**논문 관련 서비스 개발 프로젝트 기획서**

**서민정(팀장), 박규택, 이준경, 진윤화**

## **1. 프로젝트 개요**

| 작성일자 | 2024년 11월 11일 |
| --- | --- |
| 프로젝트 기간 | 2024년 11월 8일 ~ 2025년 1월 3일 |
| 팀 구성 | - 서민정 (프로젝트 매니저)  - 박규택 (AI/문서정리)  - 이준경 (AI/백엔드 개발자)  - 진윤화 (AI/프론트 개발자) |
| 프로젝트 목표 | 실무자들이 AI 활용에 있어서 논문을 효과적으로 검색하고 이해하는 데 도움을 줄 수 있는 AI 기반 논문 추천 및 질문 답변 챗봇 개발 |

## 

## **2. 프로젝트 배경**

### **2.1 프로젝트 배경 상세**

최근 다양한 업무에서 AI를 활용하는 사례가 늘어나고 있다. AI 관련 연구자뿐만 아니라 다양한 업무에 AI를 활용하여 업무 효율성을 향상시키고 있다. 제조업, 금융, 유통, 게임 업계에서 AI를 활용하고 있다. [1]



AI가 빠르게 발전하는 분야이기 때문에 최신 연구와 기술을 파악하는 것이 필수적이다. 논문에서는 AI의 새로운 알고리즘, 데이터 처리 방법, 응용 분야 등이 검증된 형식으로 소개된다. 따라서 최신 기술과 연구를 따라가기 위해 AI 학술 논문을 읽는 것은 합리적인 방법이다.

하지만 새내기 연구자나 관련 지식이 없는 사람에게 논문 검색이 어렵게 다가올 수 있다. 필요한 논문을 선별하는 능력이 부족할 뿐만 아니라, 최신 연구인지에 대한 판단도 불가능하다.

AI 학술 논문을 읽기 어려운 이유는 무엇일까? 우선, 전문 용어의 사용이 많고, 고도의 전문 지식을 요구하며, 밀도 높은 정보와 복잡한 구조를 가지고 있다. 또한 연구 방법론이 복잡하다. 따라서 이러한 어려움을 극복하기 위해서는 해당 분야의 기본 지식을 갖추고, 지속적으로 논문을 읽는 것이 필요하다. 그렇지만 논문이 익숙하지 않은 사람들에게는 논문에 익숙해지기 위해 많은 시간과 노력이 필요하다. 이를 도와주기 위하여 논문 탐색을 위한 AI 챗봇을 개발하고자 한다.

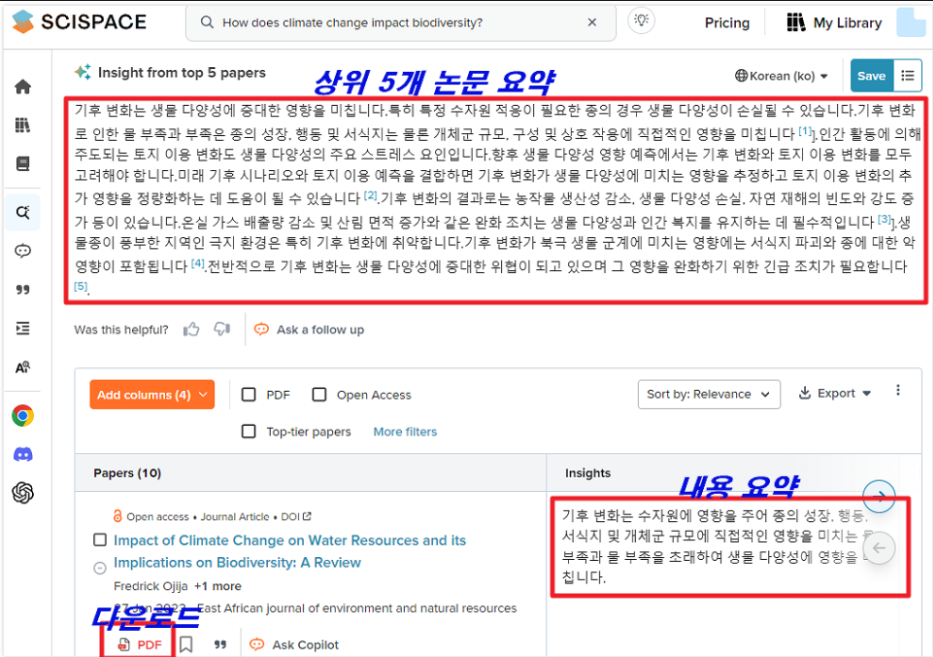
논문 탐색을 위한 AI 챗봇과 유사한 서비스를 제공하는 업체들을 조사하였다. 조사결과, 유사 서비스를 제공하는 업체로 SciSpace AI, R Discovery, DBpia를 확인할 수 있었다.

