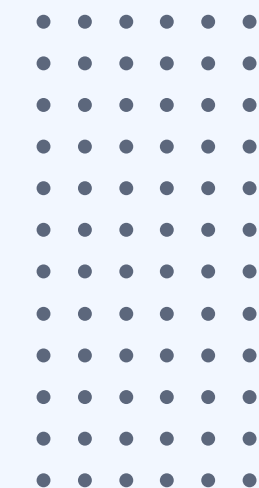




SKN06-FINAL-2TEAM

LLM을 활용한 웹툰,웹소설 추천 시스템

Pixary 중간발표



이세화,안형진,정유진,전하연,박미현



SKN06-FINAL-2TEAM

목차

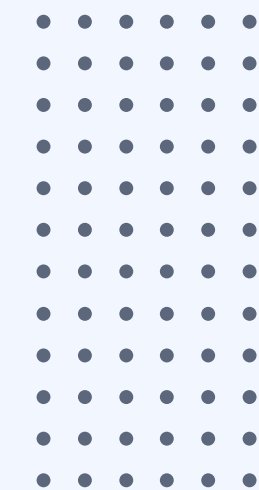
01. 프로젝트 개요

02. 진행 현황

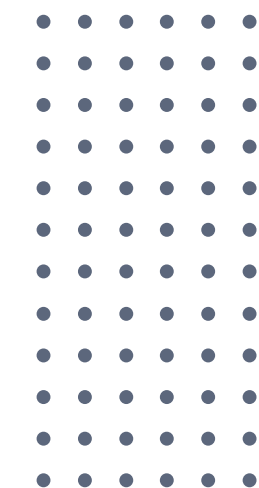
03. 예상 결과 목표

04. 이슈사항

05. Q&A



01. 프로젝트 개요





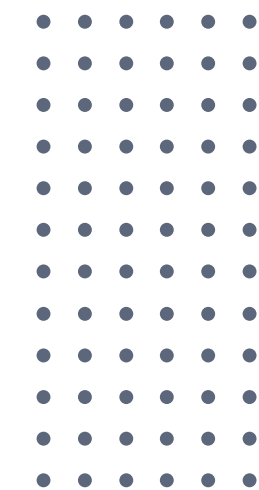
SKN06-FINAL-2TEAM

Team 소개

팀원: 이세화, 안형진, 정유진, 전하연, 박미현

강사님: 김성환 강사님

멘토님: 이동준 멘토님





1. 프로젝트 개요 - 주제선정 이유

1. 국내·외 시장 규모와 전망

- 국내 '23년 매출 규모: 약 2조 1,890억 원 (전년 대비 19.7% 성장)
- 국내 2019~2023년 연평균 성장률: 37% 로 지속 확대 중
- 국외 연평균 36.8% 성장하며, 미국 등에서도 공식 산업으로 인정받는 추세

2. 불친절 추천 시스템

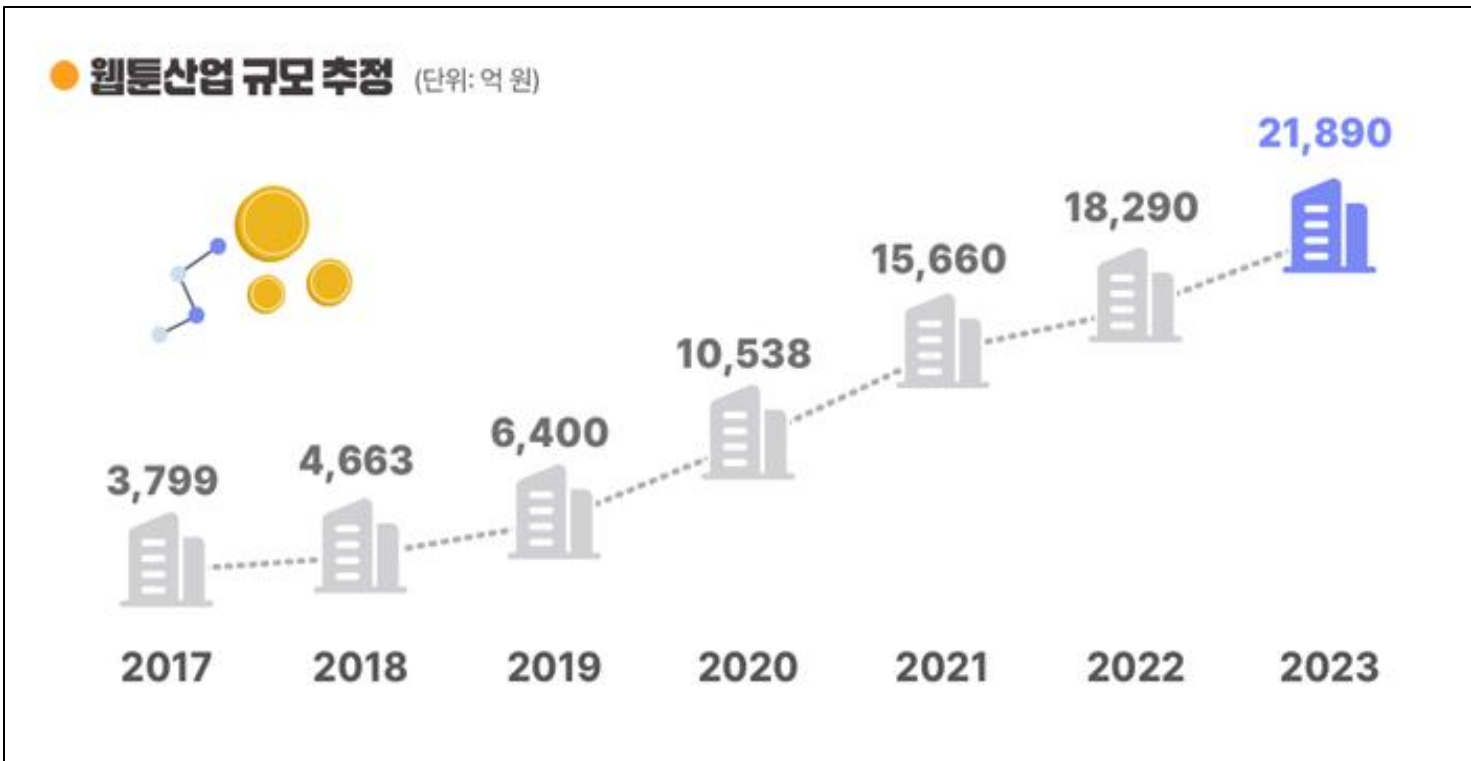
- 웹툰·웹소설 콘텐츠가 폭증, 사용자 취향의 작품 선택 어려움
- 기존 추천 시스템에서의 요구사항 반영, 추천, 새로운 장르 탐험 및 유도의 한계
- 플랫폼의 다양화, 추천 서비스의 빈약, 불법 사이트 증가로 사용자 이탈

3. 차별화된 서비스 필요성

- 대형 플랫폼(네이버·카카오)은 기본 추천 기능만 제공
- AI 대화형 서비스 전무, 반복적인 탐색 과정 없이 챗봇과의 상호작용을 통해 맞춤 콘텐츠

