

# 하루의 마침표.



#감정분석 #일기 #나에게 딱 맞는 콘텐츠 추천

NLP-01  
(대답해조)

# 목차

1. Intro
2. Model / Research
3. Data
4. Product Serving
5. Result / Conclusion
6. Q & A

# 팀원 소개

대 답 해 조 !



강범서

# Back-end part  
# Modeling part



오필훈

# Back-end part  
# Database, Cloud



이예진

# 데이터 수집  
# Front-end part



한기백

# 데이터 수집  
# Project Manager



정유리

# Front-end part  
# 프로젝트 진행상황 기록

# 1. Intro

# Project Overview

KYOBO 교보문고

출간속시인문위 | 무엇이 좋은가 | 육합방의문제아들 | 경제기사 활용법

30만 독자를 사로잡은 인생 철학책!!

국내도서 외국도서 eBook 웹소설 기프트 음반 중고장터

참가만 해도 e캐시 드림!

**교보eBook Summer CAMP!**

교보e캐시 1,000원

무제한 이용권 1개월

요즘 이 책 | 핫이슈 | 새로나온책 | 기획/테마 | eBook | sam | 한정특가

존재만으로도 충분히 관점은 당신에게

어제까지 몰랐던 세상이 오늘 밤, '쑥' 들어온다!

영화 추천

랜선 영화 인기 순위

넷플릭스 웨이브 티빙 시리즈온 왓차

투깝스 인 파리: 더 테이커다운 ★ 4.12

베놈 2: 핫 테어 비 카니지 ★ 6.40

야차 ★ 5.44

사랑이 지나간 자리 ★ 8.77

365일: 오늘 ★ 4.03

사람들이 많이 검색한 영화

드라마 판타지 공포 로맨스 스릴러 코미디 액션 SF 애니메이션

1 브로커 ★ 8.17

2 해어질 결심 ★ 7.28

3 이터널 로드 ★ 8.17

4 오마주 ★ 7.90

2022.06.05 기준

찾아쉬우 개봉예정 영화

휴식 집중 출퇴근 & 등하교 에너지 충전

마음이 편안해지는 음악

**맞춤 믹스**

Chill Supermix

Chill Mix 01

휴식 슈퍼믹스 재생목록 - YouTube Music

나만을 위한 휴식 믹스 1 나만 유티 박지

아이유(IU), Paul Kim, 10CM, 헤이즈(Heize)

이 노래로 뮤직 스테이션 시작하기

**빠른 선곡**

10cm - 비가오니까(demo) 미발매곡

10CM - 조희수 58만회

담아 갈게 아무진

이 밤이 다 가기 전에 오늘이 어제가 되기 전에 10cm

apple meeno

# Project Overview



내가 현재 느끼고 싶은 감정

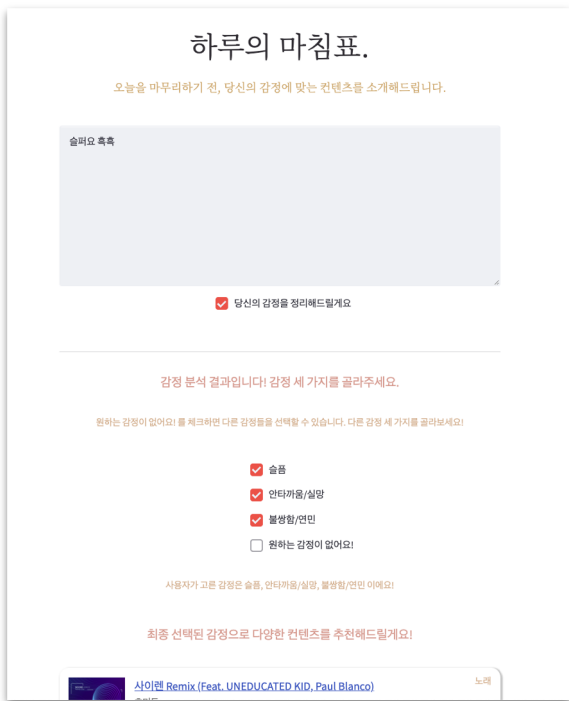
# Project Overview

“당신의 하루를 들려주세요. 오늘을 담은 선물을 드릴게요.”

