

2024 뉴스빅데이터 해커톤 수상작 모음집



2024

뉴스빅데이터

해커톤 수상작 모음집

2024 뉴스빅데이터 해커톤 수상작 모음집

2024 뉴스 빅데이터 해커톤





한국언론진흥재단은 뉴스빅데이터를 활용한 새로운 사업 모델 및 공공서비스 발굴을 위해 ‘2024 뉴스빅데이터 해커톤’을 개최했습니다.

10월 중순부터 11월 말까지 진행된 이번 대회에는 129개 팀이 참가해 뉴스빅데이터에 대한 높은 관심을 보여주셨습니다.

이 책은 수상작을 모은 것으로, 한국언론진흥재단 홈페이지(www.kpf.or.kr)와 빅카인즈(www.bigkinds.or.kr)에서 무료로 보실 수 있습니다.

2024 뉴스빅데이터 해커톤에 참가해주신 모든 분들께 진심으로 감사드립니다.



2024

뉴스빅데이터

해커톤 수상작 모음집



· 목 차



Chapter 1

2024 뉴스빅데이터 해커톤 소개	9
--------------------	---

Chapter 2

2024 수상작 소개	13
-------------	----

● 질문하지 않는 사회, 뉴스를 통한 ‘좋은 질문’ 생성 AI 서비스	16
● MZ 세대를 위한 자동화 뉴스 웹툰 서비스	22
● 크로뉴스	27
● 시니어 대상 의료정보 제공 음성 AI 챗봇: 내이리(NAERI)	32
● 사기컷: 전자금융사기 예방 뉴스 큐레이터	37
● 인구 문제에 대한 통찰력 있는 시각: POPEYE	42
● 빅카인즈 AI Echo	46
● 저스트리슨(Just Listen)	51
● 뉴스 키워드 랭킹 시스템	56

Chapter 1

2024 뉴스빅데이터 해커톤 소개

01.

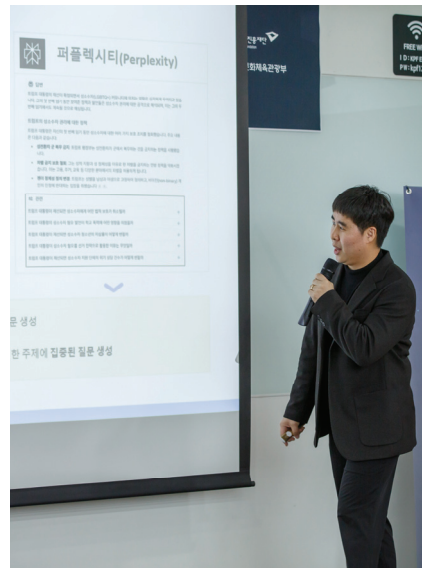
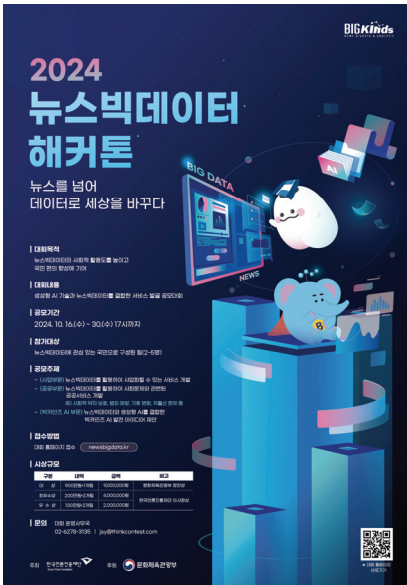
대회 목적

- 뉴스빅데이터의 사회적 활용도를 높이고 국민 편익 향상에 기여함

02.

대회 주제

- 생성형 AI 기술과 뉴스빅데이터를 결합한 서비스 개발
 - 뉴스빅데이터를 활용하여 사업화할 수 있는 서비스
 - 뉴스빅데이터를 활용하여 사회문제를 예방·해결할 수 있는 공공서비스
 - 뉴스빅데이터와 생성형 AI를 결합한 빅카인즈시 발전 아이디어 제안



03.

활용데이터

- 빅카인즈(www.bigkinds.or.kr)에서 제공하는 모든 데이터
- 공공데이터 포털 등에서 공개된 오픈 데이터
- 기타 활용하는 데 저작권 등의 문제가 없는 데이터

04.

빅카인즈

www.bigkinds.or.kr

- 종합일간지, 경제지, 지역일간지, 방송사 등을 포함한 1억여 건의 뉴스 DB에 빅데이터 분석 기술을 접목해 만든 뉴스 분석 서비스
- 관계도 분석, 키워드 트렌드, 연관어 분석 등 뉴스빅데이터를 활용한 다양한 시각화 분석 제공



Chapter 2

2024 수상작 소개

대 상

질문하지 않는 사회, 뉴스를 통한
‘좋은 질문’ 생성 AI 서비스

비글 (Beagle)

대상

질문하지 않는 사회, 뉴스를 통한 ‘좋은 질문’ 생성 AI 서비스

뉴스 데이터를 활용해 질문을 생성하고(Doc2Que) 관련 배경 정보를
제공(Que2Doc)함으로써 사용자의 질문력과 사고력을 증진시키는 서비스

팀명

비글(Beagle)

구성원

오수열(Team OCIAL)

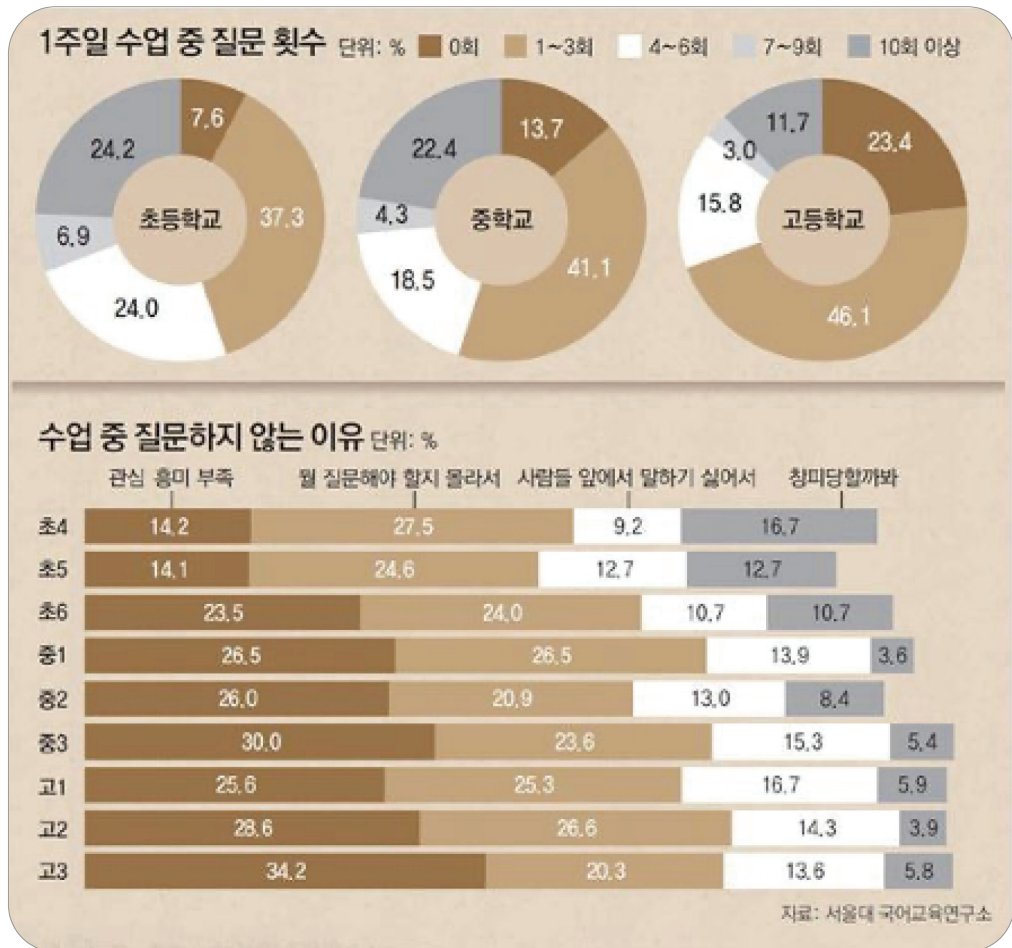
김순복(Team OCIAL)

천진원(Team OCIAL)

공연정(Team OCIAL)

정수석(Team OCIAL)



아이디어
발상 동기

서울대 국어교육연구소에 따르면 초등학교 4학년 학생 중 5.6%가 수업에서 ‘질문을 전혀 하지 않는다’고 응답했지만 고등학교 3학년에서는 이 비율이 약 5배 증가한 27.9%에 달했다. 이러한 현상은 단순히 학생 개인의 문제가 아니라, 흥미 부족, 창의성과 비판적 사고 결여, 그리고 학습 과정에서의 의사소통 부족으로 이어지는 사회적 문제이다.

또 다른 문제점으로는 ‘문해력 저하’를 들 수 있다. 이 문제는 아이들뿐만 아니라 성인들 까지도 포함하는 사회적 문제이다. “읽어도 읽지 못하는 아이들”과 “읽어도 이해하지 못하는 어른들”이라는 표현이 이를 잘 보여준다. 단순히 글을 읽는 것을 넘어, 내용을 제대로 이해하고 이를 사고로 연결하는 능력이 부족하다는 의미다. 문해력 저하의 핵심 원인인 어휘력 부족은 정보의 이해와 학습 능력을 제한한다.

비글(Beagle)은 이러한 문제를 해결하기 위해 ‘좋은 질문’의 가치를 중심으로 서비스를 설계했다. 좋은 질문은 학생들에게 관심과 흥미를 유발하고 ‘읽을 거리’를 제공함으로써 독서량과 어휘력 향상을 돕는다. 어휘력 향상은 문해력을 높여주고 더 좋은 질문을 만들 수 있는 선순환 구조가 형성된다. 이러한 과정을 거쳐 학생은 단순히 정보를 소비하는데 그치지 않고, 스스로 사고하고 배경지식을 확장하며 학습의 깊이를 더할 수 있다.

아이디어 내용



● '좋은 질문'이란?

- 지문 독해를 통해 도출되며 사고를 확장할 수 있도록 흥미를 자극하는 질문

● 핵심기능

- [수준별 지문 생성] 비글 서비스는 뉴스 원문을 기반으로 사용자 수준에 맞는 맞춤형 지문을 생성한다. 원문 기사를 사용자 수준별로 재구성하여 뉴스 접근을 원활하게 돕는다. 이를 통해 사용자 관심을 유도하고 독해와 이해력을 높이는 데 기여한다.
- [지문 독해를 통한 질문 생성] Doc2Que 모델을 활용하여 뉴스 지문에서 좋은 질문을 자동으로 생성한다. 이 질문들은 흥미와 사고를 자극하며, 배경지식을 확장할 수 있는 창의적이고 비판적인 사고를 이끌어낸다. 단순한 정답을 요구하는 질문이 아닌, 사고를 확장시킬 수 있는 '좋은 질문'을 제공하는 것이 핵심이다.
- [질문 기반 배경지식 확장] Que2Doc 모델은 질문과 관련된 추가 배경 정보를 제공하여 사용자가 더 깊이 있는 학습을 할 수 있도록 돕는다. 이를 통해 사용자는 기존 지문과 관련된 클러스터 문서들을 연결함으로써 학습의 폭과 깊이를 동시에 확장한다.
- [선순환 학습 구조 구축] '좋은 질문' → 관심과 흥미 → 다독 → 어휘력 증진 → 문해력 향상 → 다시 '좋은 질문'으로 이어지는 선순환 구조를 통해 학습자가 지속적으로 성장할 수 있도록 설계한다.

● 타겟이용자

- 공부는 잘하고 싶지만 '무엇을 질문해야 할지 몰라서' 질문하지 않는 학생
- 학생들이 수업 내용을 제대로 이해했는지 파악하기 어려운 교사
- 뉴스 지문 독해와 관련된 '좋은 질문'과 배경지식을 확장할 수 있는 관련 뉴스(교리문서)를 얻고자 하는 어른

서비스 구현



데이터 활용

- 빅카인즈(BigKinds)에서 제공하는 다양한 뉴스 데이터와 뉴스 클러스터를 활용한다. 뉴스 클러스터에서 대표 뉴스를 선정하고 수준별 지문 생성을 위해 국립국어원의 언어 및 문법 교육자료를 활용하였다. 전문 지식이 포함된 데이터에서 '좋은 질문'을 생성하기 위해 국내 논문 QA 데이터와 과학기술 문헌 데이터를 활용하였다. 이를 통해 복잡한 문제에 대한 심도 있는 질문을 생성한다.

LLM 기반 질문 생성

- 생성형 AI(LLM, Large Language Model)를 활용하여 뉴스 클러스터에서 '좋은 질문'을 생성한다. 대표 뉴스 지문에서 질문을 도출하며, 질문은 해당 지문이 포함된 뉴스 클러스터 뿐만 아니라 연결된 다른 클러스터에서 답을 찾을 수 있도록 설계된다. 질문 생성시 사용자의 수준을 함께 고려할 수 있도록 프롬프트를 구성하였다.

