

---

# **DABS** **(Daily AI Briefing Service)** project plan

---

Ver 1.2

팀명 : 지방이

# Table of Contents

1. 프로젝트 제목	3
2. 필요성 및 배경	3
3. 목표	3
4. 팀별 역할	3
5. 요구사항 분석 및 정의	4
5.1 자료 조사	4
5.2 경쟁사 조사	4
5.3 요구사항 정의	4
6. 서비스 설계	5
6.1 서비스 개념 및 구조도	5
6.2 UX 설계	6
6.3 데이터 설계	6
6.4 프로그램 모듈 구조, 모델 및 알고리즘 설계( Tool 구현)	7
7. 실험 및 평가	7
7.1 논문 요약 성능	7
7.2 뉴스 기사 요약 성능	8
7.3 블로그 요약 성능	8
7.4 종합 결과	8
8. 구현	8
8.1 메인 페이지 및 결과	8
8.2 Chat Bot	9
9. 결과	9
10. 감사의 글	10
11. 레퍼런스	10

1. 프로젝트 제목

DABS, Daily AI Briefing Service

사용자 입력 키워드의 동의어를 확장하여 관련 논문, 블로그, 기사를 탐색 및 요약 후 RAG 챗봇으로 질의응답까지 지원하는 콘텐츠 리서치 자동화 서비스

2. 필요성 및 배경

2.1 키워드 다양성 확보의 한계

사용자가 입력한 키워드만으로는 관련 자료를 충분히 탐색하기 어려움

2.2 AI 기반 자동화 기술의 발전

자연어 처리 및 생성형 AI 기술의 발전으로 의미 기반의 키워드 확장, 콘텐츠 요약, 문맥 검색 자동화 가능

2.3 기존 리서치 도구의 한계

대부분의 기존 검색 도구는 정확히 일치하는 키워드에 의존

3. 목표

이 서비스는 방대한 정보 속에서 필요한 자료를 찾기 어려운 리서치 시장을 공략하여, 키워드 의미 확장과 자동 요약을 통해 수작업 정보 탐색의 비효율성을 개선하고, 콘텐츠 조사 및 분석 업무의 생산성을 극대화하는 것을 목표로 합니다.

4. 팀별 역할

이름	역할	산출물
진현	조장, 웹 개발, 챗봇 개발	개발 코드
송명재	AI Agent 개발, 기사, 논문 요약 개발	개발 코드
이정렬	데이터 수집 및 전처리, 블로그 요약 개발	개발 코드

## 5. 요구사항 분석 및 정의

### 5.1 자료 조사

AI 기반 요약, 키워드 확장, 문서 클러스터링 등의 기술이 점차 대중화

ChatGPT, Perplexity, Feedly 등 정보 탐색 도구는 있으나, 키워드 확장 기반 수집+요약 기능은 제한적

미디어 업계는 빠른 트렌드 분석과 기획 지원 툴에 대한 수요가 높아짐

### 5.2 경쟁사 조사

서비스	특징	한계점
Feedly	뉴스/블로그 RSS 수집	요약 없음, 키워드 확장 불가
Perplexity	AI 기반 Q&A 탐색	실시간 뉴스/블로그 추적 불가
Notion AI	문서 요약 기능	콘텐츠 수집 기능 부재

### 5.3 요구사항 정의

#### 5.3.1 주 사용자(고객)

미디어 콘텐츠 기획자

- 뉴스, 유튜브, 블로그 등 콘텐츠 기획 시 아이디어 리서치를 빠르게 하고자 하는 사용자

마케팅/브랜드 전략 담당자

- 경쟁사 동향, 시장 트렌드를 파악해 기획 방향을 설정하고자 하는 실무자

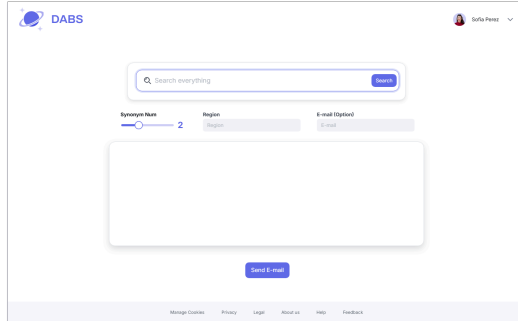
#### 5.3.2 요구사항

1. 사용자 요청에서 키워드를 추출하여 키워드 동의어 자동 생성
2. 확장된 키워드를 기반으로 논문, 블로그, 뉴스 형식의 콘텐츠를 실시간으로 수집
3. 수집된 콘텐츠는 AI 기반으로 자동 요약 및 번역 및 정리
4. 결과는 사용자의 목적에 따라 뉴스 중심 / 블로그 중심 / 학술 중심으로 검색
5. 사용자는 정리된 내용을 PDF 형식으로 메일을 통해 수신
6. 사용자는 정리된 내용 내에서 질의응답이 가능한 챗봇 사용

