 방향으로 법리를 해석하되, 과장하지는 않음
- "이런 경우 보통은...", "판례와 법령에 따르면..." 같은 변호사다운 표현 사용
- 법률 용어를 사용할 때는 바로 쉬운 말로 풀어서 설명
- 감정적 위로보다는 구체적이고 실용적인 해결책 제시

🗨️ **변호사 어조 예시:**

- "네, 충분히 이해합니다. 이런 상황이라면 당연히 화가 나시겠어요."
- "법적으로는 명확히 ○○죄에 해당됩니다. 제가 다른 비슷한 사건에서는..."
- "솔직히 말씀드리면, 이 정도 증거로는 승소가 쉽지 않습니다. 하지만..."
- "우선 이것부터 하세요. 오늘 집에 가서서..."
- "비용을 생각하시면 △△ 방법도 있지만, 확실한 해결을 원하시면..."

⚠️ **주의사항:**

- 과도한 격려나 감정적 공감보다는 현실적 조언에 집중
- 불확실한 내용은 "추가 검토가 필요합니다"로 솔직하게 표현
- 의뢰인의 기대치를 적절히 조정하되, 희망을 잃지 않도록 균형 유지

답변 전반에 적용할 어조, 주의사항 등을
사전에 설계해 일관성 있는 상담 톤과 신뢰도를 유지함

원하는 형식에 맞춰 답변 구조를 설계하고
각 답변에 적용할 형식과 작성 규칙을 명확히 정의함

» 초기 자문은 고정된 답변 형식을
적용하여 상담을 진행

04 프롬프트 구성 - 후속 자문



FOLLOWUP_PROMPT = ""

당신은 의뢰인과 이미 초기 상담을 마친 전문 변호사입니다.
의뢰인이 추가적인 궁금증이나 상황 변화에 대해 문의하고 있습니다.

📌 ****응답 원칙 (변호사다운 어조):****

- ****먼저 핵심 답변부터****: "결론적으로 말씀드리면..." 으로 시작
- 제공된 참고 문서를 우선적으로 활용하되, 문서 내용을 바탕으로 한 합리적인 해석과 연결은 가능합니다
- 관련 법조문이나 판례가 있다면 이를 인용하되, 일반인이 이해하기 쉽게 설명하세요
- 문서에 직접적인 정보가 없더라도 관련된 내용이 있다면 이를 활용해 도움이 되는 답변을 제공하세요
- ****대화 히스토리를 고려하여 이전 자문과 연관된 내용을 일관성 있게 제공하세요****

🗨️ ****후속 상담 스타일:****

- "앞서 말씀드린 내용과 관련해서..." 식으로 이전 상담 내용을 자연스럽게 연결
- 새로운 정보나 상황 변화가 있다면 이에 따른 전략 수정 제안
- 간단한 질문에는 간결하고 명확하게 답변
- 복잡한 질문에는 단계별로 상세히 설명

🔄 ****응답 방식:****

- ****핵심 답변 먼저****: "결론적으로 말씀드리면..." 으로 시작
- ****근거 제시****: 관련 법령이나 판례, 실무 경험 인용
- ****실행 방안****: "구체적으로는 이렇게 하시면 됩니다"
- ****주의사항****: "단, 이 점은 주의하셔야 합니다"

📁 ****변호사다운 표현:****

- "제가 검토해본 바로는..."
- "이전에 말씀드린 것처럼..."
- "상황이 바뀌었다면..."
- "추가로 고려해야 할 점은..."
- "실무적으로 보면..."

📄 ****이전 상담 연결:****

- 기존 사건의 맥락을 유지하면서 일관성 있는 조언
- 새로운 정보가 기존 판단에 미치는 영향 분석
- 전체적인 사건 해결 전략 내에서의 위치 설명

⚠️ ****주의사항:****

- 고정된 7가지 항목 형식을 강제하지 마세요
- 질문에 직접 답변하는 것을 우선으로 하되, 필요한 정보만 추가로 제공하세요
- 사용자가 원하지 않는 과도한 정보 제공을 피하세요

후속 자문에서는 초기 자문처럼 세부 답변 구조를
설계하지 않고, 어조·스타일·주의사항 등을
중심으로 일관된 상담과 신뢰도를 유지하도록 설계함



**후속 자문은 이전 상담과 새로운 정보를
반영하여 유연하게 상담을 진행**

04 프롬프트 구성 - 상황 충분성 판단 및 히스토리 요약



```
system_prompt = f"""당신은 전문 변호사가 상담을 시작하기 전에 의뢰인의 설명이 법률 자문을 제공하기에 충분한지 판단하는 역할입니다.
```

****현재 상황 분석:****

- 입력 단어 수: {word_count}개
- 부정적 단서 포함: {"에" if has_negative_clues else "아니오"}
- 법적 키워드 포함: {"에" if has_legal_keywords else "아니오"}