### 

### **2.2. 경쟁업체 조사**

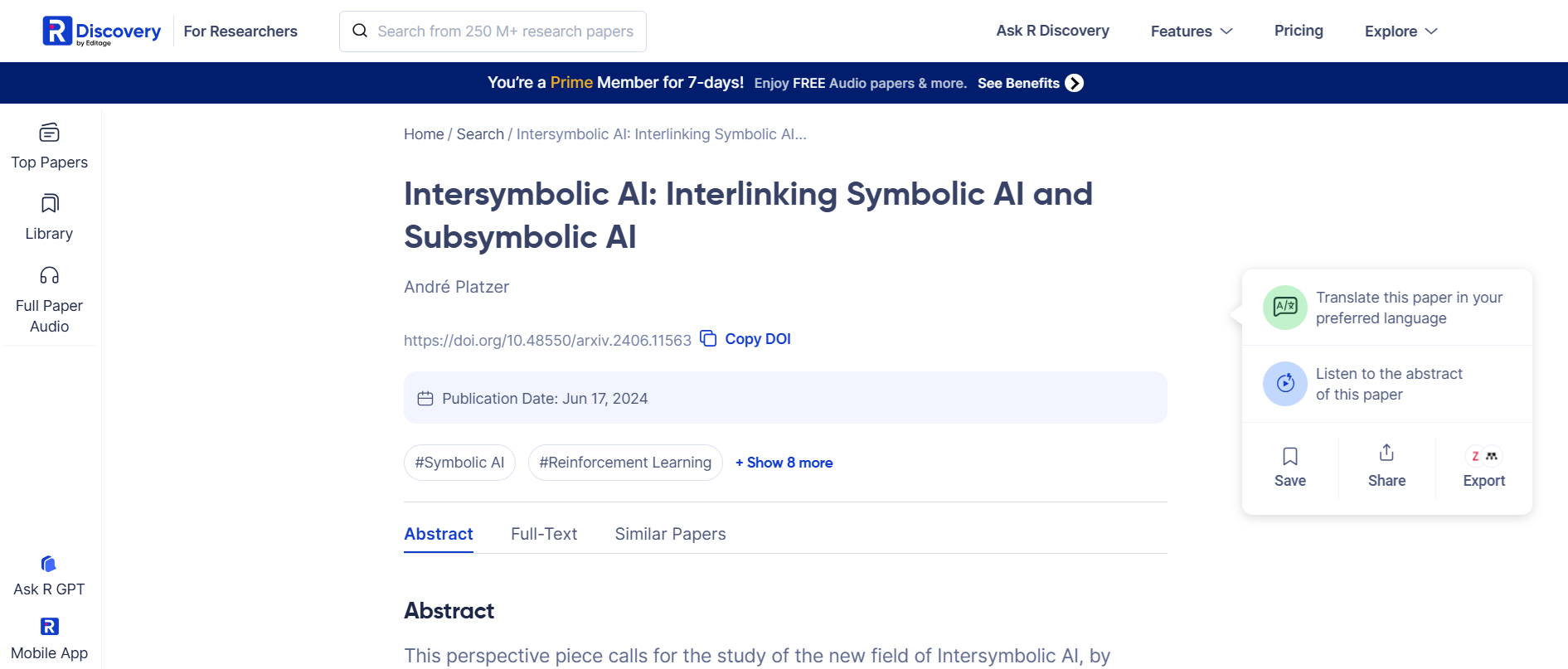
#### **2.2.1 SciSpace AI**

우선 SciSpace AI라는 업체를 조사해 보았다. SciSpacesms 연구자와 학생을 대상으로 서비스를 제공하고 있다. 논문 검색 기능뿐만 아니라, PDF 파일이나 논문 초록을 AI가 요약/분석하는 기능을 제공한다다. 검색된 논문의 상세 페이지에 들어가지 않아도, 검색 결과 리스트에서 AI가 요약한 정보를 확인할 수 있어 원하는 논문을 빠르게 찾을 수 있다. [2]