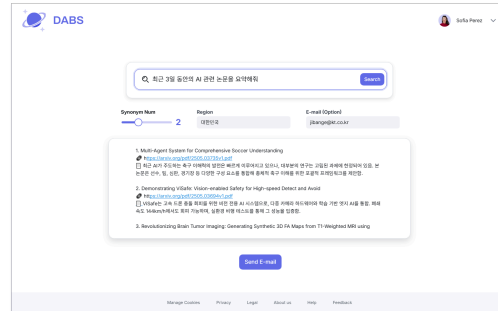
## 6.1 서비스 개념 및 구조도

## 6.2 UX 설계

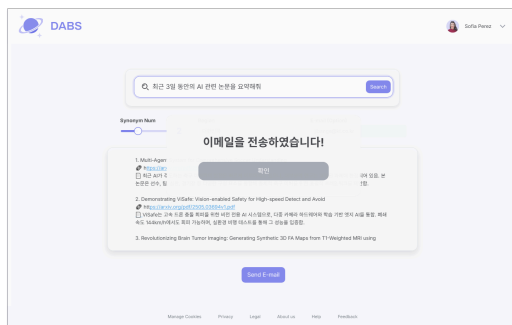
### 1. 초기화면



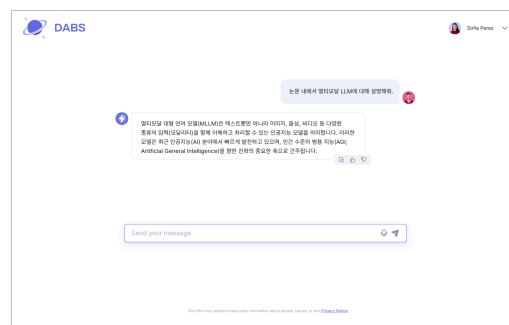
### 2. 사용자 프롬프트 입력 후 출력



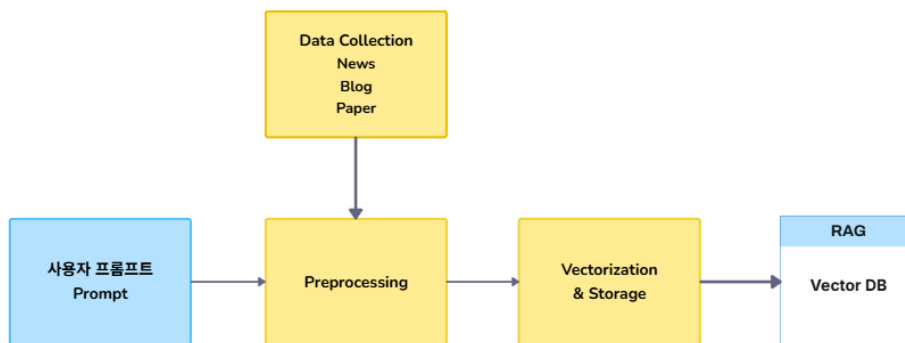
### 3. 요약 및 정리 내용 PDF 이메일 전송



### 4. RAG 기반 요약 내용 챗봇 기능

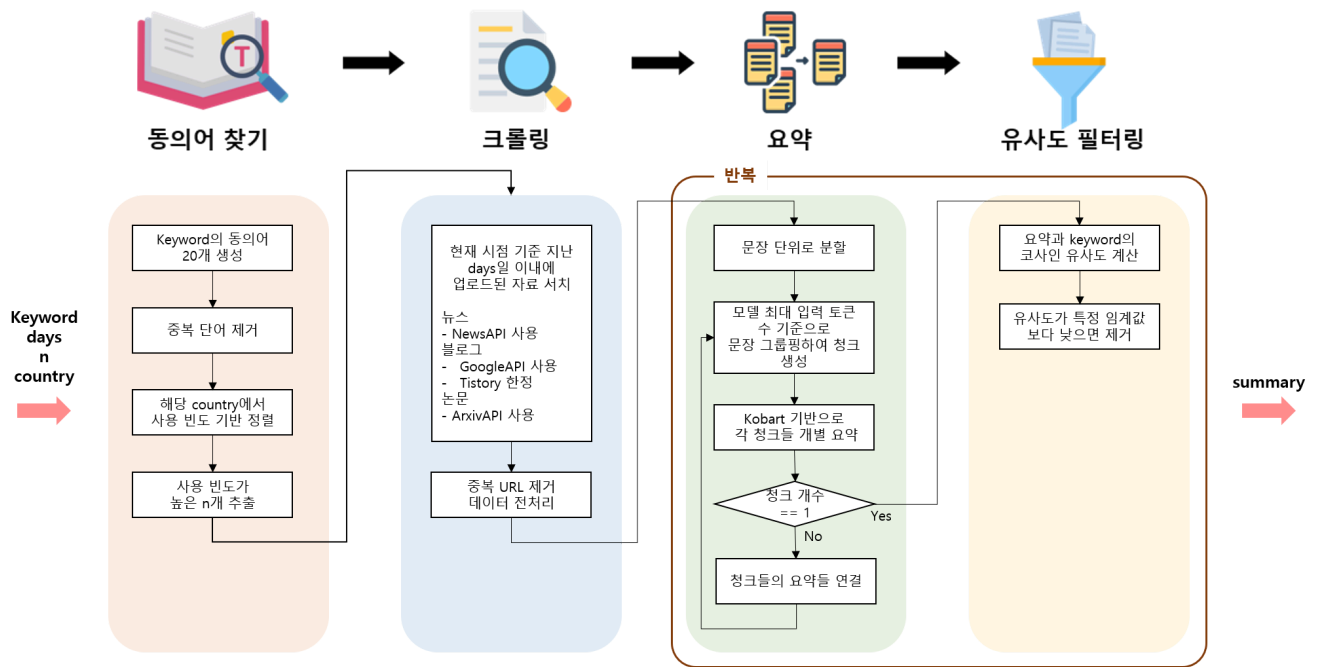


## 6.3 데이터 설계



사용자 프롬프트로부터 추출한 조건에 따라 뉴스·블로그·논문 데이터를 수집하고, 전처리 후 벡터화하여 벡터DB에 저장함으로써 RAG 기반 검색이 가능하도록 구성합니다.

## 6.4 프로그램 모듈 구조, 모델 및 알고리즘 설계( Tool 구현)



## 7. 