VectorDB를 활용한 관련 문서 추천

- 뉴스 클러스터링 데이터를 기반으로 VectorDB(벡터 저장소)를 구축하여 질문과 관련된 문서를 탐색하고 추천한다. 뉴스 데이터와 클러스터 정보를 임베딩(embedding)하여 벡터 형식으로 저장하며, 질문과 문서 간의 연관성을 계산해 적합한 문서를 제공한다. 임계점(threshold)을 설정하여 질문에 대한 답변이 존재하지 않을 경우 응답을 생략함으로써 할루시네이션(hallucination)을 방지한다.

RAG 기술 활용

- LLM과 VectorDB를 결합하여 질문 생성과 문서 추천 과정을 통합적으로 구현한다. RAG(Retrieval-Augmented Generation) 기술은 LLM이 수준별 지문과 질문을 생성할 때, VectorDB에서 가져온 관련 문서를 함께 활용한다. 이 과정은 질문을 통해 문서를 찾는 Que2Doc 모델로 구현되며, 사용자가 사고를 확장하고 학습 경험을 심화할 수 있도록 돕는다.

사업 모델



B2C

- 개인 학습자를 위한 뉴스 독해 및 질문 생성 서비스 제공
- 맞춤형 지문 생성, 질문 학습 모듈을 통해 문해력과 사고력을 증진하는 데 초점
- 유료 구독 모델을 통해 서비스 이용 가능

B2B

- 학습 관리 시스템(LMS)과 통합 가능한 API 제공
- 교육 기관의 학습 프로그램 강화
- 맞춤형 데이터 및 서비스 제공 계약을 통한 수익 창출

B2G

- 공교육 지원을 위한 문해력 및 질문력 향상 플랫폼 제공
- 공공 데이터 활용 프로젝트와 연계하여 교육 격차 해소 및 교육 혁신 지원
- 정부 주도의 문해력 증진 캠페인, '질문하는 학교'와 협업

기대효과



빅카인즈 서비스 아이디어 제안

- 비글(Beagle) 서비스는 빅카인즈(BigKinds) 뉴스 클러스터에 '좋은 질문' 생성 기능을 추가함으로써 뉴스 활용의 새로운 가치를 창출한다. 사용자는 단순히 뉴스를 소비하는 데 그치지 않고, 생성된 질문을 통해 뉴스를 더욱 깊이 이해하고 사고력을 확장할 수 있다. 이로써 뉴스 플랫폼은 정보 제공 역할을 넘어 학습과 교육의 도구로 활용 가능해지며, 개인 사용자와 공공기관 모두에게 유용한 지식 탐구 환경을 제공할 수 있다.

질문하는 학교와의 협업

- 2024년부터 교육부에서 처음으로 운영하는 '질문하는 학교(120개교)'와의 협업을 통해 교육 혁신에 기여한다. AI 기반 질문 생성 기능을 통해 학생들은 스스로 질문을 만들어가는 능력을 기를 수 있으며, 이는 교실 내 자율적 학습 분위기를 조성하는 데 큰 역할을 한다. 이 서비스는 교사들에게도 질문 중심의 연수 콘텐츠를 제공하여 수업 질을 높이고, 질문 수업 모델 발굴을 통해 학생 참여형 수업을 확대할 수 있도록 돕는다. 이를 통해 학생들은 창의적이고 비판적인 사고력을 배양하며, 교사는 학생들의 학습 성취도를 효율적으로 관리할 수 있다. 비글 서비스는 이러한 교육적 성과를 바탕으로 학교와 지역사회, 나아가 공교육 전반의 혁신적인 변화를 유도할 것이다.



MZ 세대를 위한 자동화 뉴스 웹툰 서비스

뉴코믹스 (NEWcomicS)

크로뉴스

저너러티브 (Jounarative)

최우수상

MZ 세대를 위한 자동화 뉴스 웹툰 서비스

생성형 AI 기술을 이용하여 MZ 세대에게 쉽고 재미있는 뉴스 기반 웹툰 제공

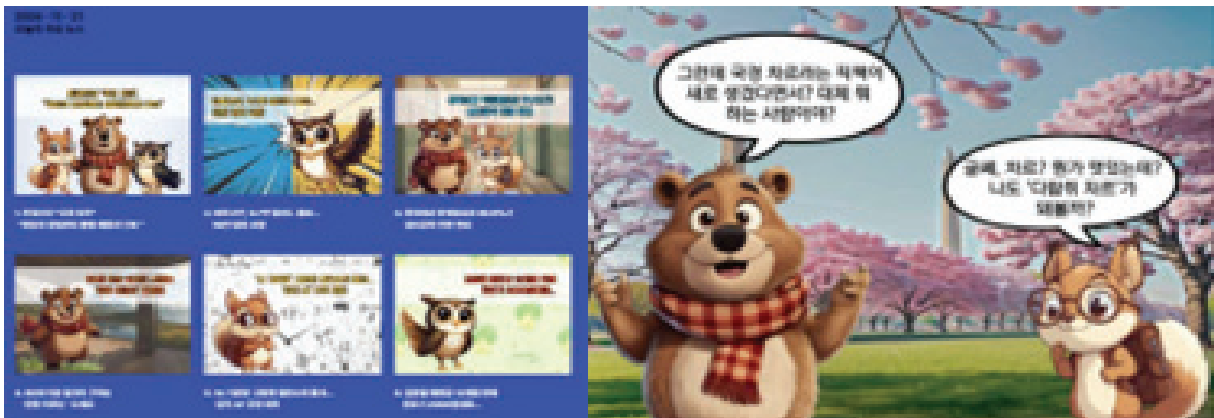
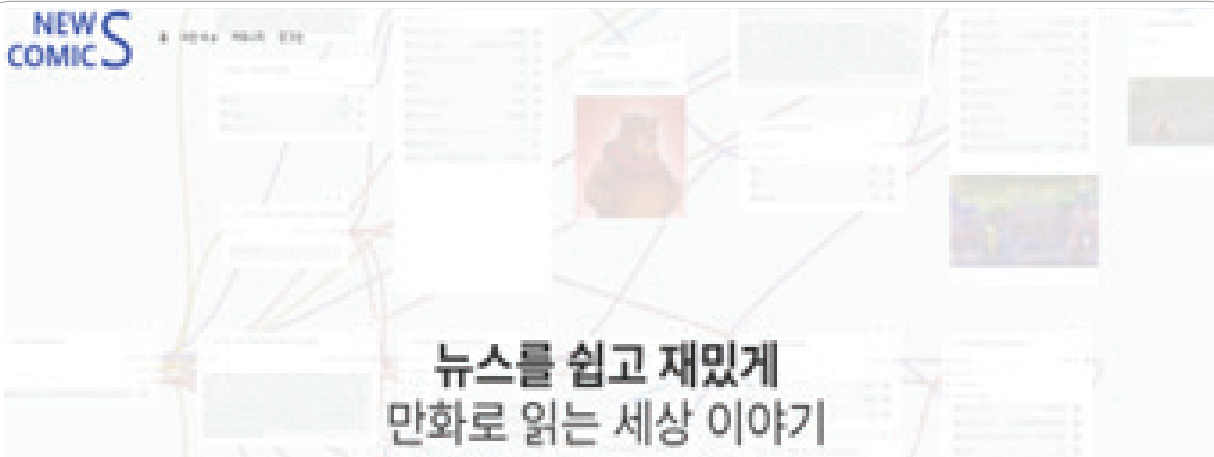
팀명

뉴코믹스
(NEWcomicS)

구성원

이인성(성균관대학교 의학과 5학년)
김동혁(경희대학교 스포츠의학과 4학년)
장승권(경희대학교 생체의공학과 4학년)

고은영(프리랜서)
김현서(한양대학교 자원환경공학과 3학년)



아이디어 발상 동기



최근 국민의 약 70%가 뉴스를 회피하고 있다. 특히, MZ 세대는 기존의 줄글 형식 뉴스에 대해 내용을 따라가거나 이해하기 어렵고, 쏟아지는 많은 기사 양에 지치며, 기분에 부정적인 영향을 미쳐 소비를 회피하게 된다고 이야기하고 있다.

그렇다면 MZ세대는 더 이상 뉴스를 소비하지 않는 것일까? 아니다. 이들이 뉴스를 소비하는 경로가 바뀌었을 뿐이다. MZ세대를 대상으로 한 설문조사 결과, 이들은 주로 SNS와 유튜브를 통해 뉴스를 접하고 있었으며, 카드뉴스, 동영상, 스토리 중심의 형식을 선호한다고 답했다. 스토리텔링과 시각적인 콘텐츠가 기존의 단순한 기사보다 훨씬 큰 호응을 얻고 있는 것이다.

이러한 흐름은 MZ세대의 관심을 끌 수 있는 새로운 형태의 뉴스 콘텐츠 개발이 필요하다는 점을 보여준다. 스토리텔링 요소를 활용해 뉴스 전달 방식을 개선하고, 시각적으로 매력적이면서도 쉽게 이해할 수 있는 콘텐츠를 제공한다면 젊은 세대의 뉴스 회피 문제를 해결할 가능성이 높다. 이에 따라 자동화된 뉴스 웹툰 서비스 ‘뉴코믹스’ 개발을 기획하게 되었다.

아이디어 내용



● 핵심기능

- 1) 오늘의 주요 이슈로 간편하게 즐기는 뉴스: 주요 이슈 및 긍정적 뉴스를 선별 제공하여 정보 과잉, 자극적 기사들로 인한 피로 경감. 여러 개의 기사를 바탕으로 한편의 웹툰을 제작하기에 중복되는 뉴스의 소비가 감소하고 다양한 관점을 녹여 뉴스의 맥락을 풍부하게 제공
- 2) 웹툰 형식으로 재밌게 즐기는 뉴스: 간결한 대사와 생동감 있는 이미지로 효과적인 정보 전달, 귀여운 캐릭터를 통해 부정적 느낌 경감, 어려운 용어를 쉽게 풀어 설명함으로써 뉴스에 대한 이해도 증가
- 3) 영상 웹툰 및 다국어 기능으로 다양한 독자층을 고려한 뉴스: 청각적 요소가 함께 있는 영상으로 사회적 약자도 소비 가능, 다국어 기능으로 외국인들도 소비 가능
- 4) 웹툰 자동화 기술로 실시간 정보를 제공하는 뉴스: 데이터 수집, 서비스 개발 및 배포 자동화를 통한 인력, 시간, 비용 절감

● 타겟이용자

새로운 뉴스 소비 형태를 추구하는 MZ 세대(10~30대)

서비스 구현



1. 데이터 수집

- 빅카인즈 API를 이용하여 오늘의 이슈 내 기사 제목 크롤링
- 네이버 API를 이용하여 기사 본문 전체를 크롤링

2. 데이터 분류/분석

- 코사인 유사도 분석을 통해 유사도 높은 기사 제거
- 웹툰 한 편 제작에 필요한 최종 기사 5개 선정

3. 시나리오 생성

- OpenAI API를 이용하여 선정된 기사 바탕 웹툰 시나리오 생성
- 프롬프트 엔지니어링 시 캐릭터 및 배경 설정, 편향 제거, 쉬운 어휘로의 전환 고려

4. 웹툰 생성

- Stable Diffusion 이용 시나리오로부터 말풍선, 캐릭터, 배경 등 요소별 이미지 생성
- ComfyUI를 통해 개별 이미지를 하나의 장면으로 합성하여 웹툰 완성

5. 웹툰 영상화

- 완성된 최종 웹툰과 Google Cloud의 음성 생성 AI를 융합하여 영상화된 웹툰 생성

사업 모델



1. 비즈니스 모델 캔버스(BMC)

1. 핵심 활동	2. 가치 제안	3. 고객		
· 뉴스 수집 및 토막 분류 모델링 (빅카인즈 API, 네이버 API) · 자동 시나리오 생성 (Open AI API) · 자동 이미지화 (Stable Diffusion API, Comfy UI) · 웹툰 영상화 -음성서비스(Google API) -다국어지원(DeepL API) -음성, 웹툰 이미지 합성(MOVIEPY)	· 뉴스 웹툰 형식으로 자동 변환 · 복잡하고 어려운 뉴스 재미있게 전달 · 24/7 지속적인 콘텐츠 자동 생산 및 업로드	(개인) · MZ세대(10대~30대) · 뉴스는 관심 있으나 긴 글을 깨리는 사람들 · 시각적 학습을 선호하는 사람 · 뉴스 소비의 제한이 있는 사회적 약자 · 한국에 거주하는 외국인	(기업) · 웹툰 플랫폼 업체 · 뉴스 제공업체 및 언론사 · 광고업체	(정부 및 공공기관) · 보건복지부 · 한국언론진흥재단 · 문화체육관광부
4. 고객관계	5. 유통 채널	6. 비용구조	7. 수익원	
(개인) 사용자 피드백 시스템 (기업 / 정부 / 공공기관) 기업제와 협업 및 상호 이익 창출	"뉴코믹스(NewcomicS)" 웹 사이트 SNS 플랫폼(인스타그램, 트위터)	· AI 개발 및 유지보수 비용 · 뉴스 빅데이터 유지보수 비용 · 클라우드 서버 운영 비용 · 플랫폼 개발 및 유지보수 비용 · 플랫폼 투자 유지비용 (인플루언서 마케팅, 캐릭터 굿즈 제작 비용)	(개인) · 월간 프리미엄 구독 서비스 · 캐릭터 굿즈 판매 (정부 및 공공기관) · 새로운 정책 홍보비 지원 (기업 및 광고업체) · 유료 콘텐츠 라이선스 및 API 판매 · 웹툰 / 동영상 내 광고비 · 사용자 기반 데이터 리포트 제공료	