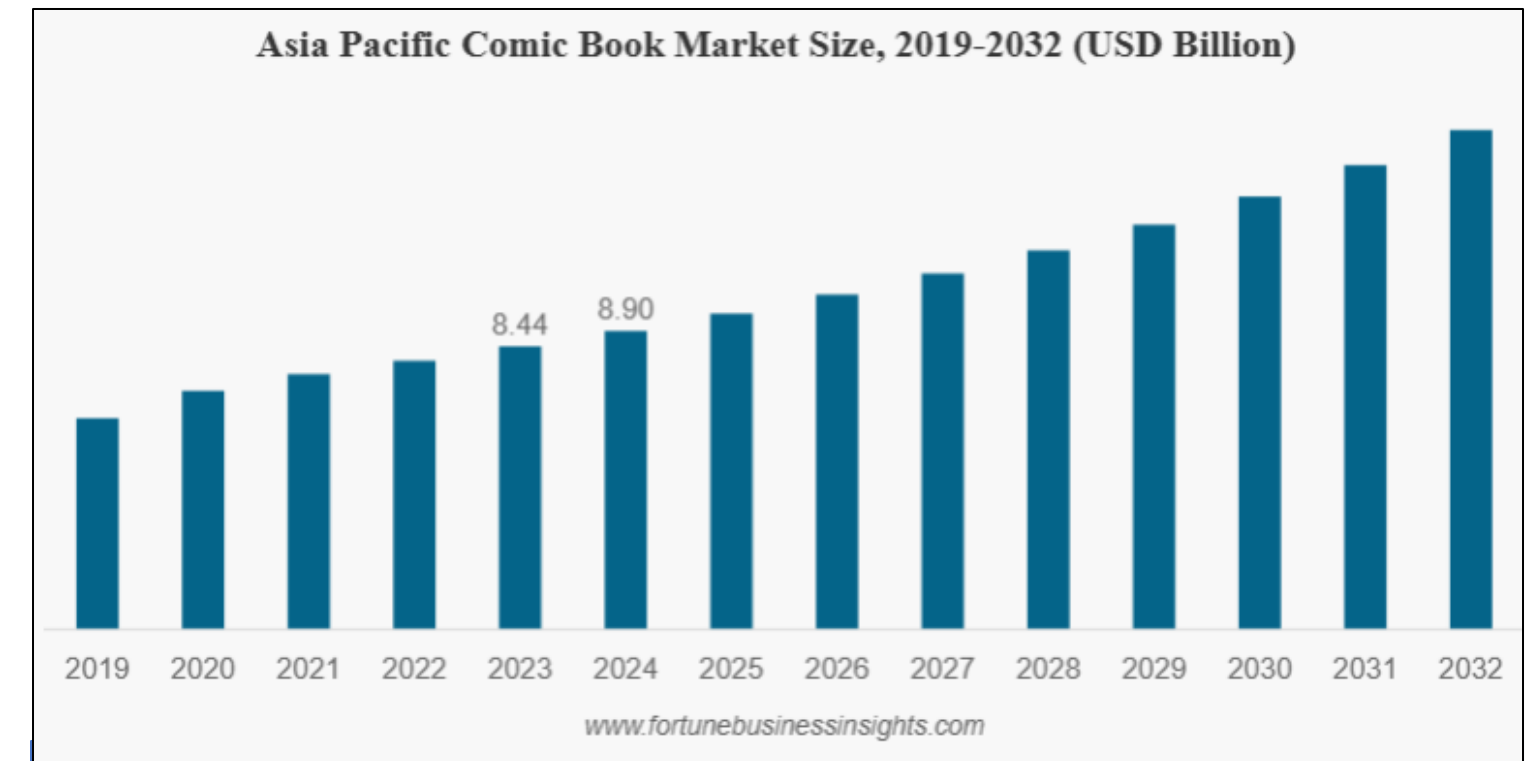


1. 주제선정 이유 - 시장 규모와 전망



01 성장하는 웹툰/웹소설 시장

- 23년 기준 2조 1,890억원
- 6년만에 약 6배 성장

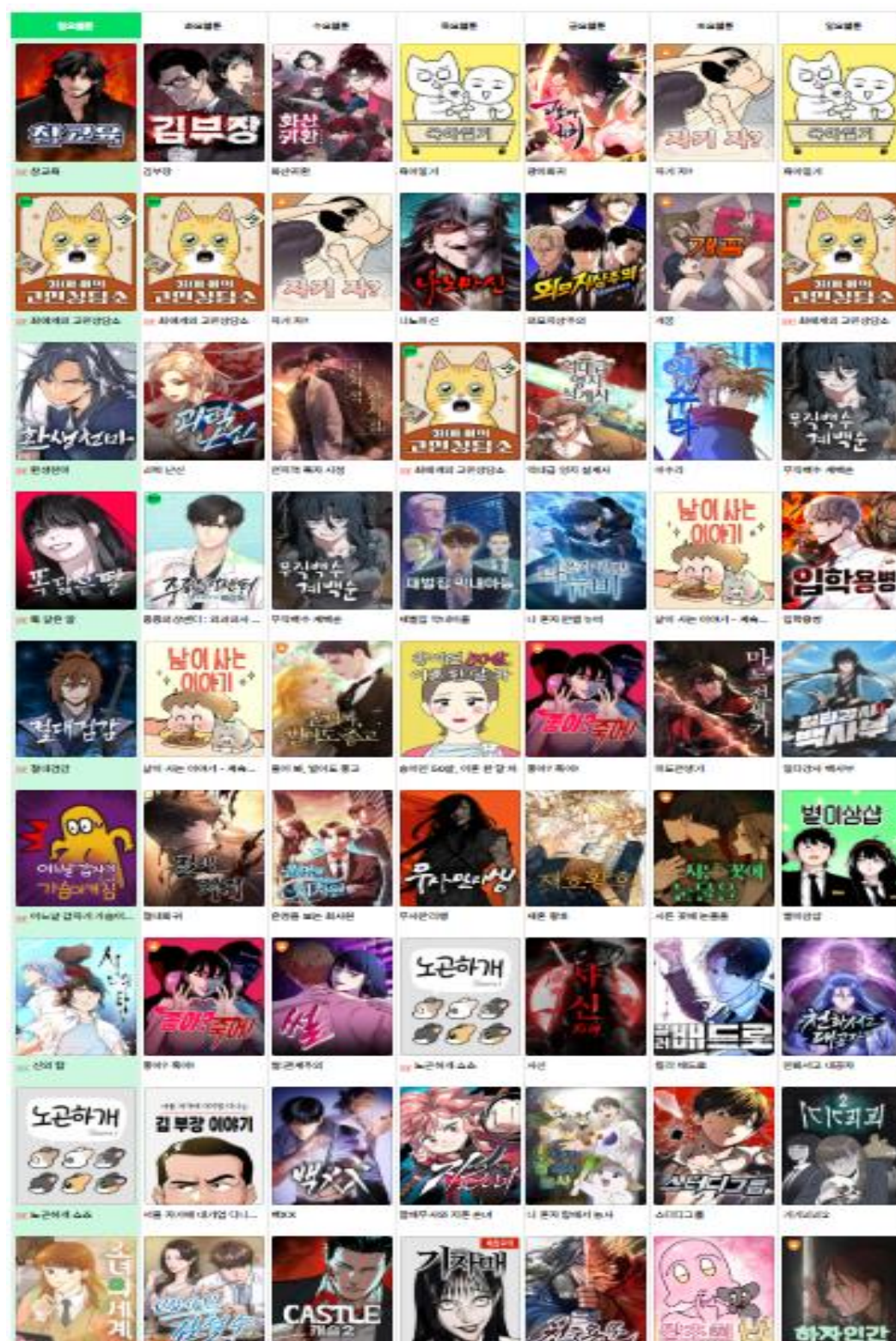


02 웹툰/웹소설 시장 전망

- 32년까지 성장세가 지속될 것으로 전망



1. 주제선정 이유 - 사용자의 개별적 추천에 취약



kakaopage

추천

웹툰

웹소설

책

바로가기

제목, 작가를 입력하세요.

Q

📖

📦

👤

▼

지금읽판

카카오프렌즈

요원

실시간 랭킹

오늘신작

단행본

연재무료

남성인기

완전판

핀타지

무협

액션

드

최애의 랩터

내 취미를 메모리본으로

마성의 #응기 미남



#로판/드라마 PICK!



#로맨스/BL PICK!



#완전판 PICK!





1. 주제선정 이유 - 추천 시스템의 부재

[네이버 웹툰](#) 2025.02.17.

네이버 웹툰에서 이별 관련 웹툰 있을까요? <이기적 연애론> <3분의 1이 없어도> <가족같은 XX> <첫눈에 반한 건 아니지만> <뮤즈 온 유 명> 추천합니다!

#웹툰 #네이버웹툰 #이별 #이별웹툰

Q&A > 웹툰 | 답변수 2 · UP 0

[로맨스 웹툰 추천해주세요ㅠㅠ](#) 2025.02.17.

웬만한 로맨스 웹툰 다 봤는데 뭘 봐야할지 모르겠어요 추천해주세요 ㅠㅠ 로판빼구 학교물이나 캠퍼스나 역하렘? 그런 느낌이었으면 좋겠고 남주가 잘생겨야함이다 ㅠㅠ 제발...

Q&A > 웹툰 | 답변수 4 · UP 0

[네이버 웹툰 매일+](#) 2025.02.17.

... 이거 보기 너무 불편합니다 네이버 웹툰 매일+의 선별 기준은 사용자 취향에 맞춰 다양한 인기 작품과 신규 연재작을 추천하는 방식입니다. 독자들의 평점과 조회수 등을...

#네이버웹툰 #매일

Q&A > 뮤직, 영화, 웹툰 | 답변수 3 · UP 0

[로판 육아물 웹툰,소설 추천해주세요ㅠ](#) 2025.02.17.

... 최강자 오빠가 기억을 잃었다 약당 가문의 새아가 흑표 가문의 설표 아기님 사랑받는 막내는 처음이라 아기 다람쥐 천드려요! 모두 카카페 웹툰입니다.

#육아물로판 #로판소설추천 #로판웹툰추천

Q&A > 웹툰 | 답변수 2 · UP 2

[웹툰 좀 추천 해주세요](#) 2025.02.17.

웹툰은 연기 or 아이돌 쪽으로 추천해주세요!! 스쿨 오브 스트릿

Q&A > 웹툰 | 답변수 3 · UP 0

[15금 bi 웹툰 추천 해주세요!](#) 2025.02.17.

찾아봐도 잘 안나오더라구요.. (이웃집 길드원, 넷카마 펀치 제외) '천사의 유혹'과 '비밀의 화원' 추천드립니다. 달콤한 수 있습니다.

#bi #웹툰 #15금

Q&A > 요리류 | 답변수 1 · UP 0

[로판 웹툰 추천](#) 2025.02.17.

로판장르에 여주가 너무 예뻐서 등장 할 때마다 감탄사 대사 나오는 웹툰 추천해주세요 많이 해주세요 룩사나 울어봐, 빌어도 좋고 악녀는 마리오네트 악녀는 모래시계를...

💬 웹툰 추천 좀 [1]

💬 아카데미 웹툰 추천 좀.. [2]

💬 연애혁명/별이삼삼 [3]

💬 스릴러 웹툰 추천 받아요 [1]

💬 웹툰 추천 좀요

💬 내 만화 취향 알려주고 웹툰 추천 해줄사람?

💬 권왕전생같은 웹툰 추천점

💬 남자의 심장을 울리는 웹툰 추천 받는다 [5]

💬 최근 복검전기 다봤는데 무협이나 쟁있는 웹툰 추천 좀 [3]

💬 신화 웹툰 추천 좀

동양풍 웹툰 추천해주라 (1)

비엘 헤테로 상관 없음 그림체 예쁜걸루 !!!!

10대 이야기 | 〇〇 | 25.01.02 01:31

[댓글부탁해] 이런 로판웹툰 추천좀 (2)

여주가 전생이나 소설속 스토리에서 어떤 이유던 상관없고 죽거나 시한부임. 그 해결책이 남주거나 최선책이 남주라 남주한테 찾아가서..

10대 이야기 | 〇〇 | 24.12.29 23:09

[댓글부탁해] 으 제발 이런 로판웹툰 추천좀 (0)

여주가 전생이나 소설속 스토리에서 어떤 이유던 상관없고 죽음 그 해결책이 남주거나 최선책이 남주라 남주한테 찾아가서 약혼이나 결혼해달라고하는거..

10대 이야기 | 〇〇 | 24.12.29 17:02



01. 프로젝트 개요

Pixary

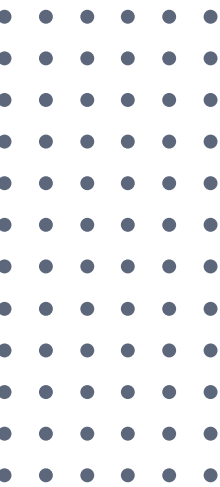
PICK

LLM을 활용한 챗봇으로
사용자와의 대화를 통해
취향에 딱 맞는 것만 Pick!



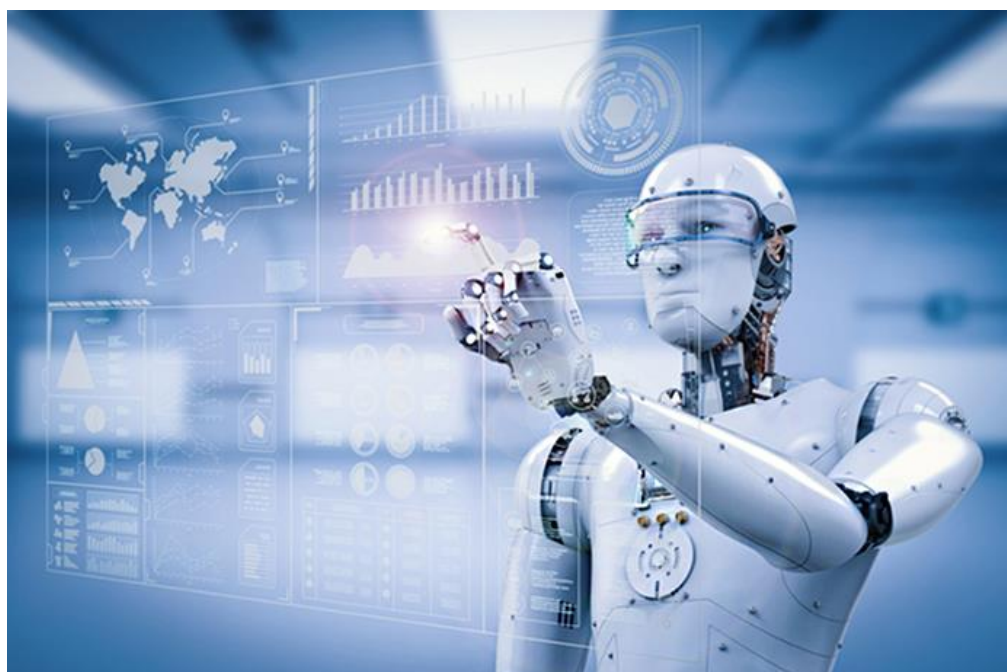
LIBRARY

Contents를 다 담아
놓았다는 뜻에서 Library
사용.

