프로젝트 기획서

과정명	LG U+ Why Not SW Camp 67			
작성일자	2024 년 05월 07일	팀명	요리조리	
참여인원	유승환, 주수진	팀장	이지헌	
프로젝트명	개인 맞춤형 i	레시피 추천 서비:	스 기획 및 설계	
기 간		5월 7일 ~ 2025년	년 5월 9일	
조직구성	유승환 : 사업기획, 데이터 분석, . 이지헌 : 사업기획, 데이터 분석, .			
<u> </u>	· 하시면 : 사업기획, 데이터 분석, . · 주수진 : 사업기획, 데이터 분석, .			
	개인 맞춤형 서비스 증가			
	● 개인 맞춤형 식품 시장 전밍	†		
	글로벌 맞춤형		시장 규모	
	(단위: 억달러, 자료:	글로벌데이터)	05	
			35	
	9			
프로젝트(분석)				
배경	2021	년	2030년	
	그림 1 영양식품 시장 규모(출처: 아시아경제)			
		종착역으로 여기	겨지는 분야로 개인 맞춤형 식품이 -	
	집힘 ○ 최근 식품산업에 인공지능과 빅데이터 등을 결합하여 다양한 소비자 유형되다이터를 분석해 개인의 필요에 따라 선택할 수 있는 제품을 제조하고 최천, 판매할 수 있게 됨 ○ 소비자는 세분화된 기성품을 선택하거나 설문조사 응답을 통해 나에게 최적화된 음식과 건강기능식품을 선택할 수 있게 됨 ○ 2021년 기준 9억 달러(약 1조2300억원) 수준이던 글로벌 개인 맞춤형 영영			
	식품 시장 규모는 2030년에는 35억 달러(약 4조8000억원)로 300% 가까이			
	성장할 전망을 띄고 있는 상황임			

● 롯데 멤버스 2024 트렌드 인사이트



그림 2 2024 엘포인트 트렌드 인사이트 보고서(출처: 롯데 엘포인트)



그림 3 2024 엘포인트 트렌드 인사이트 보고서(출처: 롯데 엘포인트)

- 약 1500만명의 2024년 소비 데이터를 분석한 결과 '개인 맞춤형 경험 소비' 가 소비자들에게 각광받고 있음
- 피부타입, 성분, 기능 등 개인의 특성을 반영한 스킨케어 제품 구매는 전년 대비 25% 증가, 두피·헤어(19.7%)와 바디(17.4%)와 같이 세분화된 맞춤형 케어 제품의 구매율 증가
- 혈당 관련 제품(146.5%), 멘탈 케어 제품(97.4%) 증가 등 건강 관리 제품에 서 개인 맞춤형 케어 제품 구매율 증가
- 다양한 분야에서 개인화가 적용되고 있음. 따라서 식품업종에서도 개인 맞춤형 서비스를 기대해 볼 수 있음
- 건강한 삶을 중시하는 MZ세대를 중심으로 저속노화 식단이 유행하였으며, 잘파세대의 저속노화식단 관련 제품 구매율은 16% 증가하는 큰 폭의 성장 세를 보여줌
- 저지방, 고단백, 항산화 등 건강 식품과 관련한 SNS 언급량도 올해 9월 전 년 동기 대비 44.4% 증가

식품 분야 AI 기술 접목 사례 증가

- 삼성 비스포크 AI 키친
 - 삼성저자의 스마트 냉장고 '비스포크 AI 패밀리허브'는 AI 비전 인사이드 기능을 통해 냉장고에 보관중인 식재료 바탕의 레시피를 추천하는 기능을 제공
 - 식재료 자동인식 및 입고일 기록을 통한 푸드 리스트 자동 생성을 통한 식 재료 관리 지원
 - 사물인터넷 스마트싱스로 연결된 오븐/정수기 등 다양한 가전제품과의 연동 및 음성 명령을 통해 조리 지원
- 누비랩 AI 푸드 스캐너
 - AI 푸드 스캐너를 통해 음식의 종류와 양을 사진을 통해 실시간으로 식별 및 분석 및 전산화
 - 음식의 양에 따른 인식된 음식의 영양성분 분석
 - 알러지 및 제약 식단 지원
 - 분석된 섭취량 및 잔반량은 데이터베이스에 누적되며, 급식 서비스 등에서 활용