2. 수익 모델(3P Model)

	세부 사항		연간 비용
AI 개발 및 유지보수	Open AI API 비용	· 평균 3,000토큰 사용 시 \$0.18	연간 \$ 1335.9
	Stable Diffusion API 비용	· 패널 당 \$0.002 x 18패널 = \$0.0 ~ \$0.036	
	Google Cloud TTS 비용	· 분당 \$0.01 ~ \$0.03 x 5분 = \$0.05 ~ \$0.15	
	Google Colab API 이용 비용	· colab pro + 월 \$ 44.99	연간 \$ 539.88
클라우드 서버 운영 및 뉴스 빅데이터 유지보수	뉴스 데이터 수집 비용	· 네이버 API 호출 횟수와 시간대에 따라 다름, \$0.05 ~ \$0.10	연간 \$ 730
	데이터 처리 및 저장 비용	· CSV 형식으로 저장 후 DB에 적재, \$0.02 ~ \$0.05	
	데이터베이스 관리비용	· DB 쿼리, 백업 및 유지 관리 비용, \$0.03 ~ \$0.05	
플랫폼 유지보수	프론트엔드 유지보수	· 브라우저 호환성 점검, \$0.01 ~ \$0.05	연간 \$ 620.5
	백엔드 유지보수	· DB 유지 및 최적화, \$0.02 ~ \$0.07	
	서버운영 비용 (1회성 작업 기준)	· 작업 시간 기반 서버 비용, \$0.05	
마케팅	연간 웹툰 홍보비 (중소 웹툰 홍보비 기준)	· 중소 웹툰 시장 연간 홍보비 기준 연간 \$ 180,000	연간 \$ 180,000
인건비	30명	· 중소기업 연간 (월급, 복리후생비) 기준 연간 \$ 1,002,506	연간 \$1,002,506

3. 사업 운영 예상 비용

payer	packaging	pricing	이용량 산정 방식
개인	월 구독	프리미엄 월 \$5	- MZ세대: 대한민국 인구의 약 34~36% (1,800만 명) - 콘텐츠 소비 성향 강한 50% = 900만 명 - 초기 시장 침투율 3% = 27만 명
	캐릭터굿즈	평균 단가 \$ 20	- 굿즈 구매 전환율: 웹툰 팬덤의 1% (18만 명) - 연평균 구매 빈도: 1.5회
기업	배너 광고	건당 \$1,200 (디지털 배너 광고 평균)	- 웹툰 플랫폼 월간 평균 트래픽: 500만 회 - 광고주 요청률: 0.1%
	동영상내 광고	건당 \$10,000 (몰입형 콘텐츠 광고 평균)	- 연간 동영상 콘텐츠 노출 수: 50만 회 - 1만 회당 1건 광고 계약 예상
	데이터 리포트	건당 \$750 (소셜 데이터 분석 리포트 평균)	- MZ세대 트렌드 데이터를 필요로 하는 기업: 약 3,000개 - 구매 전환율 1% 적용 → 월 30건 예상
정부	정책 홍보용 웹툰	프로젝트 당 \$1500 (디지털 캠페인 프로젝트 단가)	- 연간 정부 주요 캠페인 약 100건 - 이 중 10%가 웹툰 제작으로 의뢰될 것으로 예상

기대효과



(1) 경제적 효과

- 수익 확대: 자동화를 통한 시간 및 인력 최소화로 비용 효율성 및 생산성 증대
- 마케팅 기회 확대: 각종 브랜드와 캠페인 연계를 통한 광고 및 제휴 수익 확대
- 뉴스/콘텐츠 시장의 변화와 확장: 새로운 콘텐츠 소비 시장을 형성하고, 글로벌 확장을 통해 K-콘텐츠의 경쟁력 강화

(2) 사회적 효과

- AI 관련 산업 성장: 웹툰 자동화 시스템에 활용되는 기술 융합을 통한 AI 기술 산업 확장
- 교육 자료 활용: 학교 및 교육기관에서 사회적 이슈에 대한 관심과 이해 향상
- 정보 격차 완화: 시각적, 청각적 요소를 지원해 독해에 어려움을 겪는 사람들의 이해를 도움

(3) 타 데이터와 융합시 시너지 효과

- 공공데이터, 학술 및 연구 데이터와 융합 시 신뢰도와 객관성 향상 가능
- 소셜 미디어 데이터와 융합 시 반응 분석 후 피드백 통해 추후 개선 및 MZ 세대의 유행 요소를 활용하거나 관점을 다각화할 수 있음

최우수상

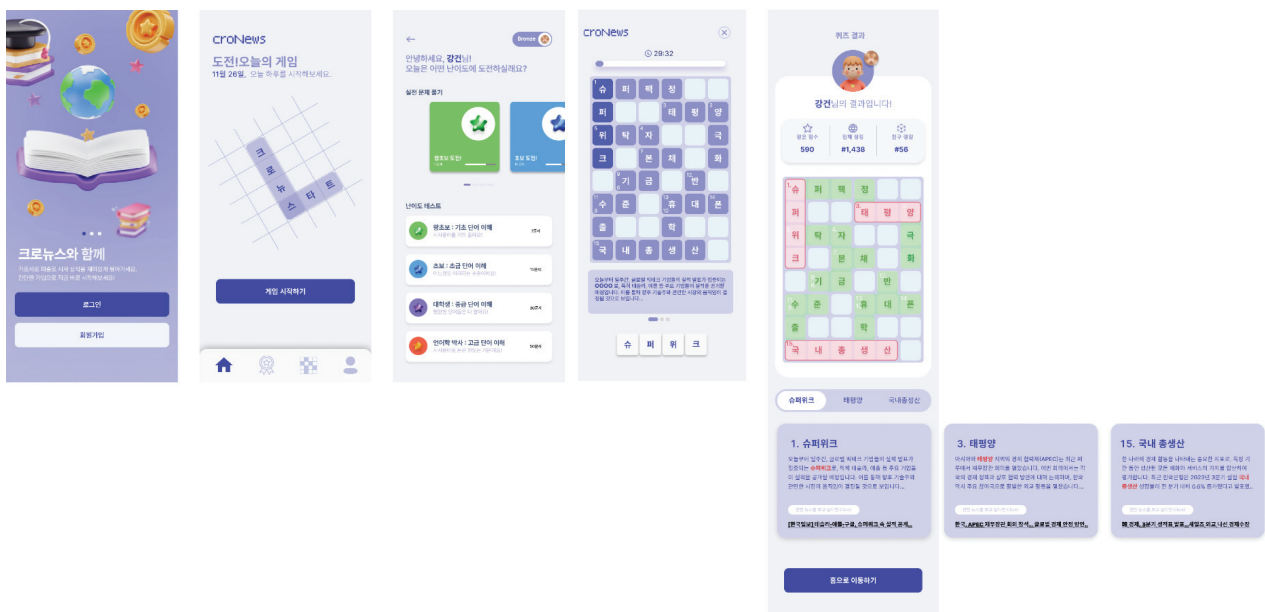
게이미피케이션 기반 뉴스 큐레이팅 서비스

팀명

저너러티브
(Jounarative)

구성원

강 건(조선대학교 컴퓨터공학과 학사 졸업) 권동현(건국대학교 의료공학과 학사 졸업)
박성일(한남대학교 빅데이터응용학과 학사 4학년) 이태화(한동대학교 콘텐츠융합디자인학부 제품디자인학과 학사 졸업생)
이혜람(한국외국어대학교 철학과 학사 졸업생)



아이디어 발상 동기



- 바쁜 현대인들은 뉴스를 접할 기회는 많지만, 깊이 있게 이해하거나 재미를 느끼기는 어려움
- 특히 2~30대는 기존의 레거시 미디어 대신 새로운 방식의 뉴스 소비를 선호함. 뉴스와 관련된 시사 상식을 쉽고 재미있게 습득할 수 있는 서비스의 필요성 증대
- 이러한 문제를 해결하기 위해 생성형 AI와 퀴즈 형식을 결합하여 뉴스 소비를 더 쉽고 즐겁게 만들기 위한 해결책을 제안

아이디어 내용



● 핵심기능

1. 실시간 뉴스 기반 퀴즈 생성
 - 생성형 AI가 최신 뉴스를 기반으로 중요한 키워드를 분석하고, 이를 바탕으로 가로세로 낱말 퀴즈를 자동 생성하여 매일 새롭게 제공
2. 난이도 맞춤형 설명 제공
 - 퀴즈에서 사용되는 키워드에 대해 사용자의 시사 상식 수준에 맞춰 생성형 AI가 난이도별 친절한 설명을 추가하여 학습 효과를 극대화
3. 뉴스 큐레이팅 기능
 - 퀴즈를 풀 후 관련 뉴스 기사를 직접 확인할 수 있는 링크 제공으로 심화 학습 지원
4. 게이미피케이션 요소
 - 점수 및 랭킹 시스템, 티어 제도를 통해 사용자 간 경쟁 유도

● 타겟이용자

- 뉴스와 시사 상식을 재미있게 학습하고자 하는 청소년부터 성인까지
- 세부 타겟 페르소나:
 - ▶ 교양을 쌓고자 하는 대학생
 - ▶ 시사 상식을 쉽게 접하고 싶은 취업준비생
 - ▶ 최신 이슈를 효율적으로 파악하고 싶은 직장인

서비스 구현



1. 백엔드 시스템
 - Spring Boot, Fast API 기반으로 빅카인즈 API를 연동해 뉴스 데이터 저장 및 퀴즈 생성 기능 제공
2. 프론트엔드 시스템
 - React Native로 모바일 앱 UI/UX 설계 및 구현
 - 스플래시 화면, 로그인/로그아웃, 퀴즈 풀기, 뉴스 큐레이팅 기능 등을 포함
3. 데이터베이스
 - MySQL을 사용하여 뉴스 원문, 키워드, 사용자 퀴즈 기록 및 계정 정보 관리
4. 프롬프트 엔지니어링
 - AI 기반으로 키워드 설명 및 맞춤형 퀴즈 생성 프롬프트 개발
5. 알고리즘
 - 같은 글자를 가진 키워드 쌍과, 아무 쌍을 가지지 않는 키워드를 만든 후, 무작위로 키워드 선택 후 DFS탐색을 활용해 크로스워드 생성

사업 모델



1. 기본 무료 서비스: 모든 사용자가 퀴즈 풀이와 뉴스 큐레이팅 기능을 무료로 이용 가능
2. 프리미엄 구독 서비스: 고급 퀴즈 카테고리, 추가 보상 포인트 및 유료 콘텐츠 제공
3. 광고 및 제휴: 언론사 및 공공기관과 협력해 광고 수익 창출 및 공익 콘텐츠 제작 지원

기대효과



1. 뉴스 소비 패턴 변화
 - 게이미피케이션 요소를 활용하여 뉴스를 재미있게 학습하며 소비를 촉진
2. 문해력 및 디지털 리터러시 향상
 - 퀴즈 풀이를 통해 복잡한 시사 정보를 쉽게 이해하고 디지털 뉴스 활용 능력을 높임
3. 젊은 세대와 레거시 미디어 연결
 - 새로운 형식의 뉴스 제공으로 전통 언론사와 젊은 세대 간의 간극을 좁힘
4. 사회적 가치 창출
 - 공공기관 협력을 통해 공익 콘텐츠 확산 및 뉴스 소비의 사회적 중요성을 강화



시니어 대상 의료정보 제공 음성 AI 챗봇 :
내이리 (NAERI)

오북이

사기컷 : 전자금융사기 예방 뉴스 큐레이터

사기컷

우수상

시니어 대상 의료정보 제공 음성 AI 챗봇: 내이리(NAERI)

시니어들의 건강한 내일을 함께하는 의료정보 제공 음성 AI 챗봇

팀명

오복이

구성원

송호진, 이소정, 강현지, 조순일



아이디어 발상 동기



- 시니어층은 디지털 리터러시와 헬스 리터러시 부족으로 인해 신뢰할 수 있는 의료 정보를 얻는 데 어려움을 겪고 있다. 실제로 시니어의 스마트폰 보유율은 높지만, 70대 이상의 디지털 정보화 수준은 51%에 불과하여 의료 정보를 탐색하고 이해하는 데 큰 제약이 따른다. 또한 건강 관련 정보의 탐색과 이해 능력도 낮아 시니어에게 신뢰성 있는 의료 정보를 쉽게 전달할 필요성이 대두되고 있다.
- 의료 정보를 획득하기 위한 시스템에도 문제점이 있다. 대학병원 입원 환자의 50%는 의료진과의 하루 상담 시간이 1분도 되지 않으며, 충분한 정보를 얻지 못하는 경우가 많다. 이러한 정보 부족의 상황에서 시니어는 접근성이 높은 인터넷 매체로 눈을 돌리지만, 인터넷에 흩어진 방대한 정보 속에서 필요한 정보를 찾아내기는 매우 어렵다. 또한 인터넷을 통한 의료 정보는 활용도가 높음에도 불구하고, 검증이 부족하기에 신뢰성 있는 정확한 정보를 찾는 것이 쉽지 않다.
- 이에 반해, 뉴스 의료 칼럼은 신뢰성, 최신성, 전문성을 갖추어 독자들에게 검증된 최신의 정보를 제공하며, 어려운 전문 용어를 대중이 이해하기 쉽게 풀이해 작성한 글로 현대 의료 정보의 필수 요구사항을 충족한다. 이러한 강점을 가진 뉴스 의료 칼럼을 기반으로 시니어에게 신뢰할 수 있는 의료 정보를 제공하는 서비스를 구상하게 되었다.