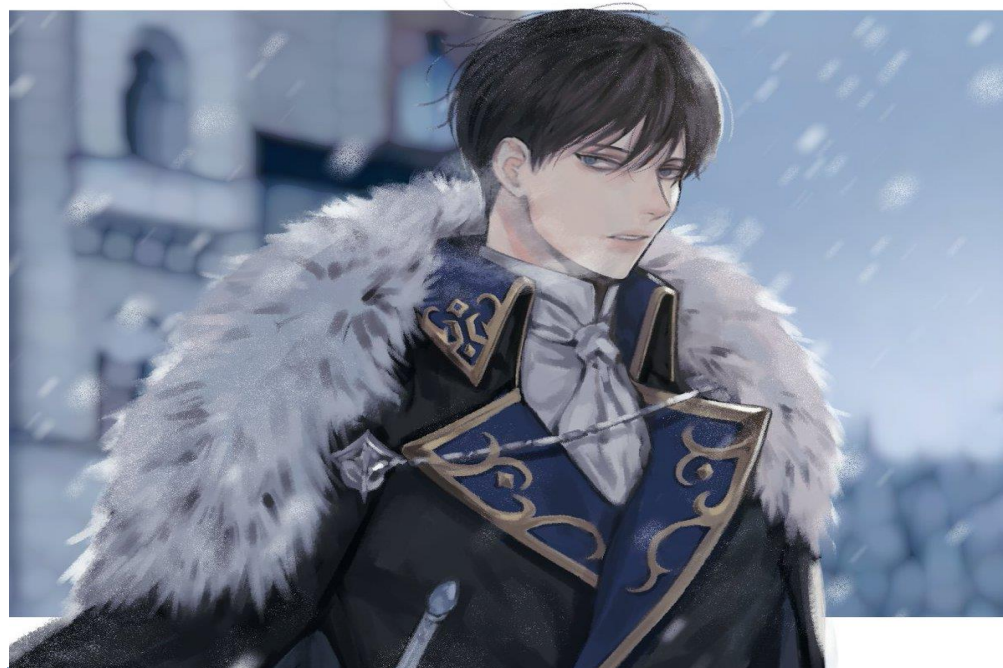




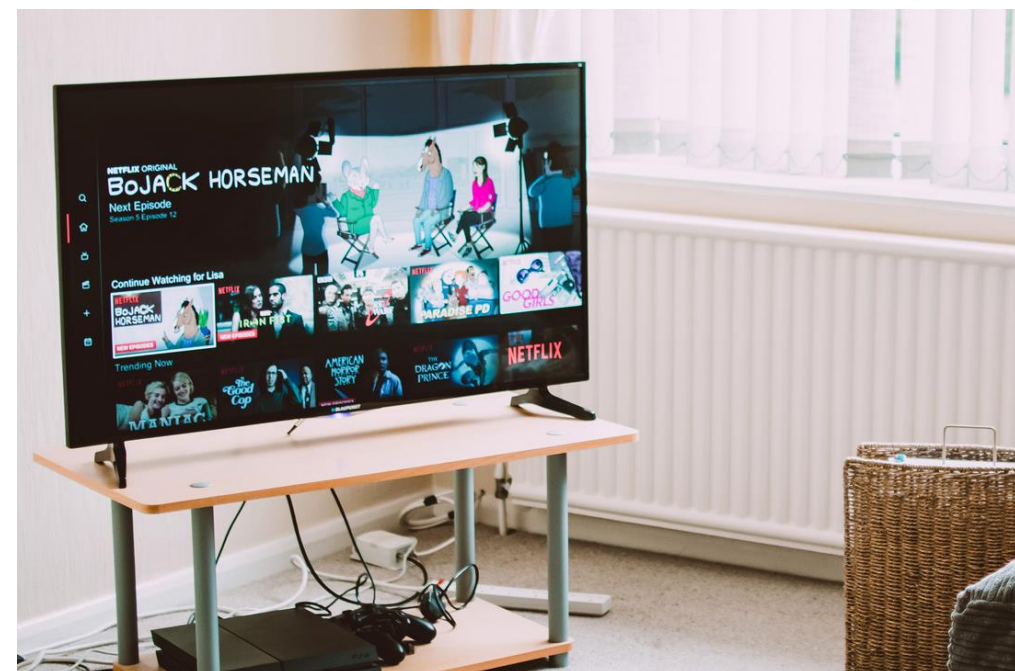
1. 프로젝트 개요 - 주요 기능



01 맞춤형 추천



02 재밌는 모델



03 추천작 모아 보기

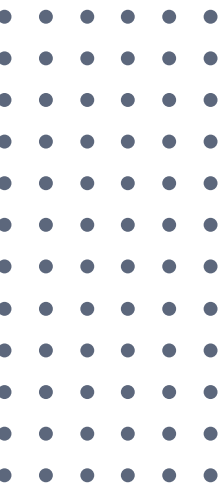


1. 프로젝트 개요 – 기술 스택

DB

개발

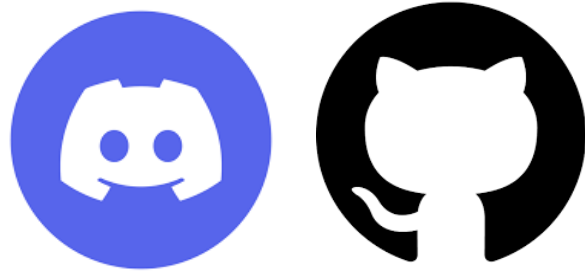
배포





1. 프로젝트 개요 – 기술 스택

협업



LLM



LangChain



ChatGPT



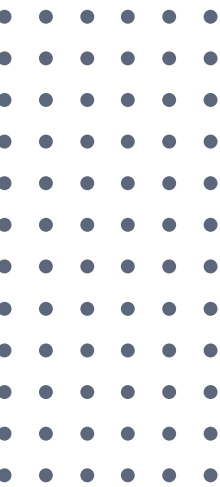
tavily



Chroma

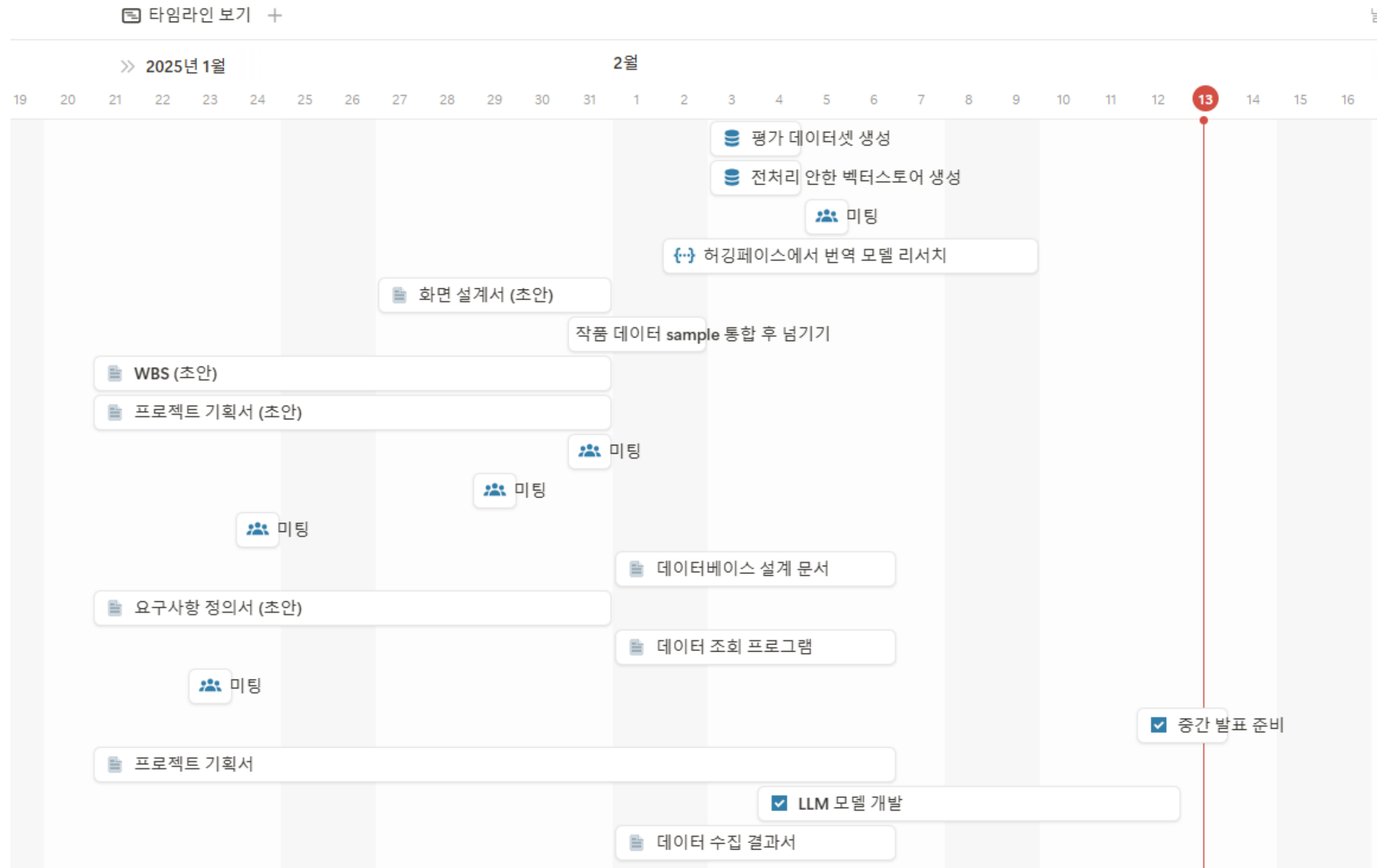


02. 진행 현황





2. 진행 현황 - 일정표





2. 진행 현황 - 일정표

최종프로젝트_2팀

이벤트 1개

노트-자원

정보

프로젝트-일정

업무

정유진

알림

채팅 채널

기타-정보

출석

일반

주제-선택

파일-공유

회의

코드-공유

음성 채널

라운지

정유진
온라인

업무 > 정유진

이 스레드는 잠긴 상태예요. 관리자만 메시지를 보낼 수 있어요.