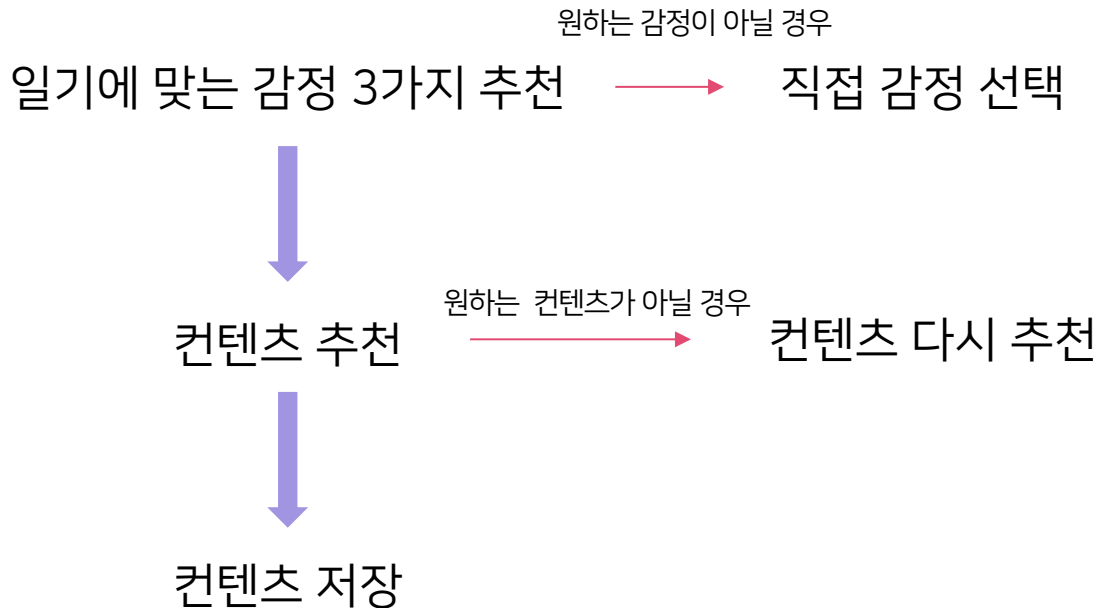
일기 텍스트에서 추출된 감정에 기반해 콘텐츠를 추천하는 서비스.



# Project Overview



## 일기 쓰러가기





# Project Overview

하루의 마침표.

"당신의 하루를 들려주세요. 오늘을 닮은 선물을 드릴게요."

일기 쓰러가기

길을 걸었다. 따뜻한 햇살과 푸르른 잎사귀. 벌써 여름이 다가오고 있음이 느껴진다. 행복한 주말의 첫 날. 오늘을 어떻게 보내야 할지 알고 할 수 있을까? 소소하게 산책하는 것만으로도 행복을 느낀다. 그렇다. 여름이든.★

#기쁨 #행복 #즐거움/신남  
2022-06-05 12:43:51

당신의 밤을 장식할 노래 한 권

이상현  
비스키 비스키  
웨이 브래드 안전 부리를 걸 지기 데어 어디 웨어 할 형 이계 뭐야 이게 아 왜 좋구나 그...

Blueming

오늘을 마무리할 책 한 권

또다시 같은 꿈을 꾸었어 (일반판)  
스미노 요루  
「너의 취장을 먹고 싶어」의 작가 스미노 요루가 지친 사람들에게 선물하는 다정하고 따스한

당신의 하루를 닮은 또 다른 누군가의 하루를 영화로

아직 사랑하고 있습니까?  
의리로 사는 10년 차 부부 '연경'과 '영욱' 빨간 불 들어온 그들 사이에 색()다르게...

코다  
24/7 함께 시간을 보내며 소리를 들을 수 없는 가족을 세상과 연결하는 코다 '루비'는 ...