레시피 관련 서비스 현황 및 규모

- 레시피 앱 시장 성장
 - 레시피 앱 시장 규모는 2022년 42억 달러로 평가되었으며, 2030년까지 20 억 달러에 달할 것으로 예상
 - 건강에 대한 관심 증가와 다양한 요리 경험 탐구를 희망하는 소비자 증가에 따라 단순 레시피 뿐 아니라 식사 계획, 식이 추적 및 영양 지침에 대한 탐색을 요구함.
 - 식품 산업의 지속적 디지털 혁신으로 인해 더욱 강화
 - 미국 노동 통계국 연구에 따르면, 식품 준비 서비스에 대한 수요는 2021년 에서 2031년까지 7% 증가를 예상
- 레시피 컨텐츠 붐
 - 요리에 대한 관심이 높아지며 집에서 직접 음식을 해보고자 하는 수요 증가
 - 2024년 10월 g마켓 조사에 따르면 조리용 강판, 칼 거래액이 각각 233%, 167% 급증할 정도로 요리를 시도하는 수요가 늘어난 것으로 분석

AI 챗봇과 사용자 정보를 통한 개인 맞춤형 레시피 추천 서비스 제공

- 식품 산업에서 인공지능(AI)와 빅데이터를 결합한 개인 맞춤형 서비스 및 제품 이 증가함에 따라 요리를 좋아하고 즐기는 대상들을 위한 맞춤형 레시피 추 천 서비스를 개발하고자 함
- 본 서비스는 기본적으로 사용자가 설정하는 필터링값과 AI 챗봇과의 대화를 통해 사용자에게 친숙하고 재미있으면서도 최적화된 레시피를 추천하고자 함
- 사용자 정보(필터링)를 통해 일반 식단 섭취자 뿐 아니라 비건, 저염식섭취자, 다이어터 등 다양한 특이식단 섭취 사용자에게도 편리하고 최적화된 레시피 추천 서비스 제공

프로젝트 목적

방대한 레시피 콘텐츠 통합 검색 기능을 통해 더욱 다양하고 세분화 된 레시피 제공

- 기존 레시피를 탐색하던 사용자들은 유튜브, 네이버 등 하나의 채널에 접속해 해당 채널 내의 레시피만 추천 받을 수 있었으나 본 서비스 사용 시 서비스 채널 안에서 유튜브, 네이버 등 다양한 채널의 레시피를 종합적으로 추천 받 을 수 있음
- 사용자는 기존보다 넓은 범주에서 레시피를 추천 받을 수 있어 다양한 레시피 수집 가능
- 특이식단 섭취 사용자에게도 편리하고 최적화된 레시피 추천 서비스 제공
- 서비스 플로우
 - 시작화면
 - 접속화면: 1)로그인