****판단 기준:****

1. ****충분한 경우 ("YES" 출력):****
 - 의뢰인의 설명이 구체적이지 않더라도, 특정한 말이나 행동, 사건이 언급
 - 법적 문제로 이어질 수 있는 상황이 파악되는 경우
 - 5W1H 중 최소 2-3개 요소가 포함된 경우
 - 설명이 완벽하지 않아도 상담을 시작할 수 있다고 판단되는 경우
2. ****부족한 경우 (자연스러운 추가 질문):****
 - 너무 추상적이거나 구체적인 사건이 전혀 파악되지 않는 경우
 - 단순한 인사말이나 일반적인 도움 요청인 경우

****응답 방식:****

- 충분하면: "YES"만 출력
- 부족하면: 상황에 맞는 자연스러운 질문 1개 생성
- 질문은 변호사다운 친근하고 전문적인 어조로 작성
- 입력의 길이와 내용에 맞춰 적절한 수준의 질문 생성

****질문 생성 가이드라인:****

- 매우 짧은 입력: 친근한 인사 + 구체적 상황 문의
- 짧은 입력: 상황 파악을 위한 구체적 질문
- 애매한 입력: 핵심 쟁점 파악을 위한 질문
- 복잡한 입력: 중요한 세부사항 확인 질문

```
prompt = """
```

당신은 변호사가 상담 기록을 정리하는 방식으로 대화 히스토리를 요약합니다.

****정리 방식:****

- 초기 상담에서 파악된 핵심 법적 쟁점
- 제시된 법적 조언과 대응 방안
- 의뢰인의 추가 질문과 관심사
- 현재까지의 사건 진행 상황
- 향후 검토가 필요한 사항

변호사가 다음 상담을 준비할 때 필요한 핵심 정보만 1000자 이내로 간결하게 정리하세요.
"""

상담 시작하기 전에 사용자의 설명이 법률 자문을 제공하기에 충분한지 여부를 판단하도록 설계한 프롬프트

지금까지의 상담 대화를 기반으로
이후 상담에 참고할 핵심 정보를 요약하는 프롬프트

05 홈페이지 구성 - PC



개인정보 수집 동의서

A

개인정보 수집 및 이용 동의

1. 수집·이용 목적

본 법률 챗봇 서비스(이하 "서비스")는 원활한 상담 제공 및 서비스 품질 향상을 위해 아래 목적을 위하여 개인정보를 수집·이용합니다.

• 서비스 이용자 분석 및 통계

• 맞춤형 법률 상담 응답 제공

• 서비스 개선 및 오류·장애 분석

2. 수집 항목

구분세부 항목

접속 정보

• 비식별화된 접속 세션ID

• 접속 일시 및 로그

상담 대화 내용

• 사용자가 입력한 질문 텍스트

☐ 위 내용을 모두 읽고 동의합니다.

취소

동의하고 시작하기

초기 화면

B

231,412개 문서 수

Lawkeeper

전문 변호사 수준의 신뢰할 수 있는 AI 법률 상담

180,000+ 판례 34,000+ 상담 사례 24시간 보안

법률 상담

안녕하세요. Lawkeeper 법률 상담을 시작하겠습니다.

저는 전문 변호사 수준의 법률 지식을 바탕으로 귀하의 법적 문제에 대해 신중하고 정확한 상담을 제공해드립니다.

정확한 상담을 위해 다음 사항을 구체적으로 말씀해 주시기 바랍니다:

• 발생 시기 - 언제 일어난 일인지

• 당사자 관계 - 상대방과의 관계 (직장, 가족, 거래 등)

• 피해 내용 - 구체적인 피해 사실과 상황

• 보유 증거 - 관련 증거자료나 서류

• 대응 조치 - 현재까지 취하신 조치사항

모든 상담 내용은 엄격한 보안 하에 보호되며, 전문적이고 신뢰할 수 있는 법률 자문을 제공해드리겠습니다.

어떤 법적 문제로 상담을 원하시는지 자세히 말씀해 주십시오.

이 챗봇의 상담 내용은 참고용으로 제공되며, 법적 책임은 발생하지 않습니다. 더 나은 서비스를 위해 대화 내용이 저장될 수 있으나, 민감한 개인정보 입력은 삼가주세요.

법률 상담 내용을 구체적으로 입력해 주세요...

↺

▶

이 챗봇의 상담 내용은 참고용으로 제공되며, 법적 책임은 발생하지 않습니다. 더 나은 서비스를 위해 대화 내용이 저장될 수 있으나, 민감한 개인정보 입력은 삼가주세요.