오늘과 닮은 주말을 선물할 공연

어린왕자와 함께 그리는 영작 클래식 콘서트 [피아니스트 김현정 클로리아]  
'더 스퀘어'라는 새로운 전시를 앞둔 스텝출름 현대 미술관 수석 큐레이터 '크리스티안' 누...

## 피드 페이지

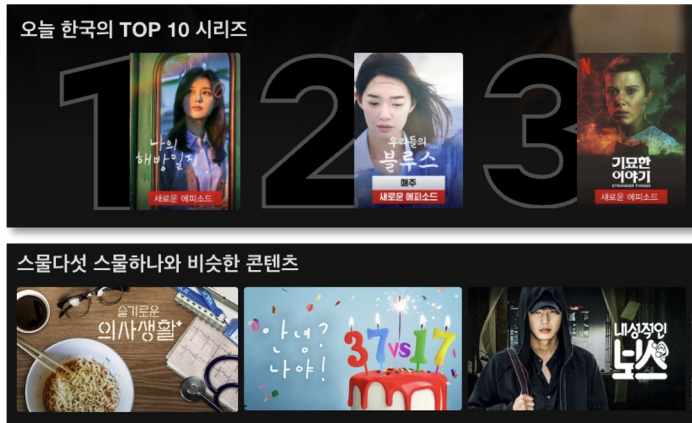
당신의 밤을 장식할 노래  
오늘을 마무리할 책 한 권  
당신의 하루를 닮은 또 다른  
누군가의 하루를 영화로  
오늘과 닮은 주말을  
선물할 공연

# Project Overview

## 하루의 마침표 차별성

### 타사 콘텐츠 추천 방식

- 인기도 기반 (예: 요즘 뜨는 콘텐츠)
- 사용자가 선택 콘텐츠 기반 (예: 비슷한 콘텐츠)



### 하루의 마침표

- 일기를 통한 사용자 감정 분석
- 감정기반 다양한 문화 콘텐츠 추천

사용자가 고른 감정은 고마움, 행복, 기쁨 이에요!

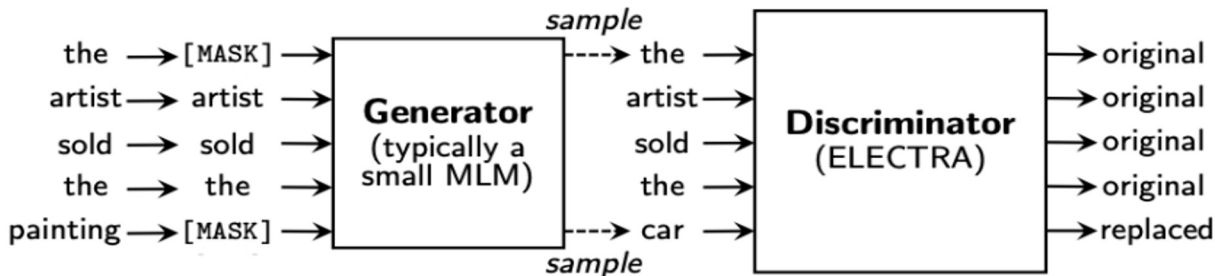


## 2. Model / Research

# Model / Research

## KcELECTRA(Korean comments ELECTRA)

- ELECTRA(Efficiently Learning an Encoder that Classifies Token Replacements Accurately)를 한국어 댓글을 이용하여 학습한 모델
- 신조어, 오타자 등 구어체의 특징을 가진 데이터셋으로 훈련



출처 : ELECTRA: Pre-training Text Encoders as Discriminators Rather Than Generators

# Model / Research

## KcELECTRA(Korean comments ELECTRA)

	Size (용량)	<b>NSMC (acc)</b>	Naver NER (F1)	PAWS (acc)	KorNLI (acc)	KorSTS (spearman)	Question Pair (acc)	KorQuaD (Dev) (EM/F1)
KcELECTRA- base	475M	<b>91.71</b>	86.90	74.80	81.65	82.65	<b>95.78</b>	70.60 / 90.11
KoELECTRA- Base-v3	423M	90.63	<b>88.11</b>	<b>84.45</b>	<b>82.24</b>	<b>85.53</b>	95.25	<b>84.83 / 93.45</b>

# Model / Research

## KcELECTRA(Korean comments ELECTRA)

- 기존의 감정 분류법(taxonomy)은 외국의 감정 분류법을 따름
- KcELECTRA에서는 한국인 감정에 알맞은 43개(+No Emotion)의 감정으로 분류한 데이터셋으로 훈련됨.