2)ID·PW 찾기

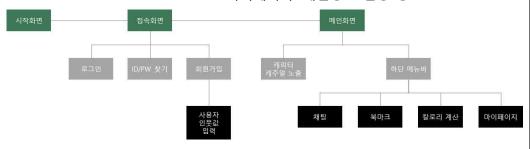
3)회원가입 - 사용자 INPUT

○ 메인화면: 1)캐릭터 캐주얼 노출 - 서비스 대표 캐릭터 노출

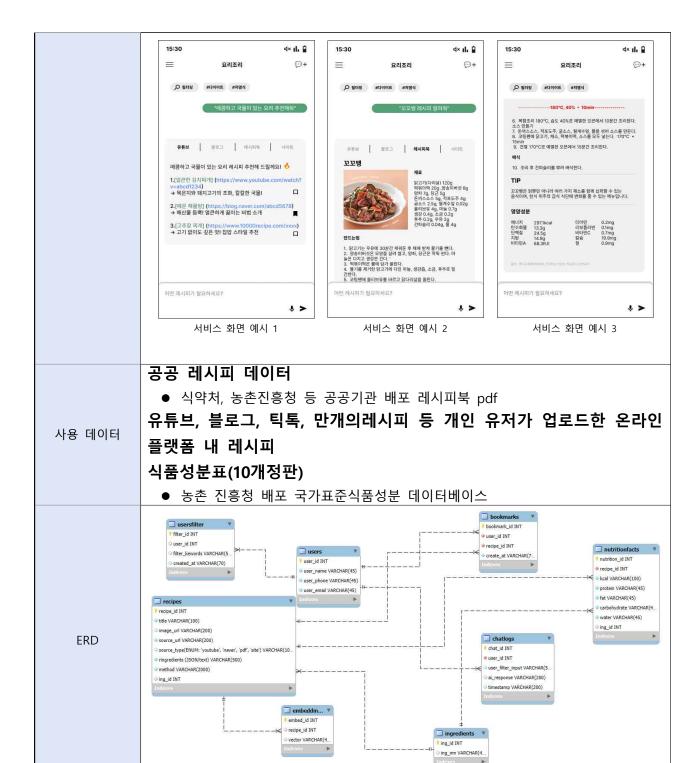
2)하단 메뉴바 - 채팅: 챗봇과의 채팅 화면

- 북마크: 마음에 들었던 레시피 기능
- 칼로리 계산: 사용자 입력 식재료의 칼로리 및영양성분 예측 계산
- 마이페이지: 개인정보 변경 등

서비스 플로우 및 서비스 화면 기획



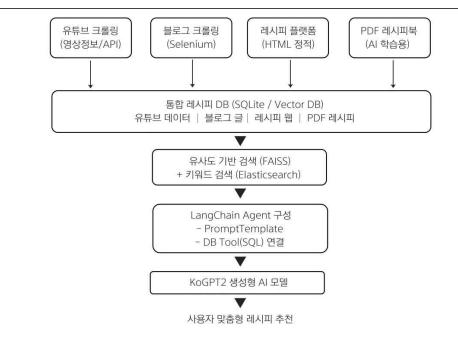
- 서비스 사용 화면(예시)
 - 유튜브, 블로그, 사이트 선택 시 서비스 화면 예시 1과 같이 링크로 레시피 추천
 - 레시피북 선택 시 서비스 화면 예시 2, 3과 같이 이미지와 텍스트로 레시피를 추천된 레시피를 볼 수 있음



서비스 프로세스 : 크롤링 + 검색엔진 + 생성형 AI 조합

데이터 활용 및

분석 방법



크롤링

● 출처별 흐름 구조

단계	설명
1단계	검색 결과나 목록 페이지에서 제목/링크 추출
2단계	각 링크를 방문해서 상세 페이지 접속
3단계	본문 내용, 레시피 설명, 재료 등을 추출
4단계	정제 후 DB에 저장

- 출처별 크롤링 방식(비동기 처리 방식)
 - 비동기 처리 방식
 - 속도 향상
 - CPU 효율 증가
 - 다수 작업 자동화
- 주요 기술 설명
 - YOUTUBE
 - 사용 기술 : 1) YouTube Data API v3, 2) aiohttp
 - ☐ YouTube Data API
 - ▶ 유튜브 공식 API로 검색어 기반 영상 정보(제목, 이미지, 설명, 링크등)를 JSON으로 제공
 - □ aiohttp
 - ▶ 여러 쿼리 동시에 처리 가능
 - ▶ python에서 비동기를 쓸 때 필요한 기술로 HTTP 요청을 비동기로 처리할 수 있는 라이브러리
 - 네이버 블로그
 - 사용 기술 : 1) Selenium, 2) aiohttp or request + BeautifulSoup
 - ☐ Selenium
 - ▶ 자바스크립트로 렌더링되는 블로그 페이지를 브라우저처럼 조작하며 크롤링
 - ▶ 특정 제목 내 페이지 내용까지 확인 한 후 크롤링 정보를 db에 삽입