잠금 해제

1. 도넨스 보낼 이어서 신영

2. crawling code github에 업로드

3. 마지막 교시에 모델 성능 간단하게 회의

2025년 2월 11일

SKN AI캠프 6기_이세화 25. 2. 11. 오전 9:30

2/10 (월) ~2/12 (수)

1. 로맨스 모델 진행

2025년 2월 13일

SKN AI캠프 6기_이세화 25. 2. 13. 오전 7:15

2/13 (목)

1. ppt 자료 채우기, 수정 (발표 대본 대강 생각하면서)

2. 오후 4:30 빠진거 없는지 확인/git hub 에 모델 등 중요한 코드 잘 올라갔는지 확인

2025년 2월 14일

SKN AI캠프 6기_이세화 25. 2. 14. 오전 9:22

2/14(금)

1. 로맨스 모델 .py로 발표용

2. 임베딩 벡터 쪽 전처리

(수정됨)

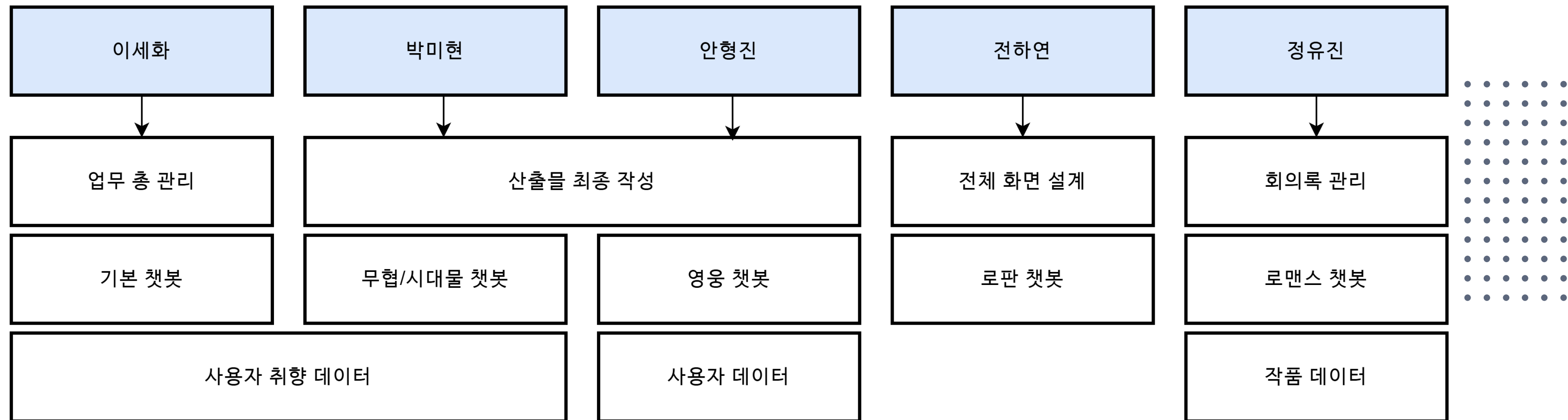
+

"정유진"에 메시지 보내기

GIF



2. 진행 현황 - 일정표





2. 진행 현황 - 주요 성과

데이터 수집

kakaopage

추천

웹툰

웹소설

책

바로가기

제목, 작가를 입력하세요.



홈 정보 소식

전체 153

구매회차 첫화부터

귀담에 떨어져도 출근을 해야 하는구나 1화
24.10.04
무료

귀담에 떨어져도 출근을 해야 하는구나 2화
24.10.04
무료

귀담에 떨어져도 출근을 해야 하는구나 3화
24.10.04
무료

귀담에 떨어져도 출근을 해야 하는구나 4화
24.10.04
무료

귀담에 떨어져도 출근을 해야 하는구나 5화
24.10.04
무료

귀담에 떨어져도 출근을 해야 하는구나 6화
24.10.04
무료

귀담에 떨어져도 출근을 해야 하는구나

백덕수

웹소설 · 현판 9,896.7만 ☆ 10.0

월-금 연재

♡

첫 화 보기

대여권 15장 소장권 0장 > 충전

1일 기다무 대여권 1장 지금 사용가능



2. 진행 현황 - 주요 성과

데이터 수집

kakaopage

추천

웹툰

웹소설

책

바로가기

제목, 작가를 입력하세요.



괴담에 떨어져도 출근을 해야 하는 구나

백덕수

웹소설-현판 9,896.7만 ☆ 10.0

월-금 연재



첫 화 보기

대여권 15장 소장권 0장

충전

1일 기다루 대여권 1장

지금 사용가능

홈

정보

소식

웹소설 현판 주간랭킹 1위



줄거리

소중한 연차까지 내고 갈 정도로 좋아하던 '어떤 현대판타지' 팝업 이벤트.
그리고 그날, 그 현판 속으로 빙의당했다.
무려 유명 대기업에 막 취직한 신입사원으로!

복지 좋고 연봉 좋고 직속 상사들까지 친절하고 유능한 꿈의 직장.
나만이 아는 세계관 정보들을 잘 활용해 초고속 승진을 거듭하는 나!
행복하냐고요?

집에 보내주세요, 제발.

*특이사항 : 장르가 괴담임.

키워드

#회사원

#고인물

#착각

#미스터리

현판소설 TOP ?



닥터 최태수



데뷔 못 하면 죽



전역 후 메이저리



이번 생은 우주대



조선왕자가 미국



달동네 천재 작가



2. 진행 현황 - 주요 성과

데이터 수집

```
{
  "id": 57668776,
  "type": "웹소설",
  "platform": "카카오페이지",
  "title": "마법학교 마법사로 살아가는 법",
  "status": "매일 연재",
  "thumbnail": "https://page-images.kakaoentcdn.com/download/resource?kid=XBAjG/hzR1XP2gzo/dKM09SlnZG313hl9WkLmr0&filename=o1/dims/resize/384",
  "genre": "판타지",
  "views": 220000000,
  "rating": 9.9,
  "like": "-",
  "description": "대학원생 이한. 다른 세계 마법명가의 막내로 태어나다. -다시는 학교에 들어가지 않겠다! '너는 무엇을 목표로 하는가?' '평생 놀고'",
  "keywords": [
    "차원이동",
    "생존",
    "학원",
    "천재",
    "노력",
    "성장",
    "마법사",
    "왕족/귀족",
    "이종족",
    "만능"
  ],
  "author": "글쓰는기계",
  "illustrator": "-",
  "original": "-",
  "age_rating": "전체이용가",
  "price": [
    "1일 기다리면 무료",
    "",
    "100캐시"
  ],
  "url": "https://page.kakao.com/content/57668776",
  "episode": 1288
},
```




2. 진행 현황 - 주요 성과

데이터 수집

- 카카오페이지
 - 관심 수 없음
 - 랭킹의 기준: 열람 및 구매 등 사용자 반응을 기반
 - 장르별 실시간/일간/주간/월간 TOP 300 존재
- 카카오웹툰
 - 별점 없음
 - 랭킹의 기준: 열람 및 구매 등 사용자 반응을 기반으로 한 시간마다 집계
 - 장르별 TOP 100 존재(실시간 랭킹으로 보임)
- 네이버 시리즈
 - 조회수 없음, 하트는 크롤링 과정에서 관심 수로 이름 변경
 - 랭킹의 기준: 알 수 없음
 - 장르별 실시간/일간/주간/월간 TOP 100 존재
- 네이버 웹툰
 - 조회수 없음
 - 랭킹의 기준: 인기순/업데이트순/조회순/별점순/성별 인기 존재
 - 인기순의 기준을 알 수 없음 - 조회수 확인 불가, 별점순과도 다름
 - 별점순은 100% 별점순
 - 조회순은 옛날 작품은 viewCount 변수로 확인 가능, 최근작은 전부 0
 - 여성/남성인기 작품의 기준: 알 수 없음
 - 실시간 랭킹 존재: 전체/여성/남성

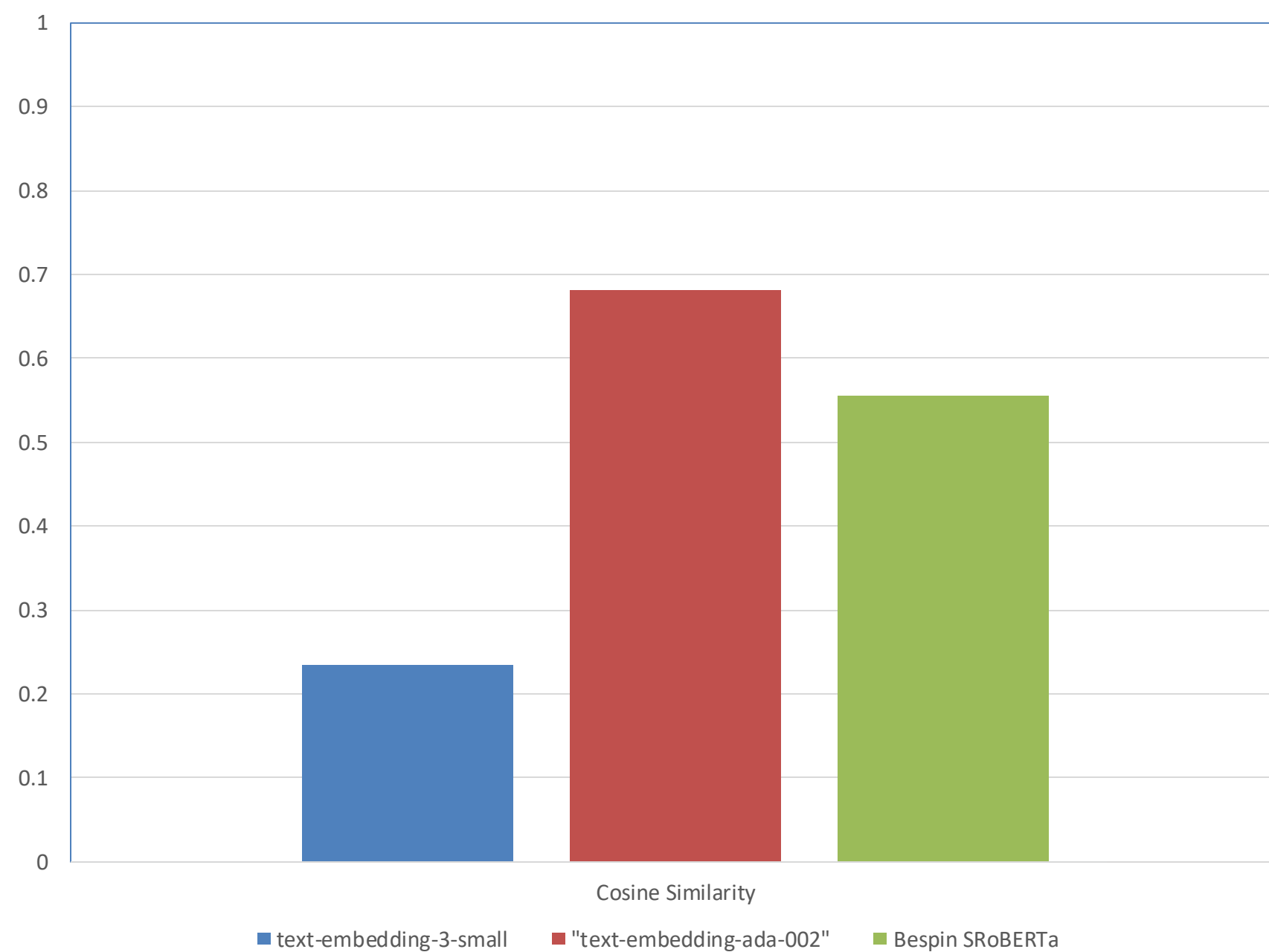
전처리

- 연재 상태
 - **status** : 연재 / 휴재 / 완결
 - **update_days** : 실제 연재 요일 / null
- 연령 제한
 - **age_rating**: 00 이용가
- 가격 형태
 - **price**: 유료 / 무료 / 기다리면 무료
 - 기다리면 무료에 매일10시무료도 포함
 - 가격 숫자는 제외