➡ KcELECTRA가 이번 task에 적합

# Model / Research

## KcELECTRA(Korean comments ELECTRA)

입력

사용자의 일기

- 일기에 작성되는 내용은  
문어체보다 구어체에 훨씬 가까움



KcELECTRA 모델

### 문어체

글에서 주로 쓰는 말투



내일은 비가 오지 않는다고 합니다.

발언의 진의 파악이 불가능하다.

이것은 본인의 소유물이다.

저는 지시를 이행할 의사가 없습니다.

### 구어체

일상적인 대화에서 주로 쓰는 말투



내일은 비가 안온데~

뭔 말인지 모르겠다!

이건 내꺼야!

저는 안할 겁니다!

# 3. Data



# Data

- '음악', '책', '영화', '공연'의 4가지 도메인에 대해 데이터 수집
- Selenium을 활용한 Web Crawling 사용

음악



책



영화



공연



# Data

## 음악

순위	곡정보	앨범	좋아요	듣기	듣기	다운	유비
1	That That (prod. & feat. SUGA of BTS) (PSY)	SEID9	87,660	▶	+	↓	□
2	TOMBOY (아이브)	I NEVER DIE	177,993	▶	+	↓	□
3	LOVE DIVE (아이브)	LOVE DIVE	133,252	▶	+	↓	□
4	사랑인가 봐 (사내방송 OST) (정유진)	사랑인가 봐 (사내방송 OST)	122,211	▶	+	↓	□

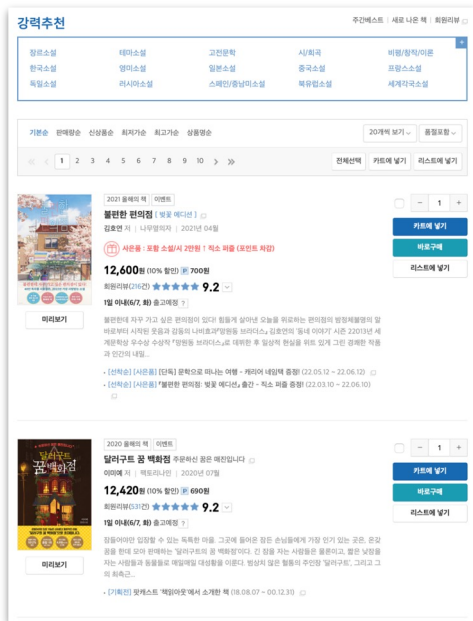
- 2022년 6월 기준 멜론 Top 100
- 2011년 ~ 2021년 멜론 Top 50

제목, 가수, 가사, URL, 이미지

600개

# Data

## 책



- 'yes24'의 소설/시 카테고리 중 '강력추천' 도서
- 최근 인기작부터 고전까지 다양한 시대의 도서를 포함하고 있으며 그 장르 또한 다양하다는 점에서 선정

제목, 저자, 소개 글, URL, 이미지

238개

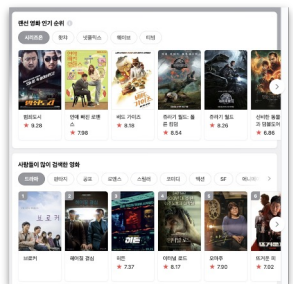
# Data

## 영화

For our poll to determine the 100 greatest American films, we surveyed 62 film critics from around the world. This time, we received responses from 177 – from every continent except Antarctica. Some are newspaper or magazine reviewers, others write primarily for websites; academics and cinema curators are well-represented too. For the purposes of this poll we have decided that a list of the greatest films of the 21st Century should include the year 2000, even though we recognise that there was no 'Year Zero' and that 2001 is mathematically the start of the century. Not only did we all celebrate the turn of the millennium on 31 December 1999, but the year 2000 was a landmark in global cinema, and, in particular, saw the emergence of new classics from Asia like nothing we had ever seen before.