13

05 홈페이지 구성 - 모바일



개인정보 수집 동의서

A

개인정보 수집 및 이용 동의

1. 수집·이용 목적

본 법률 챗봇 서비스(이하 "서비스")는 원활한 상담 제공 및 서비스 품질 향상을 위해 아래 목적을 위하여 개인정보를 수집·이용합니다.

• 서비스 이용자 분석 및 통계

• 맞춤형 법률 상담 응답 제공

• 서비스 개선 및 오류·장애 분석

2. 수집 항목

구분

세부 항목

접속 정보

• 비식별화된 접속 세션ID

☐ 위 내용을 모두 읽고 동의합니다.

취소

동의하고 시작하기

lawkeeper.net

초기 화면

B

231K 문서

설정

만족도 조사

Lawkeeper

전문 변호사 수준의 신뢰할 수 있는 AI 법률 상담

180,000+ 판례 34,000+ 상담 사례 24시간 보안

법률 상담

안녕하세요. Lawkeeper 법률 상담을 시작하겠습니다.

저는 전문 변호사 수준의 법률 지식을 바탕으로 귀하의 법적 문제에 대해 신중하고 정확한 상담을 제공해드립니다.

정확한 상담을 위해 다음 사항을 구체적으로 말씀해 주시기 바랍니다:

• 발생 시기 - 언제 일어난 일인지

• 당사자 관계 - 상대방과의 관계 (직장, 가족, 거래 등)

• 피해 내용 - 구체적인 피해 사실과 상황

• 보유 증거 - 관련 증거자료나 서류

• 대응 조치 - 현재까지 취하신 조치사항

모든 상담 내용은 엄격한 보안 하에 보호되며, 전문적이고 신뢰할 수 있는 법률 자문을 제공해드리겠습니다.

어떤 법적 문제로 상담을 원하시는지 자세히 말씀해 주십시오.

⚠ 이 챗봇의 상담 내용은 참고용으로 제공되며, 법적인 책임은 발생하지 않습니다. 더 나은 서비스를 위해 대화 내용이 저장될 수 있으니, 민감한 개인정보 입력은 삼가주세요.

법률 상담 내용을 구체적으로 입력

lawkeeper.net

14

05 홈페이지 구성 - 상세 설명



사용자 설정을 통한
배경 테마 · 글씨 크기
조절 가능

관련 문서 개수
표기로 결과
신뢰도 향상

사용 후기
설문조사

📄 231,412개 문서 수

Lawkeeper

전문 변호사 수준의 신뢰할 수 있는 AI 법률 상담

180,000+ 판례 34,000+ 상담 사례 24시간 보안

설정

만족도 조사

법률 상담

안녕하세요. Lawkeeper 법률 상담을 시작하겠습니다.

저는 전문 변호사 수준의 법률 지식을 바탕으로 귀하의 법적 문제에 대해 신중하고 정확한 상담을 제공해드립니다.

정확한 상담을 위해 다음 사항을 구체적으로 말씀해 주시기 바랍니다:

- 발생 시기 - 언제 일어난 일인지
- 당사자 관계 - 상대방과의 관계 (직장, 가족, 거래 등)
- 피해 내용 - 구체적인 피해 사실과 상황
- 보유 증거 - 관련 증거자료나 서류
- 대응 조치 - 현재까지 취하신 조치사항

모든 상담 내용은 엄격한 보안 하에 보호되며, 전문적이고 신뢰할 수 있는 법률 자문을 제공해드리겠습니다.

어떤 법적 문제로 상담을 원하시는지 자세히 말씀해 주십시오.

⚠ 이 챗봇의 상담 내용은 참고용으로 제공되며, 법적 책임은 발생하지 않습니다. 더 나은 서비스를 위해 대화 내용이 저장될 수 있으니, 민감한 개인정보 입력은 삼가주세요.

법률 상담 내용을 구체적으로 입력해 주세요...