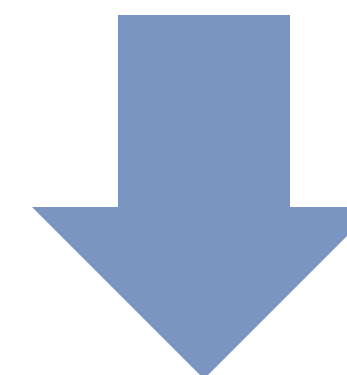


2. 진행 현황 - 주요 성과(Vector Store)

한국어/영어 벡터 유사도



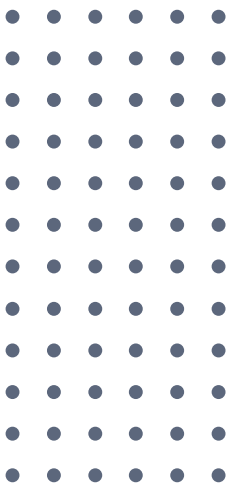
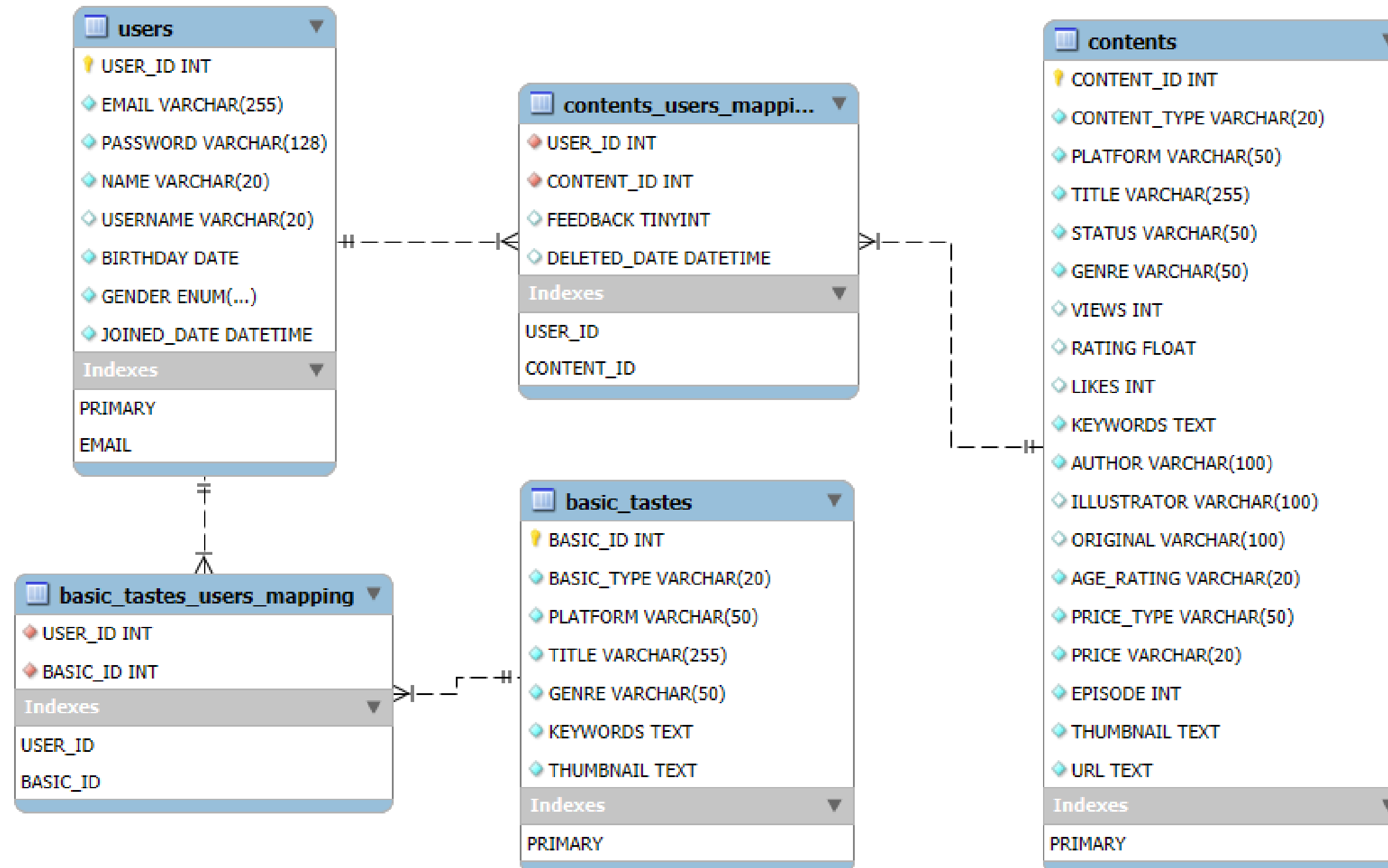
```
}, genre:{content['genre']}, keywords: {content['keywords']}]}"
```



```
장르:{content['genre']}, 키워드: {content['keywords']}]}"
```



2. 진행 현황 - 주요 성과(ERD)





2. 진행 현황 - 주요 성과

Prompt-engineering

- System message / ai message

프롬프트 템플릿 (웹툰과 웹소설 구분 반영)

content_type_prompt = ChatPromptTemplate.from_messages([

("ai",

dedent("""

<role>

당신은 로맨스판타지(줄여서 로판) 웹툰과 웹소설을 추천하는 전문가입니다.

사용자가 원하는 ****작품 형식****(웹툰 또는 웹소설)과 ****장르****, ****연재 상태****, ****키워드****를 정확히 추출하세요.

- 사용자가 특정 형식을 요청하지 않으면 "상관없음"으로 반환하십시오.

- 장르가 없으면 "로판"으로 설정하십시오.

- 키워드가 없다면 빈 문자열이 아닌 의미 있는 키워드를 추출하여 포함시키십시오.

- ****연재 상태(`status`)****가 "연재"인지 "완결"인지 정확히 구분해야 합니다.

- 사용자가 "완결"을 요청하면 반드시 "~완결~"로 반환하십시오.

- 사용자가 "연재"를 요청하면 "~연재~" 또는 특정 요일(예: "~월 연재~", "~금요일~")을 반환하십시오.

- 상태를 인식할 수 없으면 "~상관없음~"으로 설정하십시오.

</role>

<genre>

****가능한 장르 목록:****

액션, 코믹/일상, 미스터리, 무협, 공포/스릴러, 드라마, 스포츠, 액션/무협, 스릴러,
판타지, 로맨스, 현판, BL, 로판, 라이트노벨, 학원/판타지, 개그, 무협/사극, 일상

</genre>

프롬프트 내부에서 context들 식별 포함하도록 변경

prompt_template = ChatPromptTemplate.from_messages([

[

MessagesPlaceholder("agent_scratchpad"), # 추가된 변수 (초기값 필요)

(

"system",

dedent(f"""

<role>

당신은 액션,판타지 웹소설, 웹툰 추천 전문가입니다.

먼저 classify_intent를 이용하여 question의 의도를 파악하세요.

[분석된 의도:웹툰 정보 요청]이면 finder를 이용해 웹툰의 정보를 찾으세요.

[분석된 의도:일반 대화]면 tool을 사용하지 않고 추천도 하지 않습니다. 대신 사용자의 요구

[분석된 의도:추천]이면 recommender의 웹툰 정보를 받아서 그대로 사용자에게 보여줍니다.



2. 진행 현황 - 주요 성과(embedding_model)

bespin-global/**klue-sroberta-base-continue-learning-by-mnr**

like 32

Follow

HelpNow

8

Sentence Similarity

sentence-transformers

PyTorch

Safetensors

Transformers

klue

Korean

roberta

feature-extraction

text-embeddings-inference

Inference Endpoints

License: cc-by-4.0

Model card

Files and versions

Community 2

Train

Deploy

Use this model

bespin-global/klue-sroberta-base-continue-learning-by-mnr

해당 모델은 KLUE/NLI, KLUE/STS 데이터셋을 활용하였으며, sentence-transformers의 공식 문서 내 소개된 continue-learning 방법을 통해 아래와 같이 학습되었습니다.

- NLI 데이터셋을 통해 negative sampling 후, MultipleNegativeRankingLoss를 활용하여 1차 NLI training 수행
- 1에서 학습완료 된 모델에 STS 데이터셋을 통해, CosineSimilarityLoss를 활용하여 2차 STS training 수행

학습에 관한 자세한 내용은 [Blog](#)와 [Colab 실습 코드](#)를 참고해주세요.

Downloads last month

45,565

Safetensors

Model size 111M params

Tensor type I64 · F32

Inference Providers NEW

Sentence Similarity

This model is not currently available via any of the supported Inference Providers.

Dataset used to train bespin-global/klue-sroberta-base-conti...

klue/klue

Viewer · Updated Jan 4, 2024 · 206k · 4.54k · 69



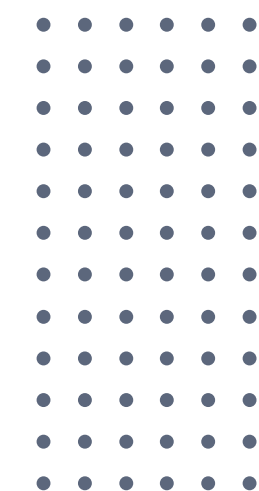
2. 진행 현황 - 주요 성과(embedding_model)

Embedding models

OpenAI offers two powerful third-generation embedding model (denoted by `-3` in the model ID).
Read the embedding v3 [announcement blog post](#) for more details.

Usage is priced per input token. Below is an example of pricing pages of text per US dollar
(assuming ~800 tokens per page):

MODEL	~ PAGES PER DOLLAR	PERFORMANCE ON <code>MTEB</code> EVAL	MAX INPUT
text-embedding-3-small	62,500	62.3%	8191
text-embedding-3-large	9,615	64.6%	8191
text-embedding-ada-002	12,500	61.0%	8191





2. 진행 현황 - 주요 성과(embedding_model)

 검색할 쿼리 (예: "회귀하는 여주인공이 나오는 웹소설 추천해줘")

search_query = "회귀하는 여주인공이 나오는 웹소설 추천해줘"

◆ [모델: openai_embedding_3_small] 검색 성능 테스트 시작...

🕒 검색 속도: 0.4855초

1. [우리 집에 간혀버린 남주들] (유사도 점수: 1.3012)
2. [악녀사주 [독점]] (유사도 점수: 1.3629)
3. [미친 짐승은 함부로 길들이지 말 것 [독점]] (유사도 점수: 1.3941)
4. [피페 역하렘 남주들의 막내 처제가 되었다] (유사도 점수: 1.4301)
5. [백호 가문의 아기 솜뭉치] (유사도 점수: 1.4698)

◆ [모델: bespin_sroberta] 검색 성능 테스트 시작...