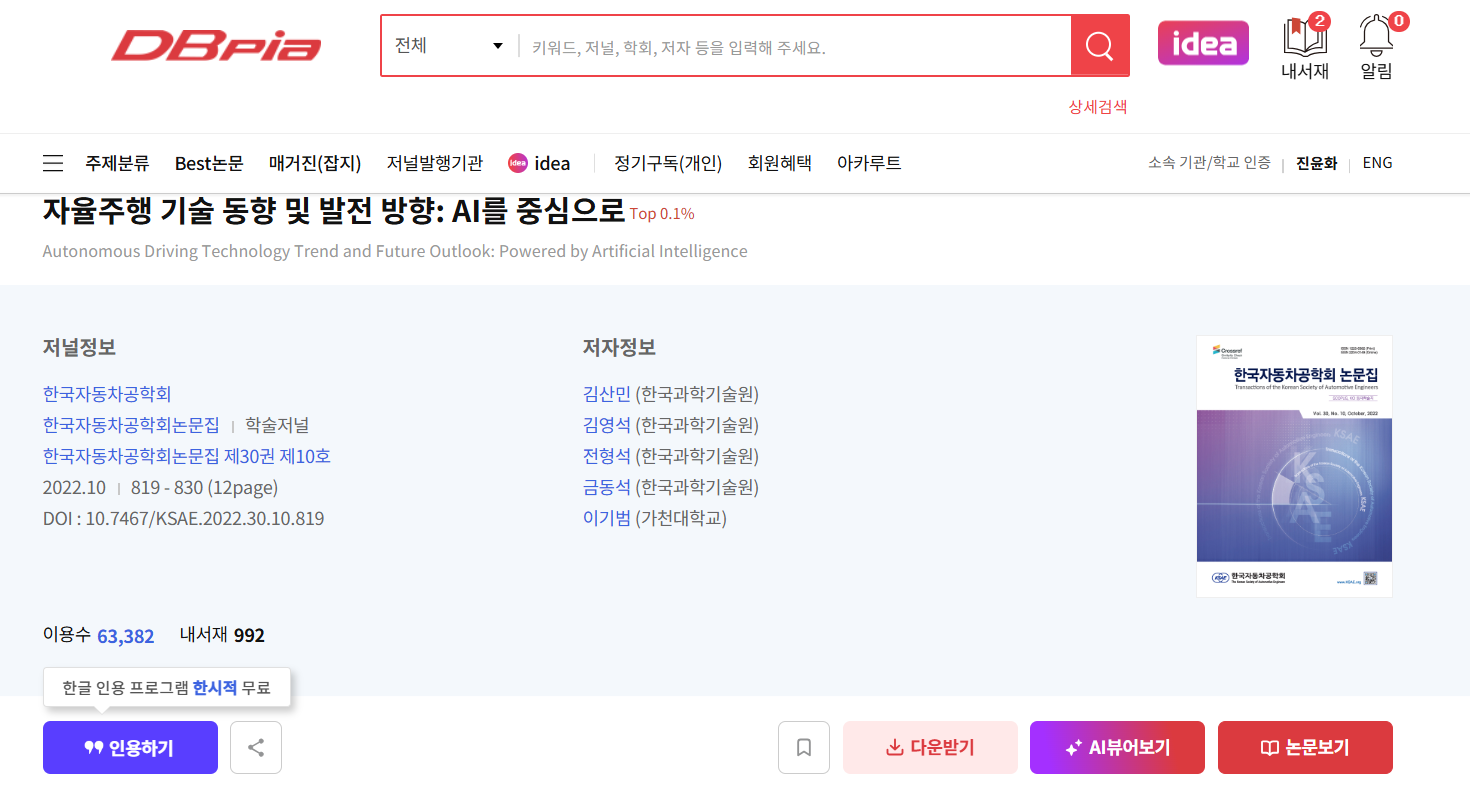
#### **2.2.2 R discovery**

R discovery는 인공지능을 활용하여 전공별, 연구 주제별 최신 논문을 추천하고, 내용을 정리해주는 애플리케이션이다. 논문 추천, 검색 결과 요약, 논문 초록 분석 등의 기능을 제공하며, 애플리케이션을 다운받지 않더라도 웹사이트를 통해활용할 수 있다. [3]



#### **2.2.3 DBpia**

마지막으로 DBpia는 한국의 대표적인 학술정보 검색 서비스로, 학술 논문, 학위 논문, 학술지, 컨퍼런스 자료등을 다양하게 검색할 수 있다. AI를 활용하여 연구주제, 연구배경, 연구방법, 연구결과 등을 요약하는 기능을 제공하고 있다. [4]



[1] <https://www.samsungsds.com/kr/insights/ai_use_cases.html>

[2] <https://typeset.io/>

[3] <https://discovery.researcher.life/>

[4] <https://www.dbpia.co.kr/>

### **2.3. 기존 업체와의 차별성**

이러한 (3과같은) 기존 서비스들과의 차별화를 위해 DocuMento는 신기술 서베이를 통해 빠르게 기술을 습득해야 하는 현업 전문가들을 타겟으로 설정하였다. 각자의 분야에서 최신 기술 동향 파악이 필요하지만, 학술 논문에 대한 경험이 부족한 실무자들은 학술 용어와 개념을 이해하는데 어려움을 겪을 수 있으며, 산업 용어와 학술 용어 간 격차를 줄여주는 접근이 필요하다. 