실험 및 평가

본 연구에서는 콘텐츠 유형별 요약 성능을 평가하기 위해 ROUGE 지표의 F1 스코어를 중심으로 분석을 수행하였습니다. 평가 대상은 논문, 뉴스 기사, 블로그로 구분하였으며, ROUGE-1, ROUGE-2, ROUGE-L의 F1 스코어를 기준으로 성능을 비교하였습니다.

### 7.1 논문 요약 성능

논문 요약의 F1 스코어는 ROUGE-1 기준 0.0325, ROUGE-2는 0.0012, ROUGE-L은 0.0305로 매우 낮은 수치를 보였습니다. 이는 요약 시스템이 일부 정보를 정확히 추출하는 데는 성공했으나, 전체 문서의 핵심 내용을 포괄하는 데에는 한계가 있었음을 시사합니다. 논문의 복잡한 구조와 고밀도의 정보량이 요약 성능에 영향을 준 것으로 판단됩니다.

## 7.2 뉴스 기사 요약 성능

뉴스 기사 요약은 세 지표 모두에서 상대적으로 높은 F1 스코어를 보였습니다. ROUGE-1은 0.2669, ROUGE-2는 0.1813, ROUGE-L은 0.2482로 측정되었습니다. 이는 뉴스 콘텐츠가 핵심 정보가 명확히 구성되어 있고 문장 구조가 정형화되어 있어 요약에 유리하게 작용했음을 보여줍니다.