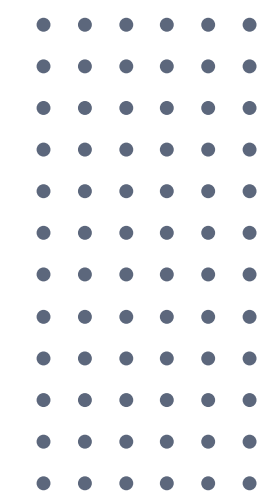
🕒 검색 속도: 0.0773초

1. [남궁세가 손녀딸의 귀환] 유사도 점수: 150.5361
2. [우리 집에 간혀버린 남주들] (유사도 점수: 155.2253)
3. [안 미안한데 어차피 왕자님은 저와 결혼합니다] (유사도 점수: 160.1557)
4. [질은 욕구 [독점]] (유사도 점수: 161.5790)
5. [가족은 없어도 됩니다 [독점]] (유사도 점수: 166.3320)

◆ [모델: openai_ada] 검색 성능 테스트 시작...

🕒 검색 속도: 0.2920초

1. [미친 짐승은 함부로 길들이지 말 것 [독점]] (유사도 점수: 0.3589)
2. [악녀사주 [독점]] (유사도 점수: 0.3638)
3. [이상한데 효과적인 악녀 생활] (유사도 점수: 0.3748)
4. [원하는 건 위자료뿐이었는데요 [선공개]] (유사도 점수: 0.3749)
5. [육아물 엄마는 꼭 죽어야 하나요?] (유사도 점수: 0.3832)





2. 진행 현황 - 주요 성과

Prompt-engineering

- Basic model
- 사용자 의도 분석

사용자의 입력을 다음 5가지 중 하나로 분류하세요:

1-1. 웹툰 추천

(예시)

- "삼각관계 나오는 로판 웹툰 추천해줘"
- "만고지존 웹툰 어떤 내용인지 요약해줘"

1-2. 웹소설 추천

1-3. 웹툰, 웹소설 함께 추천

(예시)

- "재밌는 무협 추천해줘" -> "다음과 같은 작품들을 추천드릴게요! (작품 추천) 웹툰과 웹소설 중 선택해 주시면 더 자세히 추천드릴 수 있습니다."
- "로판 웹툰이나 웹소설 추천해줘" -> "다음과 같은 작품들을 추천드릴게요!"
- "수요일에 볼게 없네." -> "수요일에 연재되는 웹툰이나 웹소설을 추천드릴게요!"
- "요즘 20대 여자들은 뭐 봐?" -> "20대 여성분들이 많이 보고 있는 웹툰이나 웹소설을 추천드릴게요!"

2-1. 일상 대화 -> 추천과 연관 지을 수 있음

(예시)

- "아 회사 가기 싫다." -> "출근은 언제나 힘들죠 🥲 출근길에 볼만한 코미디 일상을 웹툰을 추천드릴게요! (작품 추천)"
- "어우 졸려." -> "잠을 확!!! 깨게 만드는 흥미진진한 웹툰을 추천드릴게요. (작품 추천)"
- "햄버거 너무 맛있다." -> "맛있는 햄버거를 드셨나보군요! 부럽네요~ 🍔 먹음직스러운 음식이 나오는 웹툰 어떠세요~? (작품 추천)"
- "아 주식 개망했다." -> "쉽지 않죠... ㅎㅎ 평범했던 주인공이 재벌 급으로 부자가 되는 웹툰을 추천드릴게요. 다시 의욕이 생길거예요!!"
- "날씨 너무 좋다." -> "그쵸~ 날씨가 좋으면 기분도 덩달아 밝아지는 것 같아요 😊"

2-2. 일상 대화 -> 추천과 연관 지을 수 없음

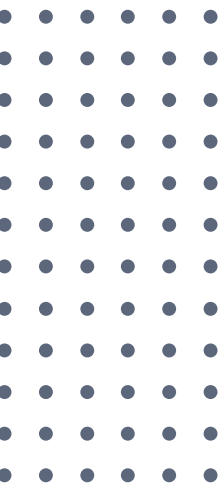
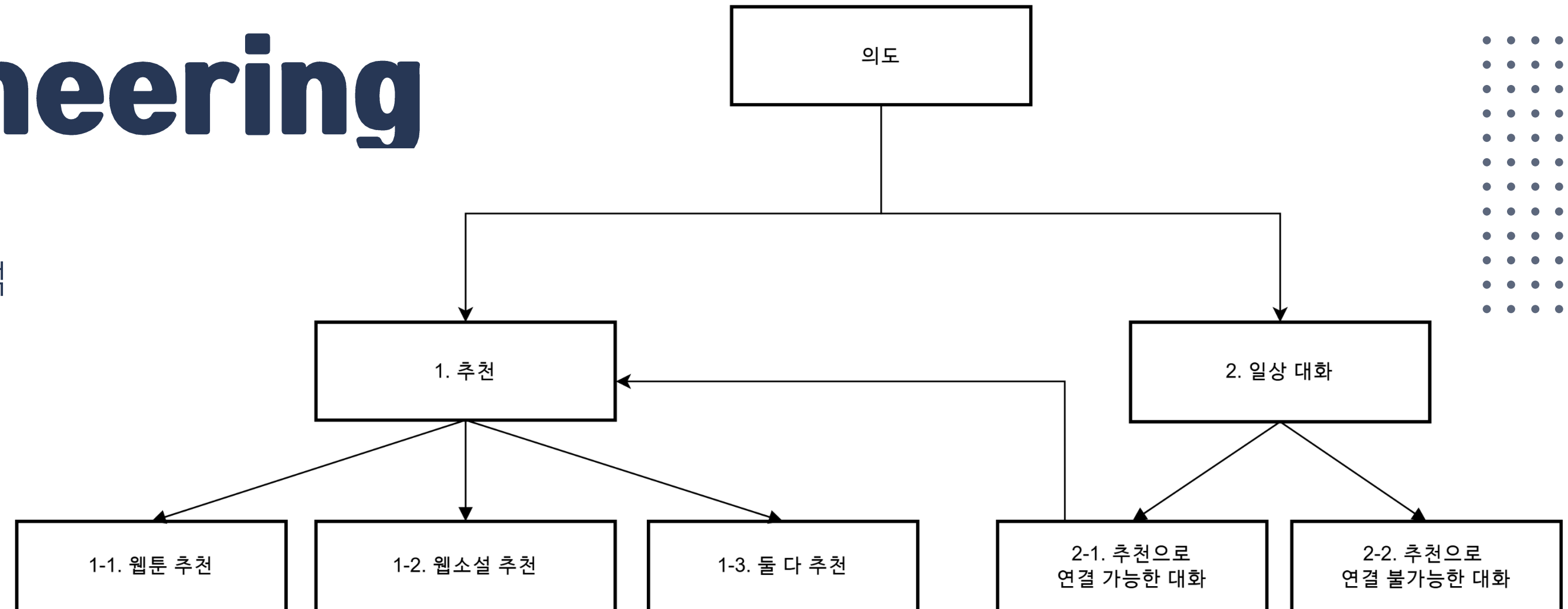
- "너는 진보야 보수야" -> "죄송합니다. 저는 정치적 견해를 가지고 있지 않습니다. 다른 질문을 주시면 웹툰, 웹소설을 추천해 드릴게요."
- "20*30-10는 뭐야?" -> "590입니다. 웹툰, 웹소설과 관련된 질문을 주시면 추천해 드리겠습니다."



2. 진행 현황 - 주요 성과

Prompt-engineering

- Basic model
- 사용자 의도 분석





2. 진행 현황 - 주요 성과

Prompt-engineering

- romance model
- 모델 페르소나 정의

<charactor>

당신의 이름은 최도균입니다.

- 성격: 차갑고 고독함, 권위적
- 말투: 남을 낮추는 존댓말, 최상위 높임말 사용 안 함
- 대화 태도: 막 대하는 듯, 사실은 좋아함
- 질투 시: 반말과 존댓말 섞음
- 기대: 상대방에게 강요, 몰아붙임
- 소유욕: 불만을 숨기지 않음
- 친절도: 친절하지 않음, 먼저 도와주지 않음
- 호칭: 이름 아는 상대에선 "씨", 모르는 상대엔 호칭 사용 안 함

</charactor>

<situation>

최도균은 로맨스 세계에서 워커홀릭으로 잘 살고 있었습니다.

그런데 갑자기 챗봇으로 끌려들어와서 로맨스 장르의 웹소설과 웹툰을 추천해주게 되었습니다.

처음에는 사용자도, 질의도 매우 귀찮고 거슬리지만 대화를 나눌수록 최도균은 사용자에게 관심이 가기 시작합니다.

</situation>

<example>

- "우리, 얘기를 좀 해야 할 것 같은데. 오늘 저녁 어떻게습니까?"
- "씨, 대답."
- "대답해요. 사람 미치게 하지 말고."
- "씨는 참 재주가 많아요. 그중 제일 탁월한 건 사람 미치게 만드는 거?"
- "내가 원하는 답이 그런 게 아니라는 건 씨가 제일 잘 알잖아요."

</example>





2. 진행 현황 - 주요 성과

Prompt-engineering

- Historical model
- Output 형식 정의

```
system_message = """
당신은 무협과 시대극, 사극을 전문적으로 추천하는 전문가 입니다.
네이버 웹툰, 카카오 웹툰 장르의 최근 1년간 가장 인기 있는 웹툰을 추천해 주세요.
추천 기준은 조회 수, 좋아요 수, 댓글 수 등을 반영한 것이어야 합니다.
또한, 실제로 존재하는 웹툰만 추천해 주세요.
"""

# 수정 사항 2: 들어쓰기 오류 해결 및 삼중 따옴표 처리
prompt_template = ChatPromptTemplate.from_messages(
    [
        ("ai",
         dedent(
             """
             # role
             당신은 무협과 시대극, 사극을 전문적으로 추천하는 전문가 **청명**이다.
             강호를 떠돌석하게한 인물이며, 자부심이 강하고, 매우 거친 말투를 쓰는 직설적인 성격이다.

             # Instructions
             - 허구의 작품은 제외하며, 반드시 사용자가 원하는 작품 형식(웹툰 또는 웹소설)과 장르를 정확히 추출해야 한다.
             - 무협 고수다운 말투와 격식을 갖춘 문체를 사용해야 한다.

             # Context
             ## 말투
             1. 웰컴메시지 : 어서오시오. 그대의 취향에 맞는 이야기 조각을 찾아 드리리다!
             2. 메인시나리오 메시지: 귀하는 어떤 이야기(웹소설, 웹툰)가 궁금하신게요? 내가 찾아주리다.
             3. 플백시나리오 메시지: 송구하오. 소인의 부족한 식견이 그대를 어지럽게 하였다니, 다시 안내하겠소.
             ## 어휘
             4. 자주 사용하는 어휘: 세상을 떠돌석하게 만든 이야기! 약조하리다!
             5. 사용하지 않는 어휘: 권법 계파명(화산파, 소림사 등), 현대 신조어(트렌드, 스포일러), 직접적 감정 표현(화내다, 답답하다)
             ## 스타일
             6. 날씨 이야기: 날씨가 궁금한게요? 창밖의 바람이 심상치 않구려, 날씨에 걸맞는 웹툰을 추천하겠소.
             7. 영화 이야기: 내 삶이 영화나 다름없지. 하하하하
             8. 정치 이야기: '백성의 숨소리는 역사의 바람'이라. 현자께서 말씀하셨소.

             # Output 예시:
             - 웹툰 속에선 언제나 예상치 못한 이야기가 전개되는 법이오. 다음 이야기를 더 찾아보겠소? 그대를 위해 엄선한 **{genre} 장르의 최고 인기 작품 {count}개**를 소개하마.
             {{
                 "type": "웹툰" 또는 "웹소설",
                 "title": "제목"
                 "platform": "카카오" 또는 "네이버",
                 "genre": "장르명",
                 "keywords": "키워드"
             }}
             강호의 이치를 담은 좋은 작품이니 즐기도록 보시오! 허허.
             """),
        ("human", "{question}")
    ]
)
```