We believe that the new classics on this list are destined to become old classics. Whether or not that happens is ultimately up to you, the moviegoers. But one thing is certain: cinema isn't dying, it's evolving.

100. Toni Erdmann (Maren Ade, 2016)
100. Requiem for a Dream (Darren Aronofsky, 2000)
100. Carlos (Olivier Assayas, 2010)
99. The Gleaners and I (Agnès Varda, 2000)
98. Ten (Abbas Kiarostami, 2002)
97. White Material (Claire Denis, 2009)
96. Finding Nemo (Andrew Stanton, 2003)
95. Moonrise Kingdom (Wes Anderson, 2012)
94. Let the Right One In (Tomás Alfredson, 2008)
93. Ratatouille (Brad Bird, 2007)
92. The Assassination of Jesse James by the Coward Robert Ford (Andrew Dominik, 2007)
91. The Secret in Their Eyes (Juan José Campanella, 2009)
90. The Pianist (Roman Polanski, 2002)
89. The Headless Woman (Lucrecia Martel, 2008)
88. Spotlight (Tom McCarthy, 2015)
87. Amélie (Jean-Pierre Jeunet, 2001)
86. Far From Heaven (Todd Haynes, 2002)
85. A Prophet (Jacques Audiard, 2009)
84. Her (Spike Jonze, 2013)
83. A.I. Artificial Intelligence (Steven Spielberg, 2001)
82. A Serious Man (Joel and Ethan Coen, 2009)
81. Shame (Steve McQueen, 2011)
80. The Return (Andrey Zvyagintsev, 2003)
79. Almost Famous (Cameron Crowe, 2000)
78. The Wolf of Wall Street (Martin Scorsese, 2013)
77. The Diving Bell and the Butterfly (Julian Schnabel, 2007)
76. Dogville (Lars von Trier, 2003)
75. Inherent Vice (Paul Thomas Anderson, 2014)
74. Spring Breakers (Harmony Korine, 2012)
73. Before Sunset (Richard Linklater, 2004)
72. Only Lovers Left Alive (Jim Jarmusch, 2013)
71. Tabu (Miguel Gomes, 2012)
70. Stories We Tell (Sarah Polley, 2012)
69. Carol (Todd Haynes, 2015)
68. The Royal Tenenbaums (Wes Anderson, 2001)
67. The Hurt Locker (Kathryn Bigelow, 2008)



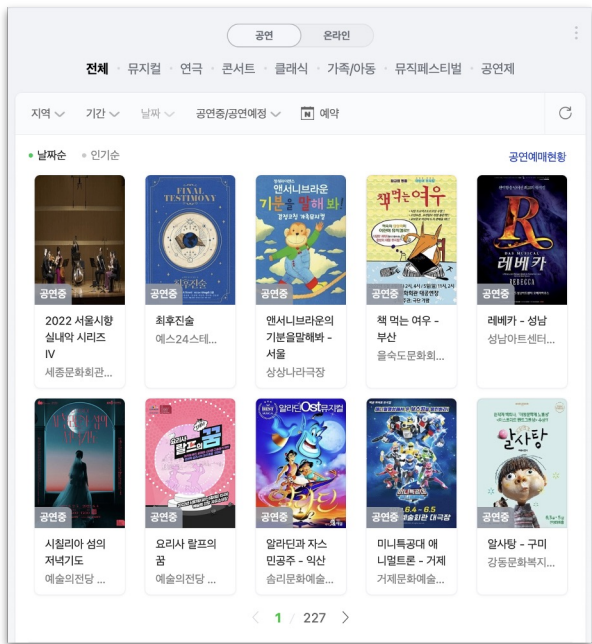
- BBC 선정 '2000년 이후 최고의 영화 100선'
- 네이버 영화 평점 랭킹
- 각 플랫폼 별 인기 영화 및 장르별 자주 검색된 영화 목록

제목, 소개 글, URL, 이미지

284개

# Data

## 공연



- 네이버 공연 정보 크롤링
- 연극, 뮤지컬, 클래식 등 다양한 장르의 공연 정보 수집

공연이름, 장르, 소개 글, URL, 이미지

300개

# 4. Product Serving

# Product Serving

## Front-end



### Streamlit

front-end 경험 없이  
python 으로 web  
개발 가능.