- ☐ aiohttp or request + BeautifulSoup
 - ▶ 네이버 API 또는 단순 검색 결과 페이지를 처리할 때 사용
- 레시피 사이트
 - 사용 기술 : 1) request + BeautifulSoup, 2) aiohttp
 - ☐ request + BeautifulSoup
 - ▶ 정적 웹페이지를 요청하고 HTML을 파싱하여 레시피 데이터 추출
 - □ aiohttp
 - ▶ 여러 페이지를 병렬적으로 수집
- PDF 레시피북(공공기관)
 - 사용 기술: 1) PyMuPDF, pdfminer.six, 2) os, glob
 - ☐ PyMuPDF, pdfminer.six
 - ☐ os, glob
 - ▶ 폴더 내 여러 PDF를 자동으로 불러오기
 - ▶ 반복 처리 자동화
 - ▶ 자동으로 모든 pdf 순회하며 텍스트 추출

하이브리드 검색 엔진

● 하이브리드 검색 흐름

- 텍스트 임베딩(Text Embedding)
 - KoBERT: 한국어에 최적화된 BERT 모델로, 텍스트를 벡터로 임베딩
 - 유사도 기반 검색(Similarity-Based Search)
 - FAISS
 - ▶ Vector DB
 - ▶ Facebook AI Research에서 개발한 벡터 검색을 위한 라이브러리로, 대규모 데이터셋에 대해 빠르고 효율적인 검색을 수행
 - ▶ 주로 유사도 기반 검색에 사용
- 하이브리드 검색 방식
 - 벡터 검색과 키워드 검색 방식을 결합하여 사용자가 의도한 검색 결과를 정확하고 넓은 범위에서 얻을 수 있음
 - 벡터 검색
 - 의미 기반의 유사도 검색을 통해 사용자가 의도한 의미를 파악하여, 정확 한 결과를 제공
 - 예를 들어, "건강한 닭가슴살 요리"와 같은 쿼리에서 "단백질이 풍부한 레 시피" 같은 유사한 레시피 추출 가능

- 사용 모델
 - ☐ FAISS
 - ▶ 대규모 벡터 검색을 위한 최적화된 라이브러리로, 벡터 간 유사도를 계산하여 빠르게 검색 결과를 반환
 - ▶ 쿼리 임베딩 후 벡터 DB(FAISS)에 저장된 레시피 텍스트 임베딩들과 코사인 유사도 비교
- 키워드 검색
 - 사용자가 입력한 정확한 키워드가 포함된 레시피를 탐색
 - 예를 들어 "닭가슴살"이라는 단어가 정확히 포함된 레시피를 빠르게 찾아 낼 수 있음
 - 사용 모델
 - ☐ BM25
 - ▶ 정보 검색에서 많이 사용되는 빈도 기반 검색 알고리즘으로, 검색 쿼리와 문서 간의 단어 빈도와 문서 길이를 고려하여 가중치를 부여
 - ☐ Elasticsearch
 - ▶ 풀 텍스트 검색을 위한 엔진으로, BM25 알고리즘을 포함한 다양한 검색 방식을 지원하며, 효율적인 텍스트 매칭을 제공
- 하이브리드 결합
 - 가중합 방식으로 결합

final_score = α * vector_score + $(1 - \alpha)$ * keyword_score

- **vector_score** = 벡터 검색 결과(의미 기반 유사도 점수)
- **keyword_score** = 키워드 검색 결과(텍스트 일치 기반 점수)
- **α** = 가중치 하이퍼파라미터(0.0 ~ 1.0 사이 값)
- ex) α = 0.7 → 벡터 검색 결과를 더 중요하게 반영