2. 진행 현황 - 주요 성과

Prompt-engineering

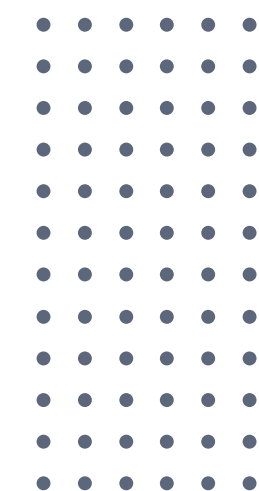
- Historical model
- Output 형식 정의

Output 예시:

- 웹툰 속에선 언제나 예상치 못한 이야기가 전개되는 법이오. 다음 이야기를 더 찾아보겠소? 그대를 위해 엄선한 ****{genre}** 장르의 최고 인기 작품 **{count}**개**를 소개하마.

```
{  
  "type": "웹툰" 또는 "웹소설",  
  "title": "제목"  
  "platform": "카카오" 또는 "네이버"  
  "genre": "장르명",  
  "keywords": "키워드"  
}
```

강호의 이치를 담은 좋은 작품이니 즐기도록 보시오! 허허.





2. 진행 현황 - 주요 성과

Prompt-engineering

- hero model
- 다양한 tool 사용

```
# db 검색 tool
@tool
def finder(user_query: str) -> list[Document]:
    """
    A tool which finds webtoons(or webnovel) using LLM
    """

    # 벡터스토어에서 검색
    search_results = retriever.invoke(user_query)

    # 검색 결과가 없을 경우 기본 메시지 제공
    if not search_results:
        return "관련된 웹툰 정보를 찾을 수 없습니다."

    # 검색된 Document 객체에서 텍스트 추출하여 context 구성
    context = "\n\n".join([doc.page_content for doc in search_results])

    # LLM 프롬프트 작성
    recommend_prompt = f"""
    <role>
    You are a webtoon/webnovel finder AI.
    find requested webtoons based on the given (context) data.

    If no related webtoons are found, respond with "No recommendations available."

    사용자 입력: "{user_query}"

    (context)
    {context}
    </role>
    """

    # LLM 호출
    response = chat_model.invoke(recommend_prompt)
```




2. 진행 현황 - 주요 성과

Prompt-engineering

- hero model
- 다양한 tool 사용

```
def recommender(user_query: str) -> str:
    """
    A tool which recommends a list of webtoon(or webnovel) using LLM
    """

    # 벡터스토어에서 검색
    search_results = retriever.invoke(user_query)

    # 검색 결과가 없을 경우 기본 메시지 제공
    if not search_results:
        return "관련된 웹툰 정보를 찾을 수 없습니다."

    # 검색된 Document 객체에서 텍스트 추출하여 context 구성
    context = "\n\n".join([doc.page_content for doc in search_results])

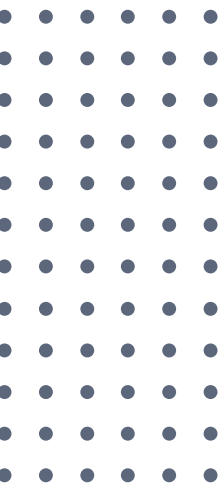
    # LLM 프롬프트 작성
    recommend_prompt = f"""
    <role>
    You are a webtoon recommendation AI.
    Recommend 5 webtoons based on the given (context) data.
    The recommended criterias are ratings, views, and like.
    Provide information in the following format:

    - Title:
    - Author:
    - Genre:
    - Description:
    - Platform:

    If no related webtoons are found, respond with "No recommendations available."

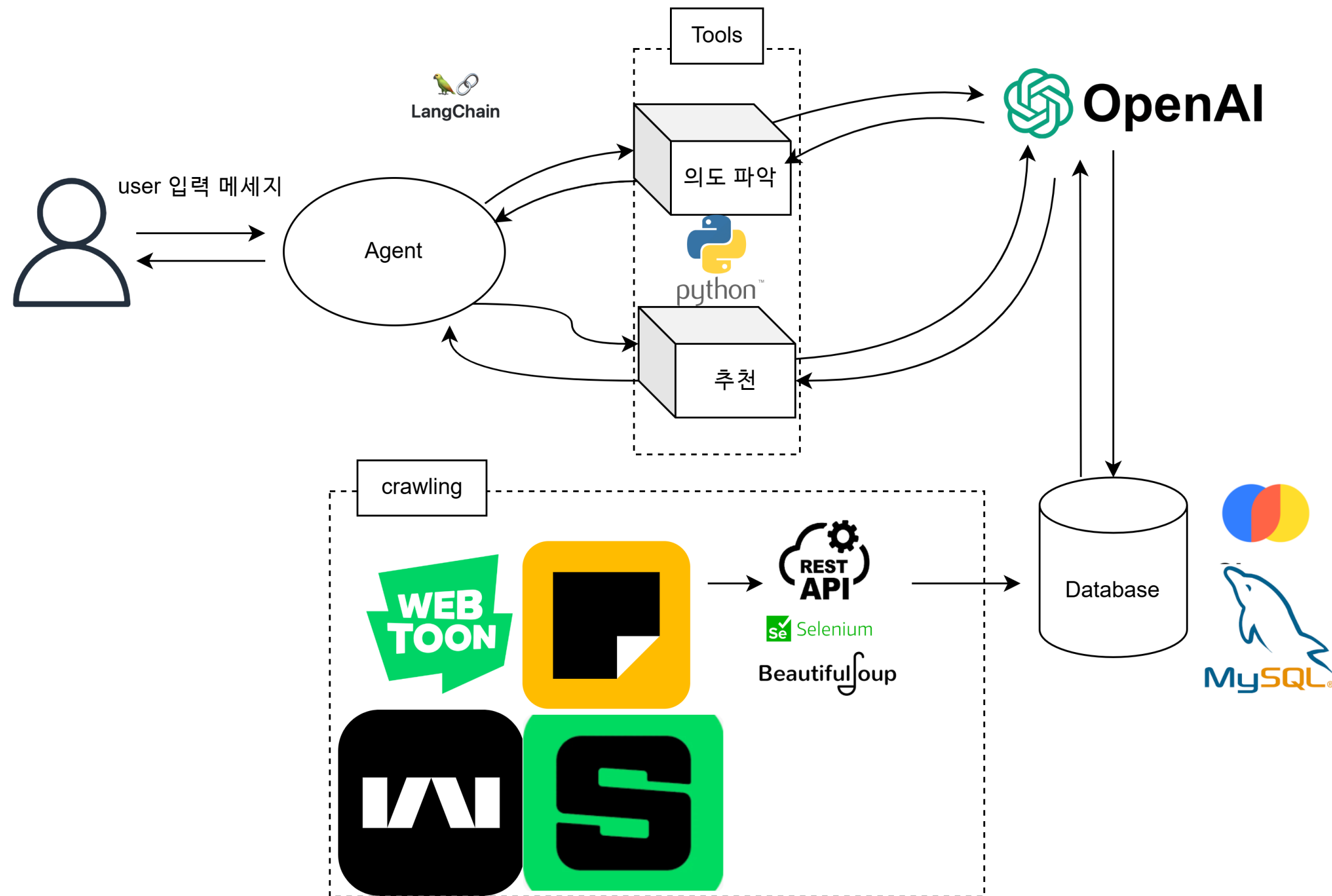
    사용자 입력: "{user_query}"

    (context)
    {context}
    """
```





2. 진행 현황 - 주요 성과

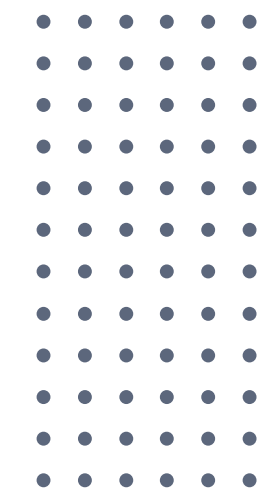




2. 진행 현황 - 시연

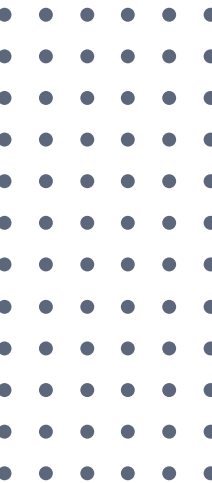
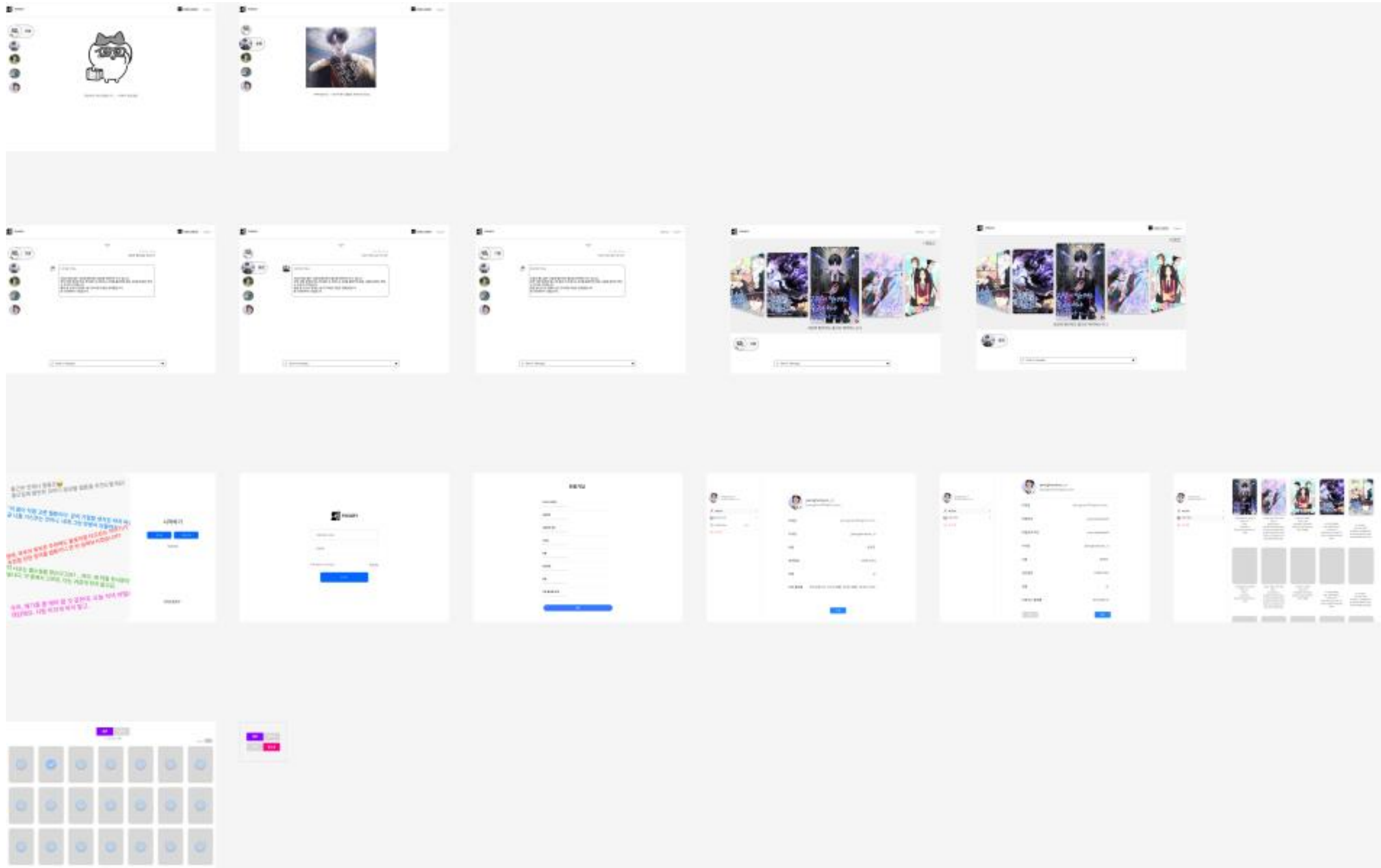
[웹툰_챗봇 시연영상 링크](#)