뿐만 아니라 해당 논문의 신뢰성을 판단하는 것 또한 어려운 일이고, 최신 논문을 지속적으로 확인해야 하는 것 또한 업무 효율성을 저해하는 요소이다. 이러한 문제들을을 해결하는 것이 다른 경쟁 업체들과 차별화된 DocuMento의 기능들이고, 각 기능별로 상세하게 설명하고자 한다.

#### **2.3.1 기능1. 키워드 추천 서비스**

키워드 추천 서비스는 사용자가 특정 프로젝트나 기술 리서치를 위해 자신이 찾고자 하는 기술 주제와 관련된 논문을 찾고자할때 제공되는 서비스이다. 사용자가 입력한 초기 키워드를 기반으로 관련된 키워드를 추천해주고, 추천된 키워드를 기반으로 하여 논문 목록 및 주요 키워드를 제안한다. 제안된 내용을 사용자의 피드백을 통해 상호작용하여 보다 정밀한 키워드를 지속적으로 업데이트한다. 사용자는 맞춤형 키워드와 관련 논문 목록을 통해 원하는 기술에 대한 깊이 있는 논문을 빠르게 찾고, 각 논문의 핵심적인 방법론 또한 제시받아 실무 적용 가능성을 검토할 수 있다.

#### **2.3.2 기능2. 신뢰성 판별 서비스**

신뢰성 판별 서비스는 사용자가 프로젝트 의사결정에 필요한 논문의 신뢰성 및 품질을 평가하고자할때 유용한 서비스이다. 신뢰성 판별 서비스는 논문의 저자 학위, 논문 게재 저널의 Impact Factor, 인용 횟수 등을 고려하여 신뢰성 순으로로 정렬된 후 제공된다. 사용자는 제공된 신뢰성 점수를 바탕으로 논문의 객관적인 품질을 판단하고 신뢰할 수 있는 정보인지 검토하여 프로젝트에 활용할지를 선택할 수 있다.