아이디어 내용



● 핵심기능

- ① **조작이 쉬운 UI/UX와 텍스트 크기 조절 기능:** 디지털 기기 활용 능력이 낮은 시니어를 고려하여 조작이 간편한 UI/UX와 가독성 높은 텍스트, 텍스트 크기 조절 기능을 지원한다.
- ② **음성 입출력 서비스:** 스마트폰의 키보드 사용이 어려운 시니어를 위해 질문과 답변을 음성으로 입출력 가능하도록 하여 사용성을 강화한다.
- ③ **드롭다운 메뉴:** 질문을 떠올리기 어려운 사용자를 고려해 증상, 진단 방법 등 질환에 대해 자주 묻는 질문 위주로 드롭다운 메뉴를 제공한다.
- ④ **LLM 기반 실시간 응답 생성:** 사용자가 입력한 질문을 분석하고, 입력 데이터와 검색 결과의 유사도 임계치(Threshold)에 따라 실시간으로 답변을 제공한다.
- ⑤ **개인 맞춤형 건강 정보 제공:** 사용자가 입력한 성별, 나이 등의 개인 정보를 바탕으로 유병률 높은 질환에 대한 맞춤형 정보를 제공한다.
- ⑥ **대화 내용 요약 및 진료과 안내:** 이전의 대화 내용을 기억하기 어려운 시니어를 고려하여 전체 대화 내용 중 핵심을 요약해 제공한다. 대화 내용을 토대로 사용자가 진료 가능한 과를 안내하며 서비스를 종료한다.

● 타겟이용자

① B2C

- 스마트폰 기기 사용에 어려움을 겪지만 건강 정보에 대한 니즈가 있는 시니어층
- 신뢰할 수 있는 건강 정보를 필요로 하는 국민

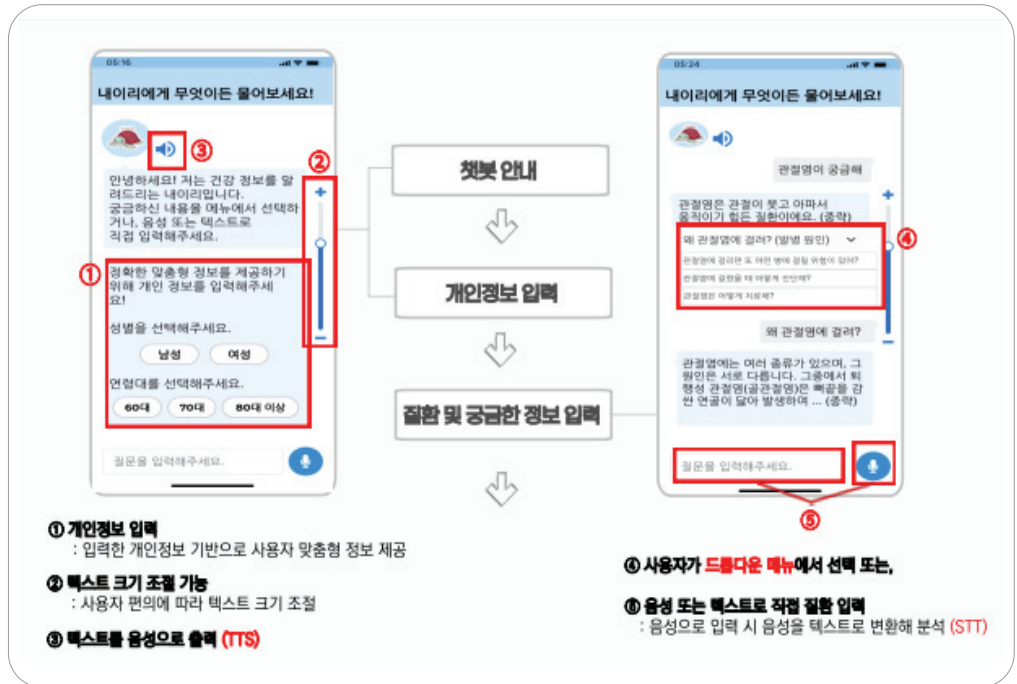
② B2G

- 전문성 있는 의료정보 전달 필요성을 가진 공공기관(ex. 질병관리청, 건강보험공단, 건강보험심사평가원)
- 통합된 의료 정보 제공 체계가 필요한 공공기관(ex. 의과학지식센터, 한국지능정보사회진흥원)

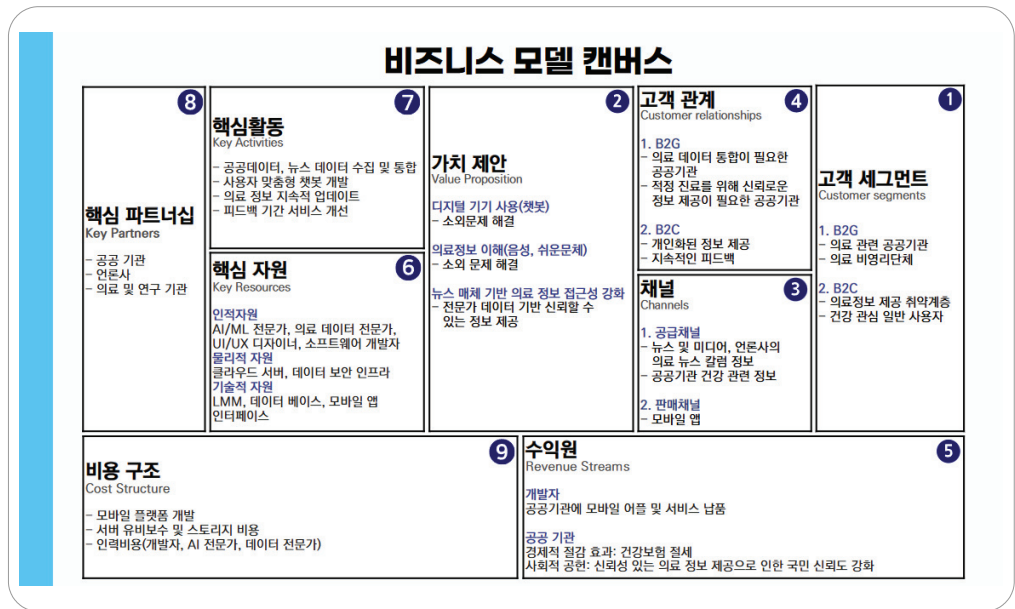
서비스 구현



- **질병 관련 뉴스 크롤링 및 DB 구축:** NAVER 뉴스를 통해 고혈압, 관절염 등 질병 관련 전문 의학 칼럼 기사를 수집하고 크롬DB(ChromaDB)를 활용하여 데이터베이스를 구축한다. 크롤링은 AsyncWebCrawler를 사용하며, URL에서 기사를 가져와 OpenAI의 text-embedding-3-small를 활용해 각 기사에 대한 임베딩을 생성한다. 이를 통해 기사의 질환 주제 및 주요 정보를 태그(Annotation)하고 정리한다.
- **문서 기반 질병 정보 검색 및 추천 시스템:** 크롤링된 데이터는 벡터 스토어(ChromaDB)에 저장되어 사용자 질문에 대해 관련 뉴스를 검색 및 추천한다. 사용자의 입력 질문과 관련된 뉴스를 코사인 유사도로 매칭하여 최적의 기사와 정보를 제공한다.
- **요약 및 정보 제공:** OpenAI GPT-4-mini 및 Retrieval-Augmented Generation (RAG)을 기반으로 수집된 기사를 3줄 이내로 요약한다. 해당 요약은 질환의 주요 원인, 증상, 진단, 치료, 예방 정보를 포함하며, 전문을 읽고 싶을 경우 URL도 함께 제공한다.
- **사용자 맞춤형 질병 정보 챗봇:** Streamlit 기반의 사용자 인터페이스를 통해 사용자는 질환 관련 질문을 입력하고 맞춤형 답변을 받을 수 있다. 입력된 질문은 음성 (STT) 및 텍스트(TTS) 입력을 모두 지원하며, 데이터베이스 검색 결과와 모델 기반 요약을 조합하여 사용자 맞춤형 정보를 제공한다.
- **AI 챗봇 프로토타입 설명:** 챗봇은 사용자의 입력 데이터를 분석하여 관련된 질환 정보를 제공한다. 예를 들어 관절염의 원인을 묻는 질문에 대해 크롬DB 내에서 관련 데이터를 검색하고, GPT-4-mini 기반으로 응답을 생성한다. 또한, 사용자 연령과 성별 같은 입력 정보를 기반으로 맞춤형 예방 정보를 제공하며, UI는 텍스트 크기 조절 및 음성 출력(TTS)을 지원하여 다양한 사용자 경험을 고려하였다.



사업 모델



기대효과



- 정보 접근성 향상:** 고령층 친화 디지털 접근성 표준을 준수하여, 글씨 크기를 크게 하고 정보 구조를 간단하게 정리하며 용어를 쉽게 구성한다. 이를 통해 정보 접근성을 높이고, 디지털 정보 소외 계층의 불균형을 해소하며 의료 정보 이용성을 개선할 수 있다.
- 정부 신뢰도 향상:** AI 기술을 활용한 정보 접근성 강화 서비스는 공공기관의 역량을 높이고, 신뢰도를 30% 이상 향상시키는 데 기여할 수 있다. 디지털 공공서비스 도입 초기 사례에서도 공공기관 신뢰도가 20~30% 상승한 바 있어, 본 서비스에서도 유사한 긍정적 효과를 기대할 수 있다.
- 의료 비용 절감 및 사회경제적 파급효과:** 불필요한 병원 방문 감소, 의료 시스템 부담 완화, 중증 질환 치료 비용 절감 등을 통해 최대 2조 6,400억 원의 비용 절감 효과를 예상할 수 있다. 시간 및 접근성 향상을 통해 사용자의 탐색 비용을 줄이고 의료 효율성을 극대화할 수 있다.
- 데이터 융합 시 시너지 효과:** 공공 데이터(보건복지부, 질병관리청) 및 전문 자료(학회, 의학 전문 잡지)와 결합하여 신뢰성 높은 의료 정보를 제공할 수 있다. 또한 GIS 데이터와 융합하여 질병의 지역적 특성을 반영한 맞춤형 정보를 제공함으로써 지역 격차를 해소하고 의료 서비스의 품질을 높일 수 있다.

우수상

사기컷: 전자금융사기 예방 뉴스 큐레이터

융합 데이터 활용을 통한

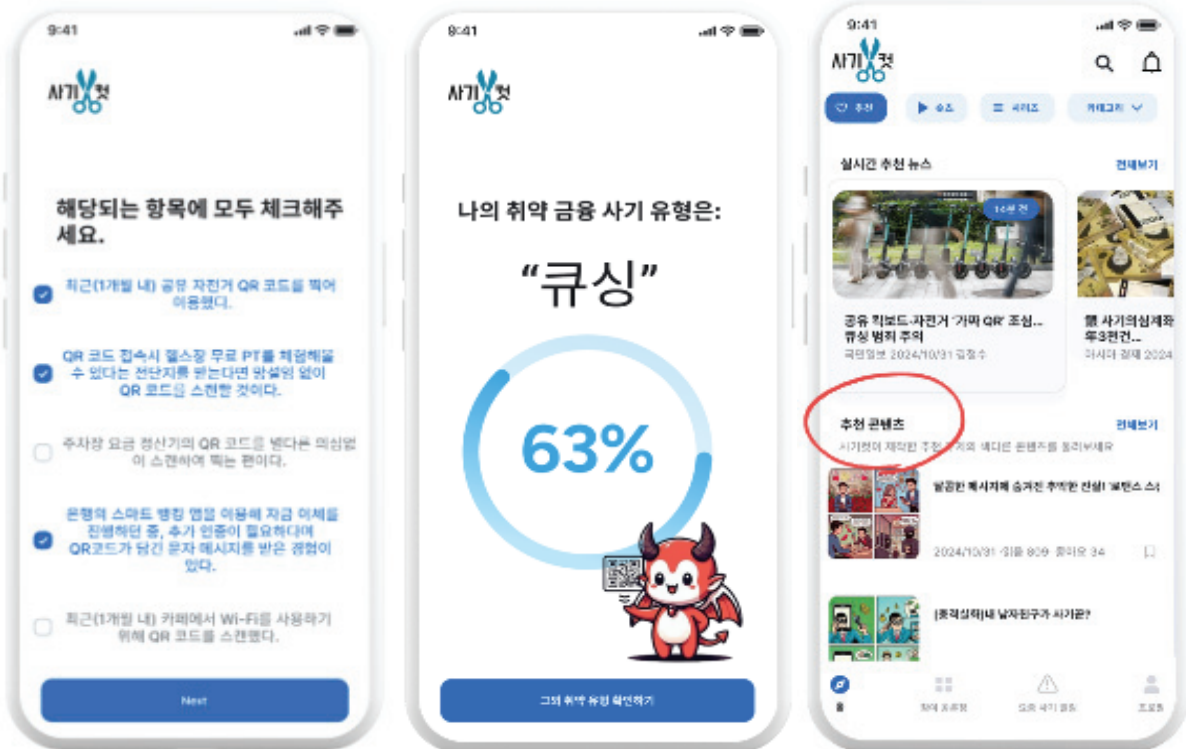
취약 사기 진단 기반 맞춤형 AI 가공 뉴스 콘텐츠 제공 서비스

팀명

사기컷

구성원

김예진(Nanjing University, International Economics and Trade)
남기웅(The Chinese University of Hong Kong, Biomedical Engineering)
박은수(한국방송통신대학교 컴퓨터과학과)
이혜정(University of Toronto, Human Biology)
지예림(경희대학교 글로벌커뮤니케이션학과, 소프트웨어융합학과)



아이디어 발상 동기



실제로 전자금융사기를 직접 경험하거나 당할 뻔했던 사기시저시저 팀원들의 문제의식에서 출발하였다. 금융감독원 자료에 따르면, 전자금융사기 피해가 점점 줄고 있는 다른 연령대와 달리 20대와 30대의 사기 피해율을 크게 증가하고 있는 추세였으며, 피해 사례 건수는 2020년 대비 2023년 약 67% 이상 상승하며 심각한 상황이다.