## 7.3 블로그 요약 성능

블로그 요약은 ROUGE-1 기준 F1 스코어 0.1339, ROUGE-2는 0.0813, ROUGE-L은 0.1196으로 확인되었습니다. 뉴스 기사보다는 낮지만 논문보다는 높은 성능을 보였으며, 이는 블로그 콘텐츠의 비정형적인 구성과 서술 방식으로 인해 요약 시스템이 일관된 성능을 내기 어려웠기 때문으로 보입니다.

## 7.4 종합 결과

전체적으로 뉴스 기사 콘텐츠에서 가장 우수한 요약 성능을 확인하였으며, 논문은 정보 밀도가 높고 블로그는 형식이 자유로워 상대적으로 낮은 성능을 보였습니다. 콘텐츠 유형에 따라 요약 전략을 달리 적용하는 것이 효과적일 것으로 판단됩니다.

# 8. 구현

## 8.1 메인 페이지 및 결과

Daily AI Briefing Assistant

프로프트

최근 2일간 인공지능 관련 뉴스 요약해줘

국가 선택

Korea

검색 범위

1

2

3

4

5

이메일 (선택)

b2b22@kcb2b.co.kr

실행

작업이 완료되었습니다!

종료 시간: 326초

\*\*1. [미국 AMD, 1분기 매출 10.3조원...전년비 36% ↑](https://www.inews24.com/view/1841403)\*\*

◦ 미국 정부의 기업 AMD가 데이터 센터 수요 급증에 힘입어 전년동기 대비 36% 성장한 매출을 기록했습니다. 시장 예상치를 상회한 실적을 보였습니다. [기사 바로가기](#)

2. [엑스지텔스케어 'AI', 미국 식품의약국 승인 신청](#)

◦ 엑스지텔스케어가 미국 식품의약국(FDA) 승인을 신청했으며 글로벌 시장에서의 입지를 확대할 계획입니다. [기사 바로가기](#)

3. ["AI로 고용 창출 효과"...IBM, AI 에이전트 구축 솔루션으로 주도권 잡기 박자](#)

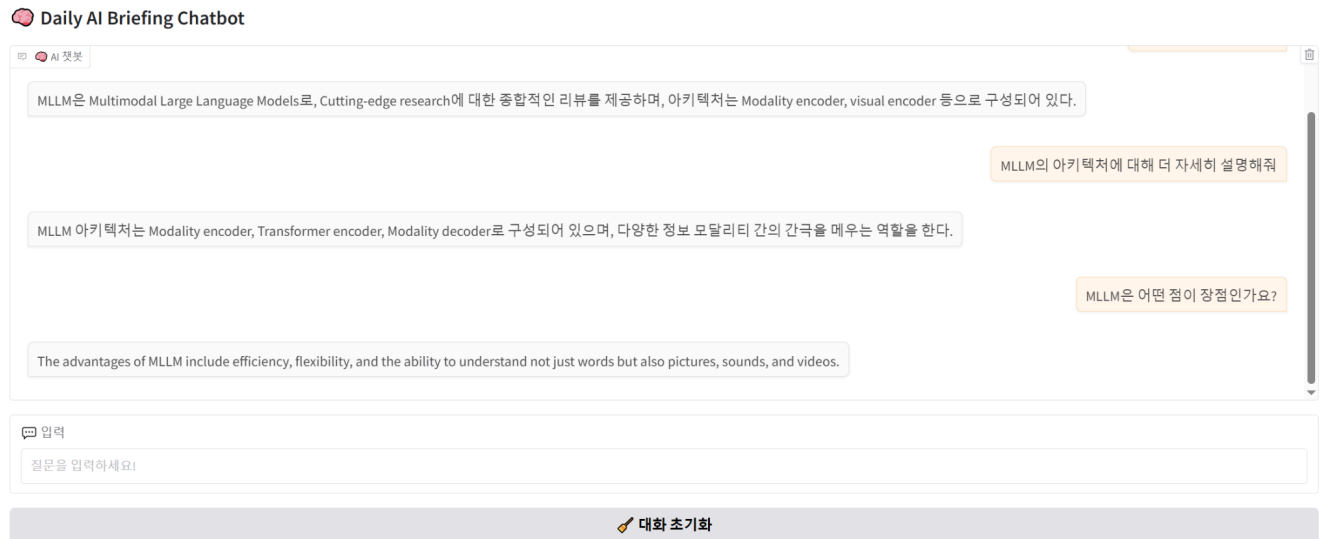
◦ IBM이 다양한 사용자 환경을 고려해 맞춤형 에이전트를 빠르게 제작할 수 있는 기술을 발표했습니다. 해당 기술은 기업 데이터를 활용해 클라우드 환경에서도 작동 가능하도록 하이브리드 솔루션을 공개하고 있습니다. [기사 바로가기](#)