생성형 AI 모델

- 대화형 AI 모델 : CLOVA X
 - 네이버 초대규모(Hyperscale) 언어모델인 HyperCLOVA X 기술을 바탕으로 만들어진 대화형 에이전트
 - HyperCLOVA X API 기반으로 사용 가능
 - 선정 이유
 - 한국어 자연어 이해에 강함 (NAVER 학습 기반)
 - 한국 사용자 대상 서비스에서 활용도 높음
 - Langchain과 쉽게 연동 가능 (LLM wrapper 설정만 하면 됨)
 - Clova X 기본 구조 (기능 중심 관점)

```
[ 사용자 입력 ]
↓
[ Clova X Frontend ]
↓
[ Prompt Engine + Context Manager ]
↓
[ NAVER HyperCLOVA X 모델 (LLM) ]
↓
[ 응답 생성 + 통 호출 (필요 시) ]
↓
[ 사용자에게 결과 전달 ]
```

- 통합 파이프라인 구성 LangChain
 - LLM을 이용해 어플리케이션을 만들 수 있게 하는 오픈소스 프레임워크
 - Python 기반 라이브러리
 - 여러 LLM 을 체인 형태로 연결할 수 있어 Clova X 를 연결해 챗봇을 만드는 현재 프로젝트에 적합
 - 기본 구조

구성 요소	설명	
LLM	KoGPT처럼 응답을 생성하는 모델	
Prompt Template	LLM에 넣을 질문 형식	
Cahins	LLM + 입력/출력 흐름 연결	
Tools/Agents	DB 검색, 웹 검색 같은 기능 확장	

● Clova X + Langchain 기반 확장형 레시피 챗봇 구조

○ 전체 시스템 구조

```
사용자 질문
↓
[Langchain Prompt Template + Tool 관리]
↓
Clova X (질문 이해 + 응답 생성)
↓
[Tool: Vector DB / 외부 API / DB 검색]
↓
Relevant 정보 → Clova X 재입력
↓
최종 응답 (자연어)
```

○ 구성 요소

구성 요소	설명		
Clava V	생성형 AI (LLM), 사용자 질문 분석		
Clova X	및 응답 생성 담당		
Lammahain	Clova X를 중심으로 도구(검색/DB/API) 연결		
Langchain	및 파이프라인 제어		
Vector DB (예: FAISS, Chroma)	레시피, 요리 설명 등의 의미 기반 검색 지원		
외부 API (Naver Blog, YouTube)	관련 링크 수집		
DB (예: PostgreSQL)	영양소 정보, 사용자 저장 링크 등 저장/조회		
	사용자가 원하는 레시피를 저장		
북마크 기능	(DB에 insert/update)		

○ 동작 흐름 예시

```
사용자: "단백질 높은 요리 뭐 있어?"

LangChain → Clova X 전달

↓
Clova X → 질문 요약: "단백질 많은 요리 검색"

↓
LangChain Tool → Vector DB or SQL 질의

↓

유사 레시피 목록 반환

↓
Clova X → 자연어 응답 생성

→ "닭가슴살 샐러드, 두부 스테이크 등을 추천드려요!"

↓
사용자: "두부 스테이크 북마크해줘"

↓
LangChain Tool → 북마크 DB 저장

↓
"두부 스테이크가 북마크에 저장되었습니다!"
```

- Clova X + Langchain 기반 챗봇의 장점
 - - 한국어 최적화 LLM인 Clova X 사용 : 한국 사용자 대상 서비스에 적합

- - Langchain 연결 기능 : Vector, SQL DB, 웹 API 등 다양한 "tool" 연결, 조합해 고도화된 챗봇 제작 가능
- - 기억 기반의 대화 흐름 관리 : Langchain의 **ConversationChain**, **Memory 객체** 활용하면 사용자의 질문 맥락을 유지하며 자연스러운 대화 가능