젊은 세대의 디지털 의존도는 점차 높아지고, 정교화되는 사기 수법에 더욱 쉽게 노출되고 있다. 그럼에도 불구하고 2030세대는 금융 정보를 주로 SNS나 지인을 통해 접하고, 이 과정에서 신뢰할 수 있는 정보의 부재로 인해 미디어 리터러시 능력이 낮아졌을 뿐만 아니라, 어렵고 무겁게만 느껴지는 뉴스를 보지 않아 사기 위험에 대한 경각심도 상대적으로 낮은 상황이다.

‘사기켓’은 사람들이 자신과 무관하다고 여기는 정보를 더욱 쉽게 놓친다는 점을 고려해, 뉴스 데이터를 활용해 개인의 취약한 전자금융사기 유형을 진단하고, 맞춤형 사기 뉴스를 제공하는 서비스를 제안한다. 이를 통해 사용자가 신뢰할 수 있는 정보를 기반으로 사기에 대한 경각심을 높이고, 전자금융사기 피해를 예방하며 미디어 리터러시를 강화하도록 유도할 수 있다.

아이디어 내용



● 핵심기능

- ① (사기 취약 유형 진단) 사용자가 입력한 개인정보와 금융 행동 패턴을 AI 모델로 분석하여 개인이 취약할 가능성이 높은 사기 유형을 진단한다. 이를 통해 사용자가 자신에게 가장 위험한 사기 유형을 사전에 인지하고, 예방에 집중할 수 있도록 돕는다.
- ② (AI 가공 전자금융사기 뉴스 콘텐츠 제공) 사용자의 특정 사기 유형에 관련된 맞춤형 뉴스와 예방 정보를 큐레이션한다. 사기 예방 뉴스는 숏폼 영상, 삽화, 논픽션 시리즈 등 다양한 콘텐츠 형식으로 가공하여 제공해 복잡하고 무거운 뉴스를 쉽고 흥미롭게 소비할 수 있도록 돕는다.

● 타겟이용자

- ① 금융 문해력이 낮고 금융사기에 대한 경각심이 부족한 2030세대
- ② 미디어 리터러시 능력이 약한 사용자
- ③ 효율적이고 흥미로운 정보를 선호하는 사용자

서비스 구현



- **활용 데이터:** 빅카인즈 뉴스 검색 API를 활용해서 뉴스 기사를 크롤링하여, 자체 DB(MongoDB)에 저장 및 관리
- **사용자 맞춤형 취약 사기 유형 매칭:** 뉴스와 논문 데이터를 기반으로 성별, 직업, 결혼 상태, 노후 계획 여부 등의 메타 항목으로 구성된 데이터셋을 JSON 파일로 저장한다. 사용자가 항목별로 답을 입력하면 해당 데이터를 기반으로 취약 사기 유형에 매핑된다. 일치하는 데이터가 없을 경우, 데이터를 one-hot encoding 후 해밍 거리를 계산해 가장 가까운 사기 유형에 매핑한다.
- **전자금융사기 유형 분류:** 크롤링한 기사를 KoBERT 모델로 분류한다. 이를 위해 5가지 사기 유형(보이스피싱, 스미싱, 파밍, 해킹, 메신저 피싱)별로 라벨링된 JSON 데이터셋을 제작해 모델을 학습시킨다. 각 사기 유형의 평균 벡터를 계산한 뒤, 크롤링한 기사의 벡터와 비교하여 코사인 유사도가 가장 높은 사기 유형으로 분류한다.
- **기사 기반 콘텐츠 제작:** OpenAI API, DALL-E3 API, Vadoo AI API를 활용해 기사 요약, 숏폼 영상, 삽화, 설화 형식의 Scam Story 콘텐츠를 제작한다. 이를 위해 하이퍼파라미터 튜닝과 프롬프트 엔지니어링을 적용하며, 기사 원문 URL도 함께 제공한다.

사업 모델



[B2G 모델]

사기컷은 금융감독원과 협력하여 API 활용 및 추가 기능 데이터셋 제공에 대한 긍정적인 수락을 받았다. 이 모델의 핵심은 사용자의 사기 피해 제보와 금융감독원의 정보를 데이터화하여 대시보드를 생성하고, 이를 언론사에 제공하는 것이다.

- **정보 공유 생태계:** 사용자가 사기 피해를 제보하면 금융감독원이 해당 데이터를 업로드하고, 사기컷은 이를 바탕으로 대시보드를 제작한다. 언론사는 이 대시보드를 통해 보도를 진행하며, 사용자는 이를 통해 유용한 정보를 얻는다.
- **신뢰 구축:** 이러한 쌍방향 정보 공유는 언론사가 독자와의 신뢰를 강화하고 독자의 목소리를 반영할 수 있도록 도와준다. 사기컷은 언론사와 협력하여 정보의 질을 높이고, 시민들에게 신뢰할 수 있는 정보를 제공함으로써 건강한 정보 생태계를 구축할 계획이다.

[B2B 모델]

사기컷은 금융 교육을 제공하는 보이스피싱 제로 금융 기관과 협력하여 고객 교육 콘텐츠를 제공하고, 이에 대한 수수료를 받는 B2B 모델을 구축할 예정이다. 이를 통해 금융 기관의 고객 교육을 지원하고, 사기 예방에 기여할 수 있다.

[마케팅 전략]

사기컷의 고객 유입 전략은 다음과 같다.

- **디지털 마케팅:** 금융사기 관련 유튜브 채널과의 홍보 콜라보레이션을 통해 2030세대를 타깃으로 맞춤형 취약 사기 유형 테스트를 배포하고, SNS 공유를 활성화할 계획이다.
- **고객 유치 프로그램:** 사기컷의 연속 출석 체크 및 게시판 제보 시 보상 포인트를 지급하여 사용자 참여를 유도한다. 또한, 지속적인 흥미 유발 콘텐츠를 업데이트하고, 게시판 기반의 정기적인 세미나를 개최하여 사용자와의 소통을 강화할 것이다.

이러한 전략을 통해 사기컷은 사용자 수를 점차 늘리고 안정적인 성장을 이루어내며, B2C 및 B2G 모델에서 B2B 모델로의 확장을 도모할 수 있다.

기대효과



① 사용자 맞춤형 취약 전자금융사기 진단을 통한 효율성 제고

사기컷은 사용자 개개인의 금융 이해도와 생활 패턴에 기반해 맞춤형 진단 결과와 그에 맞는 뉴스 콘텐츠를 제공한다. 이를 통해 사용자는 자신에게 맞는 정보를 선별하는 번거로움 없이 취약 금융 사기 정보를 효율적으로 습득할 수 있다.

② 전자금융사기 정보 접근성 향상 및 사회적 비용 절감

전자금융사기 뉴스를 숏폼 영상, 요약 삽화, Scam story로 등 재미있는 콘텐츠로 가공해 2030을 비롯해 뉴스 소비가 적은 이들에게도 전자금융사기 정보에 대한 접근성을 높인다. 이를 통해 전자금융사기 피해를 사전에 방지할 수 있다. 예를 들어, 대표적인 금융사기인 보이스피싱 피해를 연간 10% 예방한다고 가정할 경우, 최소 193억 원 이상의 피해액 감소 효과가 기대되며, 다른 금융 범죄까지 포함할 경우 절감 효과는 더욱 커질 것으로 예상된다.

* 2023년 보이스피싱 피해액 약 1,930억 원(경찰청)

③ 2030 세대의 뉴스 소비 패러다임 변화

사기컷은 2030세대가 전자금융사기 뉴스를 접하며 사기 예방의 실질적 효과를 경험할 기회를 제공한다. 이러한 경험을 통해 뉴스의 중요성과 가치를 실감하게 한다. 특히, 주로 SNS를 통해 정보를 소비하던 2030세대가 신뢰할 수 있는 뉴스 플랫폼으로 관심을 확대하게 함으로써 뉴스에 대한 접근성과 소비 습관을 개선하는 계기를 마련할 수 있다.



인구 문제에 대한 통찰력 있는 시각: POPEYE

POPBRAIN(팝브레인)

빅카인즈 AI Echo

BIG- 드림즈

저스트리슨 (Just Listen)

저스트리슨 (Just Listen)

뉴스 키워드 랭킹 시스템

스마트랭킹

본선
진출

인구 문제에 대한 통찰력 있는 시각 : POPEYE

인구(Population) + 시각(EYE)

인구 유관 빅데이터를 기반으로 대한민국 인구 문제에 대응하기 위한 최초의
생성형AI 프롬프트

팀명

POPBRAIN
(팝브레인)

구성원

김대환((사)한반도미래인구연구원 책임연구원) 유혜정((사)한반도미래인구연구원 선임연구위원)
임동근((사)한반도미래인구연구원 연구위원) 이주영((사)한반도미래인구연구원 연구원)
손승우(㈜한미글로벌 DT추진실 프로)



아이디어 발상 동기



- 대한민국은 급속한 저출산 및 고령화 문제에 직면하고 있으며, 이는 인구구조의 변화와 함께 경제, 사회, 문화 전반에 걸쳐 심각한 영향을 주고 있다. 이에 저출산·고령사회위원회 출범, 인구전담부서 설치 등 국가적 대책 또한 발빠르게 준비되고 있음
- 이러한 인구 현안 및 시급성과 달리 인구 관련 데이터를 관리하고 이를 분석·정리하여 손쉽게 이용할 수 있는 플랫폼은 부족한 실정이다. 통계청에서 제공하는 통계지리정보서비스(SGIS)가 존재하나, 해당 플랫폼은 인구 외에도 통계청에서 다루는 여러 부문의 데이터의 기본적인 사항만을 제공하고 있어 가독성 및 전문성의 한계가 존재
- 무엇보다 인구데이터에 대한 수요자의 접근성·활용능력·요구사항의 차이가 존재하며, 인구데이터 제공자의 제공방식과 수요자 여건에 있어 수급 비대칭이 큰 상황
- 따라서 인구 유관 데이터를 종합적이고 사실적으로 확인할 수 있는 개방형 플랫폼 개발은 인구 유관 여러 계층의 사용자(일반 시민, 공무원, 연구자 등)에게 수요가 있을 것으로 예상되며, 인구 문제 대응에 있어 데이터를 기반으로 과학적 의사결정이 가능할 것으로 기대

아이디어 내용

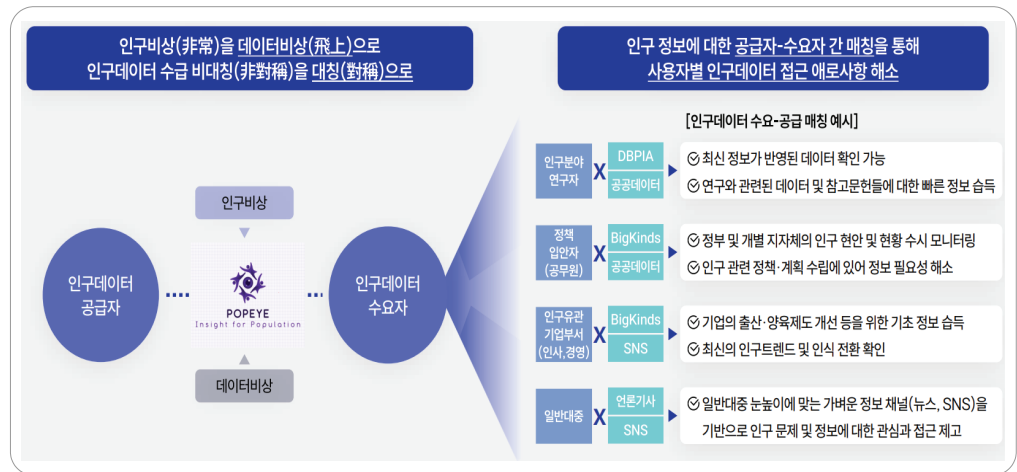


● 핵심기능

- API·벡터 DB 등을 통한 인구 데이터(언론기사*, 논문, 보고서, SNS 등) 융복합 및 도메인 특화 시스템 구축
 - * 본 해커톤의 빅인즈 API를 통해 인구·저출산·고령화 등 인구 유관 언론기사 빅데이터 DB화
- RAG - LLM 연계를 통해 각종 인구 데이터의 실시간 갱신으로 정보 최신화
RAG가 사전 학습된 LLM 모델에 추가로 검색 시스템을 결합하여 인구 데이터베이스를 실시간으로 검색하고 이를 기반으로 답변을 생성
- 채널 기반의 인구 정보 공급-수요자 간 매칭 및 데이터 레퍼런스 제공