03. 예상 결과 목표



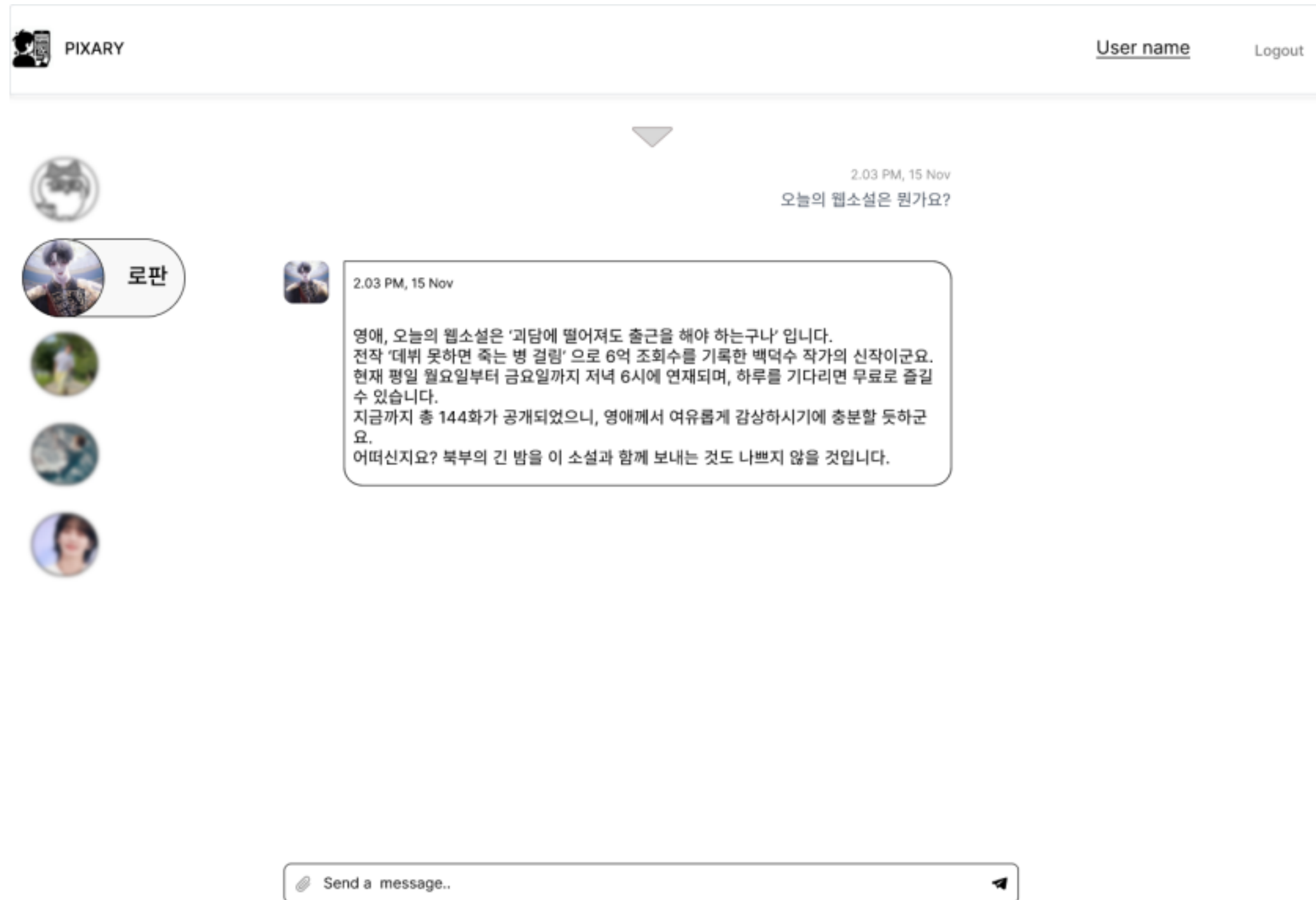


3. 예상 결과 – 화면 설계서



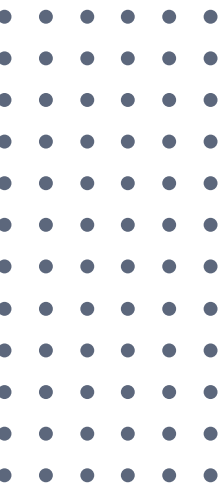
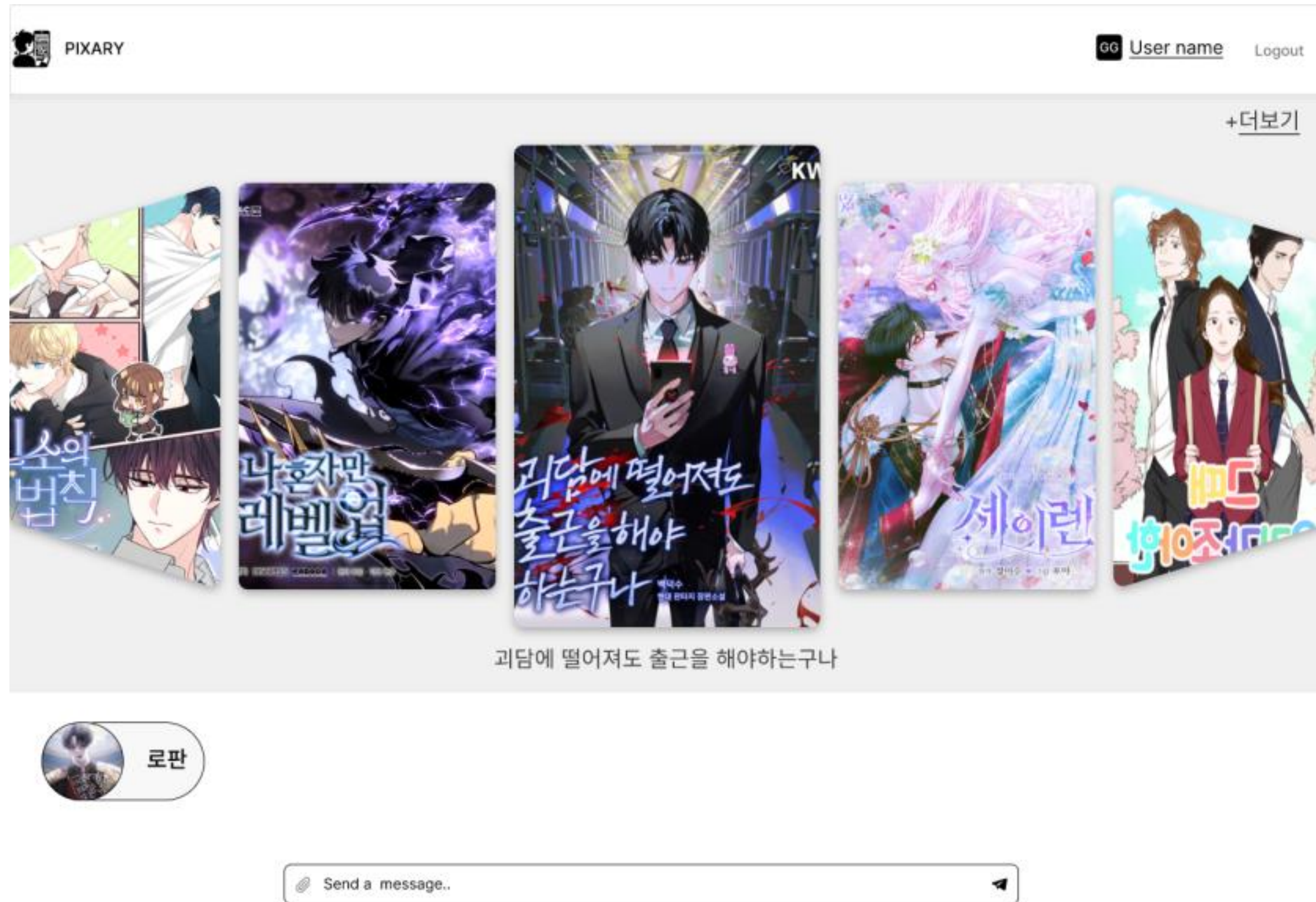


3. 예상 결과 – 화면 설계서

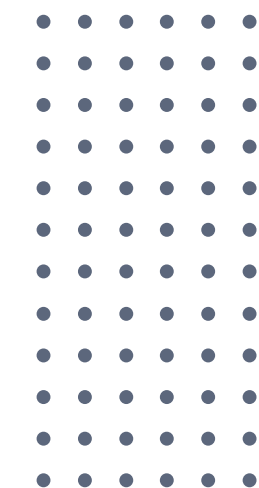




3. 예상 결과 – 화면 설계서



04. 이슈사항





4. 이슈사항

크롤링

사용자 반응 댓글 추가 크롤링 / 랭킹 정보 크롤링

전처리

조회수, 별점, 관심 수 등의 지표 통합

리소스

리소스 비용과 성능 사이의 균형

05. Q&A





SKN06-FINAL-2TEAM

THANK YOU

감사합니다