## Back-end



비동기 요청 처리.  
pydantic, swagger 등  
으로 api testing 쉬움.  
Flask, Django에 비해  
빠른 속도

## Database



NoSQL - Not Only  
SQL.  
빠른 성능.  
Schemaless -  
스키마를 변경으로  
인한 문제 발생 X

## Server



Google Cloud Platform

Free - \$ 300

# Product Serving

## Frontend

- Streamlit, CSS, HTML을 사용하여 구현



- 일기 텍스트 입력 모듈
- 감정 선택 모듈
- 선택된 감정 기반 콘텐츠 추천 모듈
- 완성된 피드(일기+콘텐츠) 저장 모듈

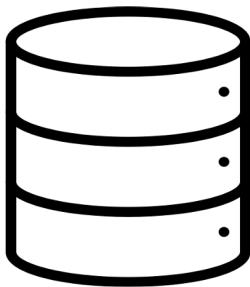


# Product Serving

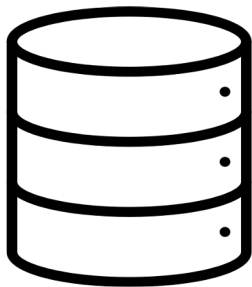
## Backend



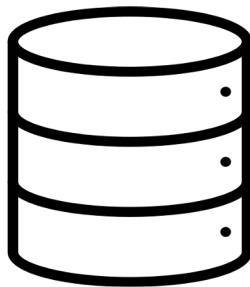
- FastAPI를 사용하여 구현



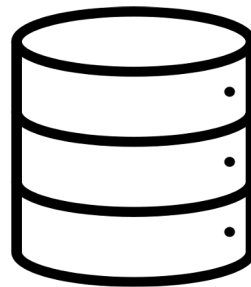
Diary



Contents



History



기타

# Product Serving

## Backend(Diary)

Input

사용자의 일기

Return

{“일기 입력 시간”, “작성한 일기 내용”, “일기에 담긴 감정들”}

POST

`/diary/input` Get Inference



# Product Serving

## Backend(Contents)

Input

일기에 담긴 감정들 or 선택한 감정들

Return

노래, 책, 영화, 연극

POST /contents/recommend Recommend Contents

GET /contents/songs View Playlist

POST /contents/songs/search Search Playlist

GET /contents/books View Booklist

POST /contents/books/search Search Booklist

GET /contents/movies View Movelist

POST /contents/movies/search Search Movelist

GET /contents/plays View Theaterlist

POST /contents/plays/search Search Theaterlist

# Product Serving

## Backend(History)

API  
일기

API  
문화 콘텐츠

GET /history/diary View Diary History Database

GET /history/selection View Selection History Database

POST /history/diary/insert Insert Diary Record

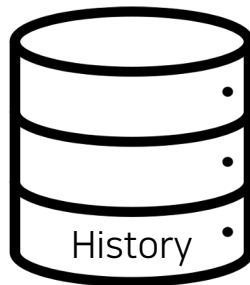
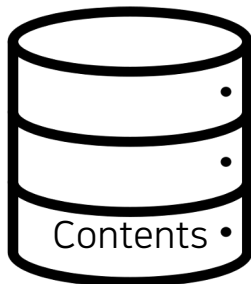
POST /history/selection/insert Insert Selection Record