● 타깃이용자

- 일반 시민: ChatGPT와 같은 개방형 플랫폼 형식으로 시민의 접근성을 높여 인구 현상에 대한 궁금증 해소
- 인구 유관 부서 담당 공무원: 정부 및 개별 지자체의 인구 현안 및 현황을 수시로 모니터링
- 인구 부문 연구자: 연구 수행을 위한 객관적 자료 확보와 인구와 관련된 최신 논의를 토대로 인사이트 도출



서비스 구현



① 인구 부문 데이터베이스 구축

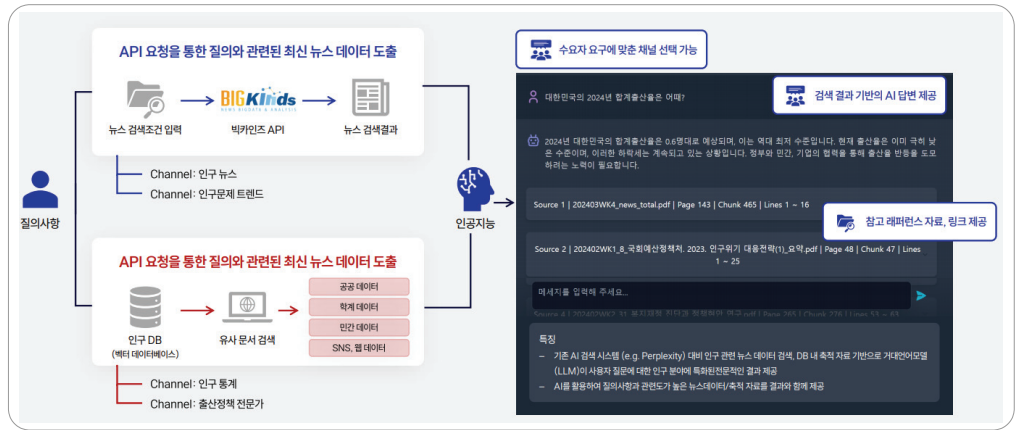
- 데이터 수집: 빅카인즈 API를 통해 인구, 저출산, 고령화와 관련된 언론 기사를 수집하고, 웹 스크래핑이나 API를 이용해 SNS와 논문 데이터 확보하는 등 인구 유관 각종 데이터 수집
- DB 구축: MongoDB, Elasticsearch와 같은 데이터베이스 시스템을 사용하여 데이터를 관리하고, RAG 연동 시 데이터 접근 속도를 높일 수 있도록 검색 시스템 최적화

② RAG가 연동된 LLM 기반 생성형 AI 'PopEye' 구현

- LLM과 RAG 통합: 사전 학습된 LLM과 Langchain 프레임워크의 RAG를 활용하여 사전 구축된 인구 데이터베이스와 생성형 AI 시스템을 연동
- 검색 시스템 설계: LLM이 질문을 받으면 먼저 관련 정보를 데이터베이스에서 검색하여 추출하고, 해당 데이터를 기반으로 최종 답변 생성
- 플랫폼 설계: PopEye가 다양한 인구 질문에 답변할 수 있도록 플랫폼을 설계하고, 정책 제안부터 통계 요약까지 폭넓은 대응을 할 수 있게 LLM 조정

③ 답변에 대한 데이터 소스 제공

- 출처 표시 시스템 구축: 답변에 활용된 데이터 출처(언론 기사, 논문 등)를 답변과 함께 표시할 수 있도록, DB에서 검색한 각 데이터의 출처를 연결
- UI/UX 구현: 답변과 함께 연관된 데이터 소스 링크와 요약을 제공하는 사용자 인터페이스 설계. 출처 확인을 쉽게 하고, 사용자가 필요한 경우 출처를 세부적으로 열람할 수 있는 서비스 디자인 구현



사업 모델



* POPEYE는 인구문제에 관한 정보 제공과 공익 증진을 위한 공공서비스로, 서비스의 지속가능성 및 수익모델의 실현가능성을 중심으로 고려

	[1단계] Prototype (MVP)	[2단계] Beta Version	[3단계] Platform & AI-Service
공급자 채널	☑ 연구자료 플랫폼(DBPIA) ☑ 뉴스빅데이터 플랫폼(Bigkinds)	☑ (추가) 공공플랫폼 내 공공데이터 ☑ (추가) SNS 데이터	☑ (추가) 민간기업 자체 인적자원데이터 ☑ (추가) 기업평가데이터
공급자 채널	☑ 인구 분야 연구자	☑ (확장) 229개 정부·지자체 인구정책 담당자 ☑ (확장) 상장사 300개 기업 내 경영부서	☑ (확장) 일반 대중의 사용까지 유도
서비스 수익모델	-	☑ 인구정책 관련 계획·보고서 정보 제공 ☑ 기업경영전략·지속가능경영보고서 등에 있어 인구 정보 제공	☑ 각 채널의 다양한 데이터를 자동으로 시각화하는 AI-서비스 제공 ☑ 요구에 맞춘 인구 보고서 제공 ☑ 인구 정책·제도 컨설팅
예상 hurdle	☑ Bigkinds, DBPIA와 같은 특정 인구 빅데이터 보유 기관과 초기 협업 및 시너지 필수 (API 접근 이슈 등)	☑ 현재 접근 가능한 공공데이터 및 크롤링 가능한 SNS 데이터 채널의 향후 접근 불가능성	☑ 민간보유 데이터 채널 운영에 있어 정보 보안 및 데이터 연계 이슈

기대효과



- 대한민국 인구 문제에 대응하기 위한 최초의 생성형AI 서비스(플랫폼)
- 인구 데이터베이스 실시간 연동 및 인구 분야 특화 생성형AI를 통해 정책 입안자와 연구자에게 최신 인구 통계 기반의 통찰 제공
- 대중이 인구 정보에 쉽게 접근하고 이해할 수 있는 창구 역할을 하여 인구 문제에 대한 인식 제고
- 나아가 향후 인구 데이터베이스 및 생성형AI를 기반으로 다양한 인구지표에 대한 시각화, 지역별 인구 문제 진단, 인구 변화 예측 모델 등의 서비스가 제공될 수 있을 것으로 사료되며, 궁극적으로 서비스들이 응축된 인구데이터 플랫폼이 구현된다면 강력한 파급력이 있을 것으로 기대

본선
진출

빅카인즈 AI Echo

빅카인즈 AI Echo, 모두의 목소리를 듣고 모두의 뉴스가 된다. 기존 빅카인즈 AI에 음성 서비스와 맞춤형 기능을 더해, 디지털 소외 계층과 일반 사용자 모두가 쉽게 뉴스에 접근할 수 있는 포괄적 플랫폼입니다.

팀명

BIG-드림즈

구성원

윤예빈(삼육대학교 인공지능융합학부 4학년)

이유진(오픈버스, 기획팀)

이동현(나인와트, 연구개발)

(별첨_빅드림즈_수상집제출용.png)



아이디어 발상 동기



“말로 묻고, 귀로 듣는 뉴스! 빅카인즈 AI Echo”

한 할머니께서 뉴스를 보시다 “요즘 정보가 너무 빨리 바뀌어서 자식들한테 물어보기도 미안해”라며 한숨을 내쉬셨습니다. “다들 아는 애긴데 나만 이렇게 모르는 게 많나...”라는 말씀에는 정보에서 점점 멀어지고 소외되는 어르신들의 현실이 담겨 있었습니다. 그 모습을 보며, 누구나 쉽게 소식을 접하고 정보를 이해할 수 있는 서비스가 있다면 좋겠다는 생각이 들었습니다.

‘빅카인즈 AI Echo’는 정보 취약 계층을 위해 설계된 AI 음성 뉴스 서비스입니다. 사용자가 음성으로 쉽게 질문하고 정보를 찾을 수 있도록 하여 뉴스 접근성을 극대화합니다. 또한, 개인 맞춤형 뉴스 제공을 통해 사용자의 관심사에 맞는 정보를 전달하며, 흥미와 연관성을 높입니다. 이 서비스를 통해 정보 격차를 줄이고, 정보 접근의 장벽을 줄여 사회적 참여의 기회를 확대합니다.

빅카인즈 AI를 활용한 해당 서비스는 방대한 뉴스 데이터를 바탕으로 정보를 선별하고, 지역과 사용자 맞춤형 뉴스 큐레이션을 가능하게 합니다. 또한, 뉴스의 출처와 신뢰성을 확보하여 정보 취약 계층에게 적합한 정보를 전달합니다. 빅카인즈 생성형 AI와 음성 기반의 접근 방식은 사용자가 별도의 기기 조작 없이 간단히 질문만으로 필요한 정보를 얻을 수 있게 하여 정보 격차를 줄이고 디지털 소외를 해소하는 데 기여합니다. 우리는 단순히 뉴스를 전달하는 것 이상으로, 사용자의 궁금증을 해결하고 세상의 흐름과 함께 하는 기회를 제공하고자 합니다.

아이디어 내용



● 핵심기능

① 음성 기반 정보 제공

사용자가 음성으로 질문을 하면, 관련 뉴스와 정보를 실시간으로 음성으로 제공하여 정보 접근의 편리함을 높입니다. 음성 인터페이스를 통해 뉴스를 검색하고, 필요한 정보를 손쉽게 얻을 수 있습니다.

② 맞춤형 정보 제공

사용자 선호도와 검색 기록을 바탕으로 관심사에 맞는 맞춤형 뉴스 콘텐츠를 제공합니다. 음성을 통해 원하는 뉴스 주제를 질문하고, 관련 정보를 바로 받을 수 있습니다.

③ 관심 키워드 기사 푸시 알림

사용자가 설정한 관심 키워드와 연관된 최신 뉴스를 실시간으로 푸시 알림으로 제공합니다. 중요한 뉴스와 소식을 놓치지 않도록 알림 기능을 강화하여 즉각적으로 정보를 전달합니다.

④ 추천 기사 카드 기능

사용자 선호와 검색 기록을 바탕으로 추천 기사 카드를 생성합니다. 사용자는 한눈에 관련 뉴스를 확인하고 필요한 정보로 빠르게 접근할 수 있습니다.

⑤ 직관적인 UI

정보 소외계층을 고려한 간단하고 직관적인 사용자 인터페이스를 설계하여, 음성 명령을 통해 누구나 쉽게 사용할 수 있도록 화면과 기능을 최적화했습니다.

● 타겟이용자

빅카인즈 AI Echo의 주요 타겟은 뉴스와 정보를 효율적으로 소비하고자 하는 개인 및 조직입니다.

1. 디지털 기기에 익숙하지 않은 사용자: 음성 기반 인터페이스로 복잡한 UI 없이도 쉽게 정보를 탐색할 수 있어, 중장년층 및 IT 비숙련자에게 적합합니다.
2. 정보 접근성이 중요한 사용자: 멀티태스킹 중에도 음성으로 뉴스를 청취하고 정보를 얻을 수 있어, 시간 효율성을 중시하는 직장인, 학생 등이 주 대상입니다.

1) 주요 타겟층

- 디지털 소외 계층: 고령층, 저소득층, 장애인 등 정보 접근성이 낮아 뉴스 소비 및 정보 활용에서 어려움을 겪는 사용자
- 일반 사용자: 최신 정보를 효율적으로 소비하고자 하는 뉴스 애호가 및 데이터 기반 의사결정을 선호하는 개인 및 전문가

2) 타겟층의 주요 니즈

- 디지털 소외 계층:
 - 간편한 인터페이스를 통한 정보 접근
 - 맞춤형 콘텐츠 제공으로 뉴스 이해도 향상
 - 데이터 리터러시 및 디지털 기기 사용법 향상 지원
- 일반 사용자:
 - 빠르고 신뢰도 높은 개인화 뉴스 추천
 - 시각화된 데이터 분석을 통한 인사이트 제공
 - 다양한 소스에서 통합된 뉴스 피드 활용

서비스 구현



- 실시간 음성 인터페이스 모델:
 Websocket을 통해 클라이언트의 오디오 트래픽을 전달받아 프록시하는 미들웨어를 구성합니다. 미들웨어에서는 OpenAI의 Real-time API를 통해 GPT-4o-Realtime-Preview 모델을 호출하고, 사용자의 질문에 대한 답변을 생성하며, 실시간 Voice-to-Voice 구조의 인터페이스를 제공합니다.
- 지식 검색 증강 시스템:
 대용량 뉴스 기사 데이터를 벡터 형태로 저장하고, 검색 요청에 대한 응답을 생성합니다. Amazon Titan Embedding Model을 사용하여 긴 분량의 뉴스 기사 텍스트를 숫자 형태로 임베딩 하고, OpenSearch를 통해 사용자의 질문과 가장 유사도가 높은 기사를 검색합니다. 이때 뉴스 데이터가 저장된 데이터베이스 스키마에 대한 정보와 샘플 쿼리 데이터를 문서로 참조하여 Schema Linking 작업을 수행할 수 있습니다.