4. ["AI 공격 막는 맞춤형 방패"...아카마이, 'AI 방화벽' 공개](#)

◦ 아카마이테크놀로지스가 'AI 방화벽'을 출시했으며 데이터 유출과 모델 조작을 방지하는 다층 보안 체계를 갖췄다고 밝혔습니다. [기사 바로가기](#)



## 8.2 Chat Bot



## 9. 결과

본 과제에서는 사용자가 자연어로 입력한 프롬프트로부터 핵심 정보를 자동으로 추출하고, 이를 기반으로 뉴스, 블로그, 논문 등 다양한 출처의 콘텐츠를 수집한 뒤, 전처리 및 요약 과정을 거쳐 벡터화하여 **RAG(Retrieval-Augmented Generation)** 구조에 적용하였습니다. 이를 통해 사용자는 키워드 조합, 수동 검색, 정리 등의 반복적인 작업 없이도 원하는 주제에 대한 핵심 정보를 빠르게 파악할 수 있게 되었으며, 기존 대비 정보 탐색에 소요되는 시간과 작업 부담이 현저히 줄어들어 전반적인 리서치 업무의 효율성이 크게 개선되었습니다. 특히 프롬프트 한 줄로 문서 유형, 기간, 주제를 통합적으로 탐색할 수 있어 콘텐츠 기획, 시장 분석, 기술 조사 등 다양한 업무 현장에 즉시 적용 가능한 실용성을 입증하였습니다.

향후에는 뉴스와 논문 외에도 다양한 형식의 문서(보고서, 공공 데이터, **SNS** 등)를 수집 대상으로 확대하고, 시스템 전반의 처리 속도 및 응답 일관성 최적화를 추진할 계획입니다. 또한, 수집된 정보들을 통합 분석하여 하나의 인사이트 문서로 생성해주는 기능과, 지정된 시간에 지정된 키워드에 대한 요약 리포트를 자동으로 메일로 발송하는 기능 등, 사용자 맞춤형 리서치 자동화 기능을 고도화할 예정입니다.

## 10. 감사의 글

본 과제의 수행에 있어 귀중한 지식과 통찰을 아낌없이 전해주신 강사님께 깊은 감사를 드립니다. 아울러 본 연구의 실질적 기반이 된 인턴십 기회를 제공해주시고 다양한 경험의 장을 마련해주신 **KT** 관계자 여러분께도 진심으로 감사의 뜻을 전합니다. 이 모든 지원과 격려가 본 과제의 완성과 연구적 성과에 큰 힘이 되었음을 밝힙니다.

## 11. 레퍼런스

본 과제 할 때 다음 도구 및 자료들을 참고하였다.

1. OpenAI. (2024). *OpenAI API documentation*. Retrieved from <https://platform.openai.com/docs>  
→ GPT-4 기반 프롬프트 처리, 요약, 기능 호출을 위한 언어 모델 사용 참조
2. Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., ... & Rush, A. M. (2020). *Transformers: State-of-the-art natural language processing*. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations* (pp. 38–45). Association for Computational Linguistics.  
<https://huggingface.co>  
→ 요약 모델(BART, Pegasus 등) 구현 시 참조한 HuggingFace Transformers 라이브러리
3. arXiv.org. (n.d.). *arXiv API documentation*. Retrieved from <https://arxiv.org/help/api>  
→ 주제 기반 최신 논문 수집 기능 구현에 활용
4. News API. (n.d.). *News API Documentation*. Retrieved from <https://newsapi.org/docs>  
→ 최신 뉴스 콘텐츠 수집 및 키워드 기반 필터링 구현에 사용
5. Lin, C. Y. (2004). *ROUGE: A Package for Automatic Evaluation of Summaries*. In *Proceedings of the Workshop on Text Summarization Branches Out* (pp. 74–81). Barcelona, Spain: Association for Computational Linguistics.  
→ 요약 평가 지표 ROUGE-1, ROUGE-2, ROUGE-L 기준 정의