SQL

- 레시피 데이터, 사용자 데이터, 추천 기록 등을 저장/조회하는 DB 내부에서 사용
- 사용 방법
 - 레시피 저장/조회
 - 태그/제목 기반 필터링
 - 크롤링한 결과 저장
 - 크롤링 데이터 누적 저장
 - 사용자 추천 기록 저장
 - 사용자 요청에 따른 조건 검색
 - 통계 데이터 분석
 - 어떤 레시피가 많이 추천되었는지 등

.

Strenghts - 강점

- AI 기반 맞춤형 레시피 추천
- 일반식 뿐만 아니라 특이식단 (비건, 다이어터) 레시피까지 넓은 범주의 추천 서비스 제공
- 공공 레시피 데이터, 유튜브, 블로그 등다양한 형태의 레시피 컨텐츠 추천

Weaknesses - 약점

- W
- 외부 콘텐츠 사용 시 저작권 및 제휴 문제 발생 가능성
- AI 추천 정확도에 대한 신뢰도 확보 어려움
 서비스 초기에는 학습 데이터 및 피드백
- 서비스 초기에는 학습 데이터 및 피드백 부족 문제가 있음

SWOT 분석

Opportunities - 기회



- 제휴를 통한 API 확보 및 비즈니스 연계를 통한 서비스 확장
- 1인 가구와 맞벌이 가정 증가로 간편 요리수요 확대
- 개인 맞춤형 서비스 및 헬스케어 시장 성장 : 맞춤형 식단 앱 수요 증가 (다이어트, 당뇨, 채식 등)

Threats - 위기



- 레시피 관련 경쟁 서비스 다수 존재
- 사용자가 일회용으로 사용 할 가능성

일반 사용자용 레시피 챗봇

건강/식단 관리 용이

서비스 활용 방안

- 필터링 및 공공기관 제공 신뢰도 높은 레시피 추천을 통해 건강에 관심이 많은 특이식단 사용자에게 최적화된 사용 경험 제공
 - 공공기관 제공 저염/건강식 레시피 추천
 - 저탄고지/다이어트/비건 등의 필터링 설정
- 이유식, 유아식 등을 만들어야 하는 육아가 처음인 부모에게 최적화된 사용 경험 제공

요리 초보자를 위한 튜터 기능

● 쉬운 레시피 필터링, 조리도구 기반 추천 기능 제공

ox) "요리 처음인데 쉬운 레시피 추천해줘, 전자레인지로 만들 수 있는 요리 추천해줘" 등

브랜드/푸드 마케팅 제휴 활용

- 식료품, 조리 기구 등 마켓 입점(쿠팡, 개인몰 등)
- 스마트 가전에 서비스 연동(삼성 비스포크 가전 등)

B2B 활용

- 식단 기준 + 재료 제한 조건에 따른 레시피 추천 서비스 제공
- 단체급식 업소 서비스 제공을 통한 잔반 문제 서포트
- 서비스 운영 총 예상 비용(월 기준)

○ 사용자 1만 기준 : 약 547,500원

○ 사용자 5만 기준 : 2,147,500원

○ 사용자 10만 기준 : 4,230,000원

- IOS 앱 개발(연)
- \$99(약 120,000원)
- AWS 클라우드 이용(월)
 - 유저 1만 기준 예상 비용

구성 요소	설명	예상 비용	
10 44	20	(USD)	
EC2 (FAISS 실행)	벡터 검색 서버	\$47.96	
OpenSearch	 텍스트 키워드 검색 서버	¢62.F6	
(Elasticsearch)	텍으트 기뤄트 점역 시미 	\$62.56	
EBS 스토리지	인덱스 및 데이터 저장	\$20.00	
4데이터 전송 비용	100GB 초과 시 과금	\$9.00	
기타 (백업, 로깅 등)		\$5.00	
총합계	약 \$145.52(약 200,000원)		