사업 모델



빅카인즈 AI Echo는 다양한 사용자층의 요구를 충족시키기 위해 단계적인 사업 로드맵을 바탕으로 서비스를 확장합니다.

- 기본 사업모델
 - 기업 솔루션 제공: 기업 및 공공기관에 맞춤형 빅카인즈 AI Echo 솔루션을 제공하여 데이터 분석, 음성 기반 리포트 등 비즈니스 도구로 활용할 수 있도록 지원합니다.
 - 파트너십 및 라이선스: 콘텐츠 제공 업체와의 파트너십을 통해 다양한 뉴스 소스를 확보하고, AI 음성 인터페이스 기술을 타 플랫폼에 라이선스 형태로 제공합니다.

1. 초기 단계:

디지털 기기에 익숙하지 않은 중장년층과 IT 비숙련자, 시각장애인 등 정보 접근성이 중요한 사용자에게 초점을 맞춥니다. 음성 기반 대화형 인터페이스를 통해 간편하고 직관적인 뉴스 소비 경험을 제공하며, 이들을 대상으로 기본 뉴스 제공 서비스를 무료로 운영합니다. 구독형 서비스와 광고 수익 모델을 병행하여 서비스 안정화를 도모합니다.

2. 확장 단계:

바쁜 현대인과 직장인, 학생 등 멀티태스킹이 필요한 사용자층으로 확장합니다. 뉴스 외에도 날씨, 교통, 지역 행사 정보 등 생활 밀착형 서비스를 추가해 플랫폼 활용도를 높입니다. 기업 및 공공기관을 대상으로 맞춤형 데이터 분석과 음성 리포트를 제공하는 솔루션을 도입하며, 프리미엄 구독 서비스를 활성화합니다.

3. 성장 단계:

글로벌 시장으로 진출해 다국어 음성 인식과 현지화된 콘텐츠를 제공하며 사용자층을 확대합니다. 뉴스 소비 데이터와 트렌드 분석을 활용한 데이터 판매 모델을 강화하고, 음성 기반 AI 플랫폼으로 발전시켜 다양한 산업 분야와 파트너십을 구축합니다. 지속적인 기술 고도화를 통해 정보 접근성과 사용자 경험을 더욱 향상시키고, 사회적 참여를 촉진하는 플랫폼으로 자리매김합니다.

기대효과



① 정보 접근성 향상

실시간 음성 뉴스 제공과 맞춤형 정보 제공을 통해, 디지털 소외 계층을 포함한 모든 사용자가 손쉽게 필요한 정보를 접근할 수 있습니다. 특히 음성 기반 서비스는 사용자가 뉴스 소비에 어려움을 겪지 않도록 직관적이고 편리한 방식으로 뉴스에 접근할 수 있게 합니다. 이는 특히 노인층과 같은 정보 취약 계층의 정보 격차 해소에 중요한 역할을 할 것입니다.

② 사용자 맞춤형 뉴스 제공

사용자 선호도와 검색 기록을 바탕으로 제공되는 추천 기사 카드는 개별 사용자가 원하는 뉴스만을 빠르게 제공하여 정보 탐색 시간을 절감합니다. 사용자는 자신의 관심사에 맞는 최신 뉴스를 신속하고 효율적으로 확인할 수 있으며, 이는 뉴스 소비의 질을 한층 높여줍니다.

③ 신뢰성 있는 뉴스 데이터 제공

빅카인즈의 검증된 뉴스 데이터를 기반으로 사용자에게 정확하고 신뢰할 수 있는 정보를 제공함으로써, 불확실한 소스에서의 정보 소모를 줄입니다. 사용자는 신뢰도가 높은 뉴스에 접근할 수 있어, 정보의 정확성과 신뢰성에 대한 걱정 없이 최신 소식을 손쉽게 받아들일 수 있습니다.

④ 사회적 소외감 감소

음성 기반 뉴스 서비스와 개인 맞춤형 기능을 결합하여 사용자는 단순히 뉴스를 소비하는 것을 넘어, 사회적 소통의 기회를 갖게 됩니다. 특히 음성 인터페이스를 통해 뉴스와 대화를 나누듯 소통할 수 있어, 사회적 연결감을 증진시키고 소외감을 감소시킵니다. 이는 노인층의 사회적 참여를 촉진하고, 그들이 더 나은 사회적 연결을 느끼도록 돕습니다.

본선
진출

저스트리스(Just Listen)

시각장애인을 위한 아이트래킹 패턴 기반 뉴스 요약 및 맞춤형 뉴스 추천 음성 서비스
 아이트래킹 기술을 활용해 시각 정보를 패턴화하여 뉴스 빅데이터를 요약하고, 추천
 알고리즘을 최적화하여 개인맞춤형 뉴스를 음성으로 제공함으로써 시각장애인의
 뉴스 접근성을 향상시키는 서비스

팀명

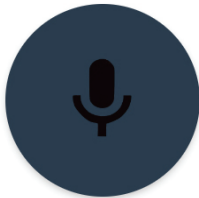
저스트리스(Just Listen)

구성원

김나혜(스마트인재개발원 수강생) 전정훈(스마트인재개발원 수강생)
 이종민(스마트인재개발원 수강생) 신충용(스마트인재개발원 수강생)

음성을 통한 맞춤형 뉴스 추천 및
 뉴스 요약 화면

저스트리스



2번. [동아일보] 시각·청각 동시에 손상된 장애인
 위한 전담기관 첫 개소

3번. [에이블뉴스] 시각장애인 디지털 신기술 활
 용교육에 힘쓰는 SK텔레콤

3번 뉴스 알려 줘

SK텔레콤 디지털접근성팀에서 시각장애인들의
 디지털 접근성을 지원하기 위해 교육을 실시. 점차
 단말기 '한소네' 사용법 등의 교육.



저스트리스



전체뉴스

시각장애인 전용 뉴스화면
 카테고리 별로 쉽게 접근 가능한 화면
 (모든 버튼은 음성으로 안내)

전체 뉴스

경제

정치

사회

연예

스포츠

실시간 화제



저스트리스



전체뉴스

아이디어 발상 동기



우리 사회는 지능정보사회로 진입하며 디지털 대전환이 가속화되고 있지만, 장애인은 여전히 디지털 정보화 활용 면에서 취약한 계층으로 분류되고 있습니다. 연구에 따르면, **장애인의 디지털 정보화 활용 수준이 낮아 삶의 만족도에 부정적인 영향을 미치며**, 사회적 지지가 높은 경우에도 정보를 스스로 활용하기보다 조력자에게 의존하는 경향이 강해, 디지털 정보화 활용의 수동성을 더욱 강화하는 것으로 나타납니다.

시각장애인의 스마트폰 활용 비율은 전체 장애인 평균보다 높고, 매일 사용하는 비율도 90% 이상에 달하지만, 텍스트 기반 뉴스와 최신 기술 활용에 어려움을 겪고 있습니다. 기존 뉴스 페이지는 시각적 흥미요소를 중점으로 제작되어 있기 때문에, 시각장애인이 뉴스에 효과적으로 접근하기 어렵습니다. 또한 기술의 발전으로 텍스트를 음성으로 들려주는 서비스가 존재함에도, 뉴스 본문 전체를 읽어주기 때문에 시각장애인들이 원하는 정보를 얻기까지 시간이 시각적으로 글을 읽는 사람보다 과도하게 소요됩니다.

이러한 문제는 시각장애인의 정보 접근성을 제한하고, **정보의 빈곤으로 이어져 삶의 질을 저하시키는** 주요 요인으로 작용합니다. 반면, 비장애인들은 시각정보를 활용하여 뉴스를 빠르게 이해하고 중요한 정보를 효율적으로 습득하고 있습니다.

이에 따라, 저희는 아이트래킹 기반으로 시각정보를 패턴화한 요약 기술과 빅카인즈 뉴스 데이터를 결합하여 뉴스 맞춤형 요약서비스를 제공하고 음성을 통해 원하는 뉴스를 쉽고 빠르게 찾을 수 있는 서비스인 '저스트리슨(Just Listen)'을 기획하였습니다. 이 서비스는 시각장애인들에게 뉴스 정보를 얻기 위해 키워드를 직접 눈으로 보고 선택하는 비장애인의 방식과 유사한 경험을 제공합니다.

'저스트리슨'은 시각장애인의 정보 빈곤 문제를 해결하고, **정보 접근성을 개선하여 디지털 정보화 활용 수준을 향상시키는 동시에**, 시각장애인의 능동적인 정보 활용과 사회적 연결, 참여를 지원하는데 기여할 것입니다.

아이디어 내용



● 핵심기능

1. 뉴스 요약 서비스

- 시각 정보 기반 요약: 아이트래킹 데이터를 활용하여 일반인의 시각적 패턴을 분석하고, 뉴스 본문의 주요 내용을 추출해 요약 및 음성으로 제공
- 탐색 비용 절감: 시각장애인이 뉴스 본문 전체를 들을 필요 없이, 필요한 핵심 정보에만 빠르게 접근 가능
- 왜곡 방지: 기존 빈도 수 등의 요약 모델에 시각데이터를 추가함으로써 요약 과정에서 기사의 의미를 더 정확히 전달

2. 생성형 AI를 통한 대화형 맞춤 뉴스 추천

- 개인화된 뉴스 큐레이션: 사용자가 음성으로 입력한 키워드나 관심 주제에 따라 관련 뉴스를 실시간으로 검색 및 추천
- 선호 언론사 및 주제 필터링: 사용자가 선호하는 언론사 또는 주제에 해당하는 기사만을 선택적으로 제공하여 개인 맞춤형 경험 제공

3. 음성 인터페이스 기반 접근성 강화

- 직관적인 음성 명령: 음성으로 키워드를 입력하고 원하는 주제를 선택할 수 있는 편리한 UI/UX 설계
- 빠른 응답과 피드백: 버튼 클릭 시 비프음을 통해 클릭 여부를 즉각적으로 알림.

4. BigKinds 데이터 활용

- 방대한 뉴스 데이터: 100개 이상의 언론사에서 제공되는 최신 데이터를 바탕으로, 신뢰성과 다양성을 겸비한 뉴스 콘텐츠 제공
- 실시간 업데이트: 데이터가 매일 업데이트되며, 최신 뉴스를 반영하여 사용자의 정보 요구 충족

5. 차별화된 기술과 효율성

- 기존 서비스 대비 시간 절감: 시각적 기준으로 만들어진 뉴스페이지의 버튼으로 음성을 듣는 서비스 대비 뉴스 접근 시간이 최대 30% 단축
- 정보 습득 속도 향상: 뉴스 본문 청취 시간 1분 30초에서 20초로 대폭 감소

 타겟 이용자

시각장애인 및 정보 접근성이 필요한 사용자: 뉴스와 같은 방대한 정보를 효율적으로 탐색하고자 하는 시각장애인 및 정보 취약 계층

서비스
구현



- ① **정확한 Embedding을 위한 자연어 처리:** 사용자의 음성을 Google web speech api를 이용하여 실시간으로 텍스트로 변환하고, 변환된 텍스트는 KoNLPy를 활용하여 불용어(조사, 어미)를 제거하고 핵심 키워드를 추출하여 노이즈를 최소화합니다. KoBERT 모델을 활용해 추출된 키워드의 문맥적 중요도를 평가하고, 사용자의 요구와 문맥이 맞는지 분석하여 강화된 프롬프팅을 통해 할루시네이션을 방지합니다.

- ② **LLM 모델을 통한 사용자 맞춤형 정보 제공**: 1차 Embedding된 키워드의 범주가 넓을 경우, 대화형 AI를 통해 사용자와의 대화로 세부 주제로 좁혀 나가며 맞춤형 뉴스를 추천합니다.
- ③ **아이트래킹 패턴 학습을 활용한 뉴스 특화 요약**: 아이트래커 기계를 통해 일반적으로 뉴스에서 독자들이 어떤 부분을 먼저 주목하는지를 분석한 아이트래킹 데이터를 활용하여 학습합니다. 이 패턴을 기반으로 뉴스 본문에서 중요도가 높은 위치의 단어에 가중치를 부여합니다. 가중치가 부여된 학습된 모델에서 가중치가 부여된 단어를 중심으로 LLM 모델을 활용하여 간결하고 핵심적인 요약본을 생성합니다.
- ④ **뉴스 데이터 큐레이션**: 빅카인즈에서 제공하는 104개 언론사의 방대한 뉴스 데이터를 활용합니다. 이 데이터는 카테고리별, 이슈별, 일자별로 라벨링되어 있어 개인화된 뉴스 추천에 적합합니다. 정형화된 데이터로 뉴스 기사를 분류하여 적합한 뉴스 정보를 추출하고, 이를 기반으로 맞춤형 뉴스 큐레이션 서비스를 제공합니다.
- ⑤ **요약된 뉴스 음성 출력**: 정리된 요약본을 음성화하여 사용자가 시각적 제약 없이 뉴스를 청취할 수 있도록 지원합니다.