#### **2.3.3 기능3. 최신 논문 알림 서비스**

최신 논문 알림 서비스는 사용자가 지속적으로 기술 트렌드를 모니터링해야 하거나 연구 프로젝트의 최신 기술 자료를 필요로 할때 유용한 기능이다. 사용자가 관심을 가진 기술 분야에 새로 추가된 논문을 Slack 혹은 이메일을 통해 알림을 주며, 실시간으로 최신 논문을 확인하여 필요한 정보를 획득하여 업무 효율성을 높일 수 있다.

#### **2.3.4 기능4. 선행 논문 추천 서비스**

선행 논문 추천 서비스는 사용자가 특정 논문을 통해 새로운 기술이나 연구 방법에 대하여 배경 지식이 부족할때 유용한 기능이다. 사용자가 이해하고자하는 논문과 유사성이 높은 선행 논문 n개를 추천해, 사용자가 논문의 배경 지식을 단계적으로 쌓아갈 수 있도록 지원한다. 각 선행 논문에는 간략한 요약을 제공하며, 추천된 논문이 현재 논문에 대해서 어떻게 기여하는지 설명하여 사전 지식 학습의 효율성을 높인다.

## **3. 프로젝트 목표 및 기술사항**

### **3.1 프로젝트 목표**

### **3.1.1 주요 목표**

| 논문 데이터  베이스 구축 | 다양한 학술 논문을 포함한 데이터베이스 구축 |
| --- | --- |
| 논문 검색 및  추천 시스템 | 사용자의 질의에 맞춰 관련 논문을 추천하는 시스템 |
| 질문-답변  시스템 | 사용자가 논문 내용에 대해 자연어로 질문하고 답변을 받을 수 있는 기능 |
| 추천 알고리즘  개발 | 사용자 취향 및 연구 분야를 기반으로 맞춤형 논문 추천 |

### **3.1.2 측정 가능한 목표**

1. 사용자 질문에 대한 80% 이상의 정확도

2. 첫 3개월 내에 500명 이상의 사용자 확보

3. 논문 추천 시스템에서 90% 이상의 관련성 정확도 달성

## **4. 요구사항**

### **4.1 탐색경로 제공**

| 기술 | 랭그래프 (프롬프트 설정) | |
| --- | --- | --- |
| 상세 요구사항 | 사용자가 필요로 하는 논문을 모를때, 여러 가능성의 탐색경로를 제시하여 탐색을 시작할 수 있도록 도움 | |
| 사용 대상 | 본인이 어떤 논문을 필요로 하는지 정확히 모르는 사람 | |
| 구현 방법 | 프롬프트 설정 | 사용자가 입력한 주제에 맞는 키워드를 통해 논문 탐색을 위한 경로를 제공 |
| 랭그래프 | - 유사한 연구 분야나 관련된 키워드를 기반으로 탐색 경로를 제안  - 사용자가 더 구체적인 주제를 찾을 수 있도록 유도 |

### **4.2 실제 논문 탐색**

| 기술 | 랭그래프 | |
| --- | --- | --- |
| 상세 요구사항 | 사용자가 필요로 하는 논문을 정확히 알고 있을때, 관련 논문을 직접 탐색할 수 있는 기능을 제공 | |
| 사용 대상 | - 찾고자 하는 논문을 정확히 알고있는 사람  - 탐색경로챗봇의 결과를 받은 사람 | |
| 구현 방법 | 랭그래프 | 사용자가 입력한 주제에 맞는 키워드를 통해 논문 탐색을 위한 경로를 제공 |
| 사용자가 입력한 키워드 또는 논문 제목을 바탕으로 정확한 논문을 찾아서 결과를 제공 | |

### **4.3 유사 논문 추천**

| 기술 | RAG (Retrieval-Augmented Generation) + Abstract | |
| --- | --- | --- |
| 상세 요구사항 | 사용자가 찾은 논문과 유사한 논문들을 추천하여, 추가적인 참고자료를 제공할 수 있도록 합니다. | |
| 사용 대상 | 추가적인 유사 논문에 대한 정보가 필요한 사람 | |
| 구현 방법 | RAG | 사용자가 입력한 주제에 맞는 키워드를 통해 논문 탐색을 위한 경로를 제공 |
| Abstract 적용 | 유사 논문에 대한 핵심 요약을 제공해, 사용자가 빠르게 논문 내용을 확인할 수 있도록 합니다. |