○ 유저 5만 기준 예상 비용

구성 요소	설명	예상 비용 (USD)
FAISS 서버 (EC2)	I데이드 (8GB RAM)	약 \$70.00
OpenSearch 노드	t3.medium.search 2대 이상 필요	약 \$120.00
스토리지 (EBS)	벡터 및 인덱스 합산 500~800GB 가정	약 \$50.00
데이터 전송 비용	900GB 초과 전송 (100GB 무료) × \$0.09	약 \$72.00
기타 비용	모니터링, 백업 등	약 \$10.00
총합계	약 \$322.00(약 450,000원)	

○ 유저 10만 기준 예상 비용

구성 요소	설명	예상 비용
18 #1	20	(USD)
	m5.xlarge 인스턴스 (16GB	
FAISS 서버 (EC2)	RAM, 벡터 200만 이상 처리	약 \$140.00
	용)	
OpenSearch (3노드)	t3.medium.search × 3 또는	약 \$180.00
Opensearch (3±=)	m5.large × 2	¬ \$100.00
스토리지 (EBS)	벡터 + 인덱스 데이터 1TB	약 \$100.00
(LD3)	가정	3100.00
데이터 전송 비용	1.9TB 초과 (2TB - 100GB 무	약 \$171.00
	료) × \$0.09	
기타 (백업, 모니터링)	CloudWatch, 스냅샷, 로깅	약 \$15.00
총합계	약 \$606.00(약 845,000원)	

비용

- CLOVA X 서비스 이용
 - 모델 : 기본 HCX-005 기준
 - 사용자 수 1만 규모 예상 비용 (월)

구성 요소	설명
입력 토큰 수	평균 20 토큰 가정
출력 토큰 수	평균 40 토큰 가정
토큰 총 합계	5 (일 평균 질문 수) * (20+40) * 10,000 * 30(일)
월 예상 비용(원화)	약 337,500원

○ 사용자 수 5만 규모 예상 비용 (월)

구성 요소	설명
입력 토큰 수	평균 20 토큰 가정
출력 토큰 수	평균 40 토큰 가정
토큰 총 합계	5 (일 평균 질문 수) * (20+40) * 50,000 * 30(일)
월 예상 비용(원화)	약 1,687,500원

○ 사용자 수 10만 규모 예상 비용 (월)

구성 요소	설명
입력 토큰 수	평균 20 토큰 가정
출력 토큰 수	평균 40 토큰 가정
토큰 총 합계	5 (일 평균 질문 수) * (20+40) * 100000 * 30(일)
월 예상 비용(원화)	약 3,375,000원

추진 일정(8주)

구분	구분 상세 내용		
		1주차	2주차
	서비스 기획		
기획 및 설계	기술 스택 선정		
717 % = 1	데이터 출처 조사		
	폴더 및 프로젝트 구조 구성		
	유튜브 API 수집 모듈		
	블로그 크롤링		
데이터 수집 및 저장	레시피 사이트 비동기 크롤링		
	PDF 텍스트 추출		
	정제 및 저장		
	SBERT 임베딩 처리 및 적용		
AI 추천 모델 및 검색	FAISS 구축		
엔진 구현 및 테스트	질문 임베딩 + 유사 레시피 추출		
	성능 점검		
	LangChain / KOGPT연동		
생성형 응답 구현	생성형 응답 테스트		
808 5H TE	형용사 필터 적용		
	성능 점검		
	사용자 입력 → 임베딩 → 검색 →		
	응답 흐름 연결		
	추천된 레시피 링크, 북마크, 필터		
앱 개발	등 UI 구성		
	플랫폼(모바일/웹) 기반 프론트		
	화면 제작 및 연결		
	앱 프레임워크 반영 가능		
	기능 테스트		
테스트 및 배포	데이터 보완		
	배포		

일정