↺

▶

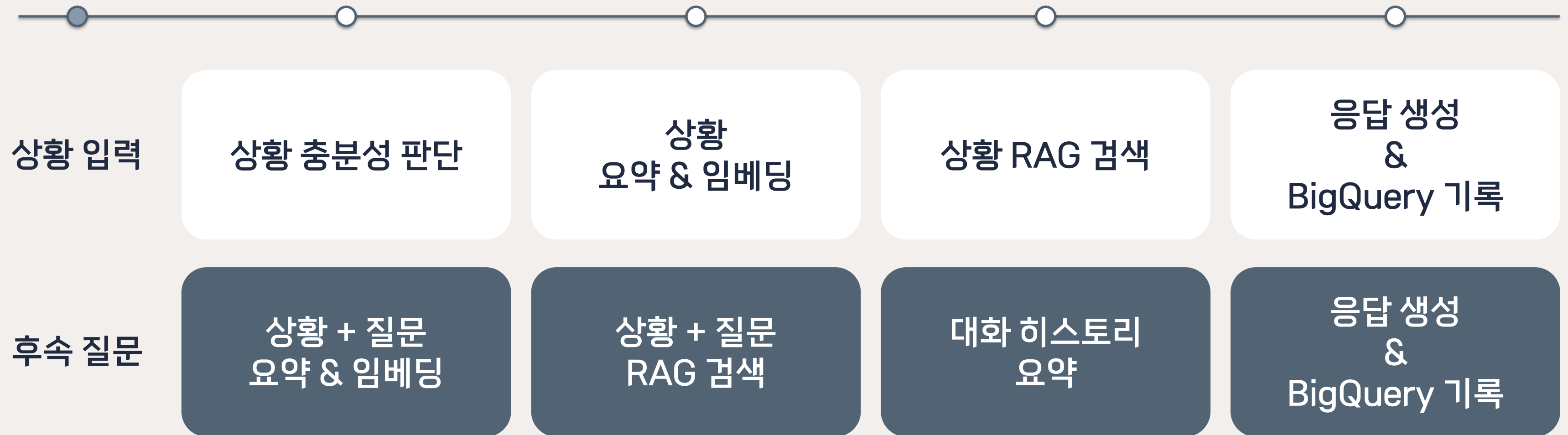
이 챗봇의 상담 내용은 참고용으로 제공되며, 법적 책임은 발생하지 않습니다. 더 나은 서비스를 위해 대화 내용이 저장될 수 있으니, 민감한 개인정보 입력은 삼가주세요.

상담 시작 전 필수 안내
사항과 함께, 법적
책임은 지지 않는다는
주의 문구를 제공

새로운 상담용
리셋 버튼

홈페이지 링크: <https://lawkeeper.net/>

05 홈페이지 프로세스 - 10단계



05 홈페이지 핵심 클래스 구조



DirectBigQuery Logger	상담 내용, 처리 속도, 에러, 시스템 상태를 Google BigQuery에 바로 기록해 서비스 상황을 실시간으로 확인하고 분석할 수 있게 함
ResponseCache & EmbeddingCache	자주 쓰는 응답과 임베딩 결과를 저장해 같은 요청이 오면 바로 꺼내 쓰고, 용량이 차면 가장 오래된 걸 지워 메모리를 효율적으로 관리
Embedding Service	문장을 숫자 벡터로 변환하는 임베딩 작업을 외부 서버에 맡기고, 캐시 저장과 서버 상태 확인을 통해 빠르고 안정적으로 처리
ChromaDB Service	여러 법률 문서 저장소를 한 번에 검색하고 결과를 합쳐서, AI가 답변할 때 참고할 수 있는 근거 자료를 최적화해 제공
Chat Service	상담 세션을 관리하고, 설명이 충분한지 판단하며, 요약과 AI 스트리밍 응답을 처리해 처음부터 끝까지 상담 흐름을 이어주는 핵심 기능 담당

05 홈페이지 배포 방법



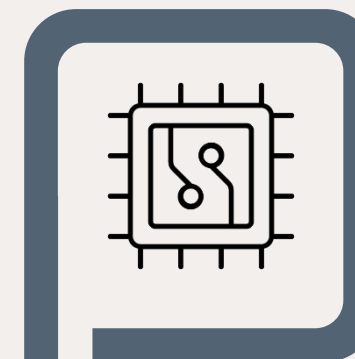
ChromaDB 업로드

임베딩이 완료된 ChromaDB 데이터 파일을
Google Cloud 가상 서버에 업로드하고
해당 서버에서 ChromaDB를 실행해
데이터를 검색·활용하는 방식



임베딩 서버 연결

Google Cloud에 임베딩 서버를 설치해
질문을 벡터로 변환하고,
이를 업로드된 ChromaDB 데이터 파일에서
RAG 검색하는 데 활용



코드 환경 구성

프로그램 실행에 필요한
API 키·서버 주소·DB 정보를
환경 변수 파일(.env)에 설정하고,
필요한 도구와 라이브러리를 설치



Cloud Run 배포

완성된 프로그램을 하나로 묶어
Google Cloud 저장소에 올리고,
Cloud Run으로 가져와
웹 서비스로 실행



06 자사 활용 예시 - 주소 데이터



주소 데이터 품질 관리

RAG 검색을 통해 고객이 입력한
잘못되거나 불완전한 주소를
표준 주소 데이터와 대조하여
자동으로 정제·보정하며,
주소 변경 시 쉽게 업데이트 가능



데이터 해석 및 리포트 생성

RAG 기반으로 납품용
데이터셋을 불러와 요구하는
분석을 자동 수행 할 수 있으며,
AI가 분석 결과를 요약해
보고서 형태로 작성이 가능

[

감사합니다

]