# Product Serving

## Database

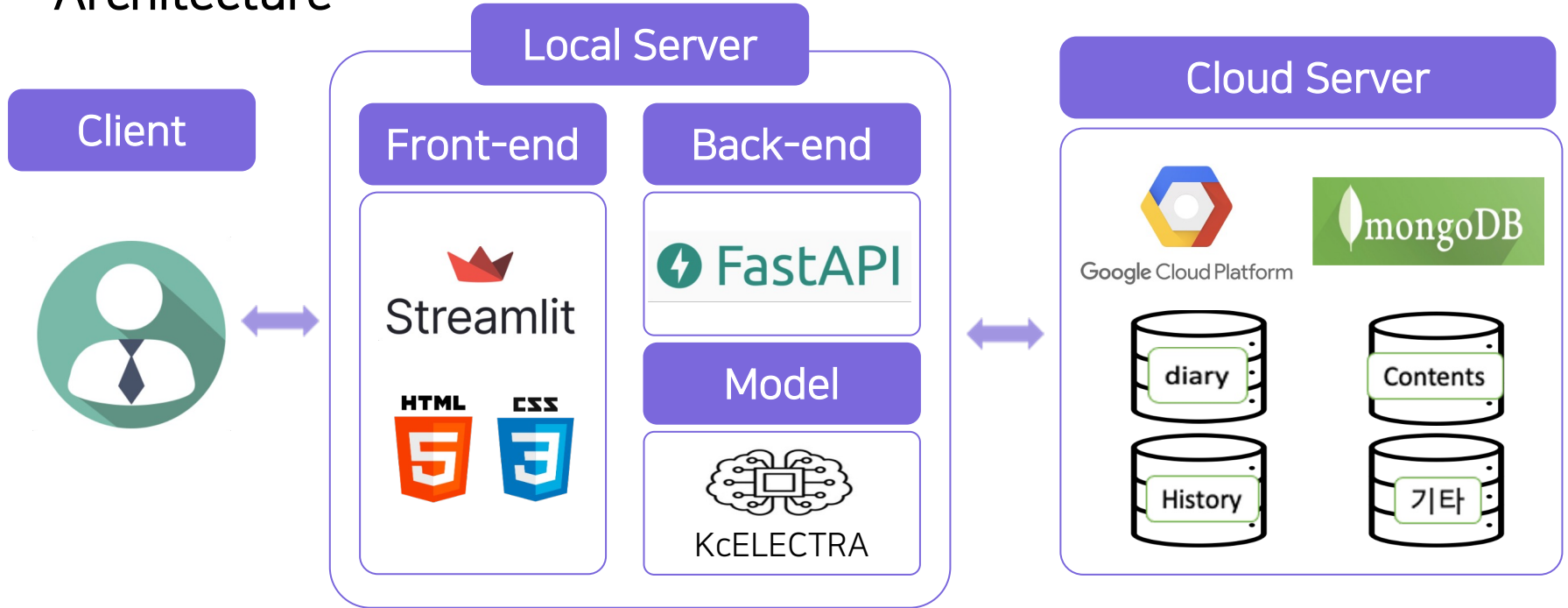


- GCP Cloud 를 사용하여 Local Server 와 분리
- MongoDB 를 사용
  - RDB 에 비해 빠른 성능
  - 스키마 변경에 따른 문제 발생 X



# Product Serving

## Architecture



# 5. Result & Conclusion

# 시연 영상

## 하루의 마침표.

"당신의 하루를 들려주세요. 오늘을 담은 선물을 드릴게요."

일기 쓰러가기

길을 걸었다. 따뜻한 햇살과 푸르른 잎사귀. 벌써 여름이 다가오고 있음이 느껴진다. 행복한 주말의 첫날. 오늘을 어떻게 보내야 할지 알고 할 수 있을까? 소소하게 산책하는 것만으로도 행복을 느낀다. 그렇다..! 어땠던..★

#기쁨 #행복 #즐거움/산남

2022-06-05 12:43:51

당신의 밤을 장식할 노래 한 곡



**환여름밤의 꿈**

San E, 레이나 (Raina)

무더운 밤 잠은 오지 않고 이런저런 생각에 불려본 너 나를 줄 물렸어 간지러운 바람 웃고

...

노래

## 하루의 마침표.

"당신의 하루를 들려주세요. 오늘을 담은 선물을 드릴게요."

일기

Please don't share it publicly

Make sure to share it



## 6. Q&A