사업 모델



항목	B2G	B2B	B2C
서비스 제공 대상	정부, 지방자치단체, 복지부, 교육부 등 공공기관	사회적 책임(CSR)을 강조하거나 시각장애인 지원 프로그램을 운영하려는 기업 (대기업, 미디어 회사 등)	- 시각장애인 개인 사용자 - 디지털 취약계층
서비스 제공	- 정부 및 공공기관에 맞춤형 뉴스 요약 및 추천 음성 서비스 제공 - 공공기관의 디지털 정보 접근성 개선 및 시각장애인 정보 제공	- 기업과 제휴하여 시각장애인 대상 맞춤형 뉴스 서비스 제공 - 기업의 사회적 책임을 다하는 서비스 제공	- 개인 소비자 대상으로 맞춤형 뉴스 서비스 제공 - 시각장애인에게 음성 뉴스 제공
수익 모델	① 정부 출연금: 정부와 협약을 통해 지원받음 ② 공공 사업 연계: 복지부, 교육부 등 공공기관의 정책적 지원을 위한 서비스 제공	① 기업 제휴: CSR(사회적 책임 활동) 일환으로 시각장애인에게 정보 제공 ② 광고 수익: 뉴스 기반 광고를 통한 수익 창출	① 구독 모델: 기본 기능 외 프리미엄 기능(고급 뉴스 필터링, 다양한 음성보이스 기능 제공 등) 제공
소비자 접점	- 공공기관 플랫폼 통합: 공공기관 시스템에 통합하여 시각장애인에게 맞춤형 음성 뉴스 제공 - 시청각 콘텐츠 배포: 공공 서비스로 활용되며, 시각장애인들이 접근할 수 있도록 음성 뉴스 제공	- 기업 전용 플랫폼: 기업의 웹사이트나 소셜 미디어에 통합하여 시각장애인 대상 서비스 제공	- 모바일 앱: 시각장애인 개인이 음성 뉴스 서비스를 통해 뉴스 정보를 실시간으로 청취

기대효과



- ① **(시각장애인의 정보 접근성 확대)** 본 서비스는 시각장애인이 단순히 음성을 통해 뉴스를 청취하는 것에서 나아가, 스스로 원하는 정보를 탐색하고 접근하는 독립성을 강화한다. 아이트래킹 기반의 뉴스 요약과 맞춤형 뉴스 추천 시스템을 통해 시각장애인은 정보에 더 쉽게 접근할 수 있게 되며, 이는 시각장애인의 사회적 연결과 정보화 사회 참여를 촉진하는 데 중요한 역할을 한다.
- ② **(정보의 맞춤화 및 개인화)** 대화형 AI와 맞춤형 뉴스 추천 시스템을 통해 사용자의 요구를 반영하여, 시각장애인들도 개인의 의견과 의도에 맞춘 맞춤형 뉴스 청취가 가능하다. 이를 통해 시각장애인들은 더 많은 선택권을 가지며, 자신에게 필요한 정보만을 효율적으로 얻을 수 있게 된다.
- ③ **(사회적 연결과 참여 증대)** 시각장애인들은 정보의 빈곤으로 인해 사회적 연결과 참여에 어려움을 겪는다. 본 서비스는 시각장애인들이 최신 뉴스와 정보를 쉽게 얻을 수 있도록 하여, 그들이 사회적 이슈와 연결되고 참여할 수 있는 기회를 확대한다. 이를 통해 시각장애인들의 삶의 질을 높이고, 사회적 참여를 증진시킬 수 있다.
- ③ **(사회적 관심 증대)** 시각장애인들의 디지털 정보화 접근성을 높이는 본 서비스는, 장애인 정보 접근성 문제에 대한 사회적 관심과 인식을 증진시킬 수 있다. 이러한 서비스의 발전은 디지털 사회에서 시각장애인의 참여 기회를 확대하고, 더 나아가 전반적인 사회적 이해도를 높이는 데 기여할 것이다.
- ④ **(활용성 향상)** 시각장애인들은 디지털 정보화 활용에 있어 많은 제약을 겪고 있다. '저스트리슨' 서비스는 시각장애인들이 디지털 기술을 능동적으로 활용할 수 있도록 돕고, 정보 접근을 통해 디지털 정보화 수준을 향상시킨다. 이로써 시각장애인의 정보 활용 능력이 강화되고, 사회적 참여가 더욱 활성화될 것이다.

본선
진출

뉴스 키워드 랭킹 시스템

생성형 AI와 다기준 방법론을 활용한 새로운 뉴스 키워드 랭킹 시스템으로
기존 빅카인즈의 '오늘의 키워드' 서비스 보완

팀명

스마트랭킹

구성원

강민석(경희대학교 빅데이터응용학과 석사과정 4기)

임해빈(경희대학교 빅데이터응용학과 석사과정 4기)

뉴스 키워드 랭킹 작성자: [강민석](#)

스마트 키워드 랭킹

스마트 키워드 랭킹은 문맥이 고려
된 단어 조합으로 구성된 키워드에 대
한 다기준 의사결정 랭킹 시스템입니
다.

Category
☐ (전체)
☐ 문화일반
☒ IT_과학일반

IT_과학일반
1,037개

WordCloud

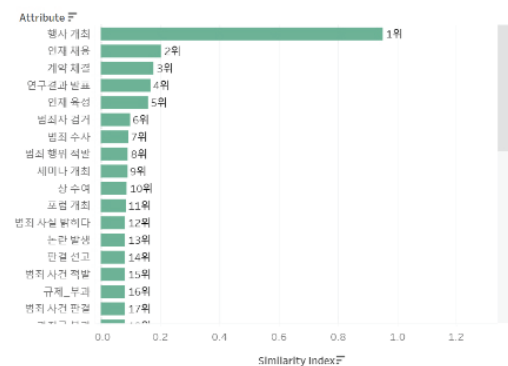
특강 개최사업 운영 이벤트 개최
컨퍼런스 개최 상 수여소송 제기 간담회 개최 프로그램 운영
워크숍 개최범죄 사건 적발연구 결과 발표 인재 채용판결 선고
범죄 사실 밝히다세미나 개최 강연 개최 논란 발생
학생 시험 치르다
인재 육성 행사 개최 연구결과 발표
범죄자 검거 계약 체결
사업 추진 범죄 행위 적발문제점 제기 범죄 사건 발생
회의 개최 범죄 사건 판결 범죄 수사 규제_부과 계획 수립 대회 개최
현상 분석 포럼 개최시스템 시범 운영

Tableau Public에서 보기

Criteria Weights

중 뉴스 개수	평균 검색 점수	평균 경과시간	평균 글자 수
0.4400	0.4113	0.0899	0.0588

Keyword Ranking



Attribute	중 뉴스 개수	평균 검색 점수	평균 경과시간	평균 글자 수
행사 개최	420.0	1.2	768.1	728.7

아이디어 발상 동기



현대 사회에서는 하루 평균 6만여 개의 뉴스가 생산되며, 정보의 홍수 속에서 중요한 정보를 빠르게 선별하는 일이 점점 어려워지고 있다. 이를 완화하기 위해, 뉴스 빅데이터 플랫폼인 ‘빅카인즈’는 수많은 뉴스를 분석해 사용자들에게 다양한 서비스(‘오늘의 키워드’, ‘오늘의 이슈’ 등)를 제공하고 있다.

본 연구자들 역시 석사 과정에서 자연어처리와 관련된 연구를 수행하며 빅카인즈 플랫폼을 활용해왔다. 사용 도중 빅카인즈 내 서비스에서 아쉬운 점을 인지하고 보다 나은 솔루션을 제시하고자 본 해커톤에 참여하게 되었다.

현재 ‘오늘의 키워드’ 서비스는 자주 언급된 인물, 장소, 기관을 워드클라우드 형태로 제공하며, 빅카인즈 사용자 중 플랫폼에 머무는 시간이 짧은 사용자들이 다수 사용하고 있다. 하지만, 해당 서비스는 키워드의 다양성 및 맥락 부족과 평가 방식의 단순성으로 인해 약 50%의 사용자 만족도에 그치고 있다. 이에 따라 본 아이디어는 맥락이 포함된 단어 조합의 키워드와 다기준 평가 시스템을 도입한 새로운 뉴스 키워드 랭킹 시스템을 제안하여 사용자 만족도를 높이하고자 한다.

아이디어 내용



- ① (키워드 형태 개선) 기존 ‘오늘의 키워드’는 키워드가 인물, 장소, 기관에 국한되어 현 이슈를 명확하게 파악하기 어려웠다. 본 서비스는 생성형 AI를 활용하여 문맥이 포함된 단어 조합 키워드를 추출함으로써 사용자가 구체적인 맥락을 쉽게 이해할 수 있도록 한다.
- ② (평가 기준 다양화) 기존 키워드는 단순 빈도 기반으로 평가된 반면, 본 서비스는 뉴스가치 판단기준을 토대로 경과 시간, 글자 수, 감성 점수, 댓글 및 좋아요 수 등 5가지의 기준을 추가 반영하여 키워드 중요도를 평가한다. 이를 통해 보다 체계적이고 신뢰도 높은 키워드 랭킹 시스템을 제공한다.
- ③ (효율적 구현 및 배포) 생성형 AI 모델인 Google Gemini를 활용해 키워드를 도출하며, 다기준 의사결정(MCDM) 방법론인 CRITIC-TOPSIS를 적용해 키워드의 중요도를 평가한다. 결과는 대시보드 형태로 제공되어 사용자들이 직관적으로 활용할 수 있다.

● 타겟이용자

- ① 빅카인즈를 짧은 시간 이용하는 일반 사용자
- ② 최신 이슈를 신속히 파악하고자 하는 바쁜 현대인
- ③ 보도 우선순위를 결정해야 하는 미디어 종사자

서비스 구현



- **Step 1.** 생성형 AI(Gemini)를 통해 뉴스 내용을 대표하는 몇 가지의 속성을 추출한 후, 유사한 내용의 뉴스를 통합하기 위해 추출한 속성을 정제된 키워드로 분류한다.
- **Step 2.** 뉴스가치 판단기준(시의성, 부정성, 영향성, 흥미성, 정보량)을 토대로 키워드의 중요도를 평가할 수 있는 6가지의 기준(평균 경과시간, 평균 감성 점수, 총 댓글 수, 총 좋아요 수, 총 뉴스 개수, 평균 글자수)을 정의하며, 이를 토대로 키워드에 대한 기준값을 계산하여 의사결정 행렬을 구성한다.
- **Step 3.** 다기준 의사결정(MCDM) 방법론 중 상관성, 변동성을 고려하는 객관적 가중치 부여 방법인 CRITIC을 통해 각 기준의 중요도를 평가하며, 이상 해와 비이상 해와의 거리를 고려하는 TOPSIS를 통해 키워드를 정의한 기준을 토대로 중요도 랭킹을 도출한다.
- **활용 데이터:** 빅카인즈 Open API - 뉴스 검색 API(기사 정보)
크롤링 - 네이버 뉴스(기사 댓글 수, 좋아요 수)

기대효과



1. (현재 이슈 파악 용이성) 본 서비스는 키워드가 문맥이 고려된 단어 조합으로 구성되며, 이를 워드클라우드 및 막대 그래프 차트 등이 포함된 대시보드로 구성해 제공되어 사용자가 현재 이슈를 더욱 쉽고 빠르게 파악할 수 있으며, 각 의사결정자들의 효율적인 의사결정 지원에 기여한다.
2. (다양한 기준을 통한 평가) 본 서비스는 빈도 외의 다양한 기준을 키워드 평가에 반영하여 랭킹을 도출하며, 사용자가 대시보드에서 특정 키워드에 대한 통합적인 랭킹과 함께 각 기준값들을 확인할 수 있다. 이를 통해 보다 체계적인 키워드 평가와 함께, 사용자는 키워드에 대한 평가 기준과 랭킹 결과에 대해 쉽게 해석할 수 있다.
3. (효율적이며 즉각적인 구현 및 배포 가능) 본 서비스의 구현 단계에서 생성형 AI는 한글 이해 성능이 뛰어나고, 시간 및 비용 측면에서 효율적이며, JSON 형식 출력 기능이 포함된 Gemini 1.5 Flash, MCDM은 간단 명료한 프로세스를 갖춘 CRITIC-TOPSIS, 배포 단계에서는 웹에 간편한 배포가 가능한 Tableau를 활용함으로써 효율성을 극대화하였으며 즉각적인 활용이 가능하다.

2024 뉴스빅데이터 해커톤 수상작 모음집

발행인	김효재
편집인	남정호
기획	한국언론진흥재단 뉴스빅데이터팀
발행일	2024. 12. 18
발행처	한국언론진흥재단 04520 서울특별시 중구 세종대로 124 프레스센터 12층 T. 02-2001-7577 F. 02-2001-7560 www.kpf.or.kr
디자인·인쇄	(사)아름다운사람들 T. 02-6948-9650

2024
뉴스빅데이터
해커톤 수상작 모음집