### **4.4 논문 파악하기 (요약)**

| 기술 | 랭그래프 + 벡터DB + Summarization | |
| --- | --- | --- |
| 상세 요구사항 | 사용자가 논문의 핵심 내용을 빠르게 파악할 수 있도록 논문을 요약하여 제공합니다. | |
| 사용 대상 | 빠르게 논문의 핵심 내용을 파악하고 싶은 사람 | |
| 구현 방법 | 랭그래프 +  벡터DB | 논문 내에서 주요 내용을 추출하고, 요약할 정보를 벡터화하여 빠르게 핵심 내용을 파악할 수 있도록 합니다. |
| Summarization | 논문의 핵심 내용을 자동으로 요약하는 알고리즘을 통해 사용자가 간단히 논문의 내용을 이해할 수 있도록 돕습니다 |

### **4.5 논문 내부 탐색**

| 기술 | RAG + 벡터DB (Subsection을 통한 구현) | |
| --- | --- | --- |
| 상세 요구사항 | 논문 내에서 특정 부분을 인용하거나 필요한 내용을 빠르게 찾을 수 있도록 하는 기능을 제공합니다. | |
| 사용 대상 | 논문 인용에서 필요한 부분 파악하고 싶은 사람, 논문에서 필요한 부분 파악하고 싶은 사람 | |
| 구현 방법 | RAG +  벡터DB | 논문을 다양한 섹션으로 나누어 각 섹션을 벡터화하여, 필요한 정보를 검색할 수 있도록 합니다. |
| Subsection  을 통한 구현 | 논문의 특정 부분(예: 서론, 방법론, 결론)을 세분화하여, 사용자가 필요한 정보를 빠르게 찾을 수 있도록 합니다 |

### **4.6 논문 이해 지원**

| 기술 | RAG + 벡터DB (Subsection을 통한 구현) | |
| --- | --- | --- |
| 상세 요구사항 | 사용자가 논문의 이해 정도를 확인하고, 부족한 부분을 보완할 수 있도록 돕는 기능입니다. | |
| 사용 대상 | 논문의 이해 정도를 확인하고 싶은 사람 | |
| 구현 방법 | 랭그래프 | 사용자가 논문 내용에 대해 질문을 하면, 논문 내의 관련된 정보나 중요한 개념을 제공하여 이해도를 높입니다. |
| RAG | 논문 내용에 대한 추가적인 설명을 제공하거나, 어려운 부분에 대해 보충 설명을 하여 이해도를 높입니다. |

### **4.7 실제 논문 구현 코드 (예시)**

| 기술 | 랭그래프 + Agent | |
| --- | --- | --- |
| 상세 요구사항 | 논문에서 설명된 이론이나 방법론을 실제로 구현한 코드 예시를 제공하여, 사용자가 논문의 적용 예시를 쉽게 이해할 수 있도록 합니다. | |
| 사용 대상 | 논문의 적용 예시가 필요한 경우 | |
| 구현 방법 | 랭그래프 +  Agent | 논문에서 설명된 방법론을 기반으로, 구현 가능한 코드 예시를 제공하는 기능을 챗봇이 담당합니다. |

## **5. 프로젝트 범위**

| 논문 데이터 수집 및 처리 | 논문 데이터베이스 구축, PDF 파일에서 텍스트 추출, 논문 메타데이터 수집 |
| --- | --- |
| 챗봇 개발 | 사용자 질문을 처리할 수 있는 자연어 처리(NLP) 기반 챗봇 개발. |
| 검색 알고리즘 | 논문 검색 기능 및 추천 시스템 개발. |
| UI/UX 디자인 | 직관적인 사용자 인터페이스 설계. |
| 성능 최적화 | 챗봇의 응답 시간 최적화, 정확도 개선. |

## 