

# 프로젝트: 모두봇

## 〈내가해〉

팀장: 추영재

팀원: 손호진, 정소민, 최하송, 김순호  
2025.10.28

## [ 목 차 ]

1. 문제 정의와 목표: 페르소나
2. 모두봇 개요
3. 서비스 소개
4. 코드 이슈 & 해결 방안
5. 프로토타입 체험
6. 기대효과 및 추후 개발 제언

DISCORD

많은 문서

고용24

출결  
관리

LMS

훈련 장려금

면접

예비군

행정  
처리

NOTION

ZEP

단위 기간

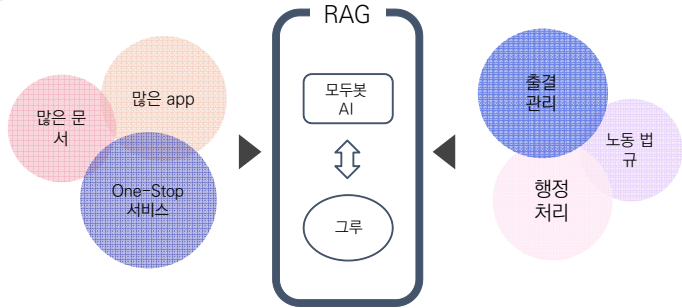
아놔,,,망할!!!!  
아직.. 잘 모르는 사람들.. Meta환경에 익숙하지 않아.. 뭐가 이리...휴...

내가 원하는 시간과 장소에서 필요한 정보를 몇 줄로 간략하게 정리해서  
얻을 수는 없을까?



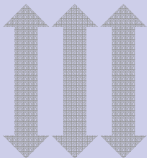
# 모두봇 개요

# 1 개요



[현재]

매니저



그룹



[RAG]

매니저



모두봇

그룹

## 2

## UI 구성과 기능배치

## 모두의 연구소

## 날짜 선택

원하는 날짜를 선택하세요:

2025/10/28

오늘은: 2025-10-28입니다.

## 관련 사이트

모두의 연구소 홈페이지

대차 5기 노선 워크스페이스

대차 5기 ZEP

LMS 홈페이지

## 첨부파일

휴가신청서 다운로드



안녕하세요! 저는 모두봇입니다.  
즐거운 모두연 생활을 위한 정보를 제공합니다. 😊

무엇이든 물어보세요!

질문을 입력해주세요. >

04



# 서비스 소개

## 1 개발 환경 및 기술 스택



## 2

## 팀원 협업:github

The screenshot shows a GitHub repository page for a user named 'anhyzn' with the repository name '-D55-LangchainThon'. The repository is public and has 4 branches and 0 tags. The main branch is 'main'. The repository contains several files and folders, including 'notebooks', 'visualization', '.gitignore', 'README.md', 'poetry.lock', 'pyproject.toml', and 'requirements.txt'. The README.md file is selected and shows the following content:

**[D55]LangchainThon**

"내가해" 팀의 모두봇 프로젝트입니다!

**Introduce**

모두의 연구소 행정업무를 간소화하여 운영하시러님과 그루님을 모두 편하게 사용할 수 있는 모두 "모두봇"을 구성합니다.

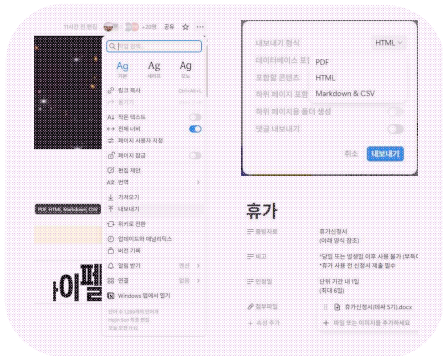
The right sidebar shows the repository's statistics: 0 stars, 0 watching, and 0 forks. It also includes sections for Releases, Packages, and Contributors.

### 3 데이터 개요



## 4

## 데이터 수집



## 5 데이터 전처리 (csv)

1	이름	전산신청(운매체크)	Files & media	Text	공결여부(그루체크)	날짜	부재시간	상태	서류제출
187									
188	(예시) *** 지각	No			Yes	09/04/2023	2시~4시	출결	No
189	*** 조퇴	No			No	03/19/2024		조퇴	No
190	*** 지각	No			No	04/08/2024		출결	No
191	***휴가	No			Yes	02/26/2024		휴가	No
192	***휴가	No			Yes	01/22/2024		휴가	No
193	***결석	No			Yes	02/02/2024		출결	No
194	☎***.공결 △	No			Yes	12/22/2023	종일	공휴일	No
195	***결석	No			No	11/08/2023		결석	No
196	***휴가	No			Yes	12/22/2023		휴가	No
197	재량휴일	No			No	10/10/2025		휴가	No
198	***지각	No			No	09/07/2023	10시~1시	지각	No
199	***결석	No			No	11/13/2023			No
200	***조퇴	No			No	03/08/2024	15:00 ~ 18:30	조퇴	No
201	***	No			Yes	01/21/2024			No
202	***결석	No			No	12/19/2023		결석	No
203	***결석	No			Yes	01/10/2024			No
204	***조퇴	No			No	09/22/2023	16시 30분 이후	조퇴	No
205	공휴일	No			No	12/25/2023		출결	No
206	추석	No			No	10/06/2025 → 10/08/2025		공휴일	No
207	여름방학	No			No	08/04/2025 → 08/08/2025		휴가	No

46	***		2025-09-15 17시~	조퇴
47	***	개인사유	2025-09-16 15:00-18:00	조퇴
48	***	개인사유	2025-09-24	휴가
49	***	개인사유	2025-09-29	휴가
50	***	개인사유	2025-09-29 15:05 ~ 18:00	조퇴
51	***		2025-09-30	휴가
52	***	예비군/민방위	2025-09-30	공결
53	***	자동차점검	2025-09-30	지각
54	***	개인사유	2025-10-01 ~ 13:00	지각
55	***	개인사유	2025-10-01 14:42 ~ 18:00	조퇴
56	***	개인사유	2025-10-02 11:00 ~ 15:00	외출
57	***	병원진료	2025-10-02	결석
58	개천절		2025-10-03	휴일
59	추석		2025-10-06	휴일
60	추석		2025-10-07	휴일
61	추석		2025-10-08	휴일
62	한글날		2025-10-09	휴일
63	재향휴일		2025-10-10	휴일
64	***	병원진료	2025-10-13 10:00~18:00	공결
65	***		2025-10-13	휴가
66	***	개인사유	2025-10-14 14:30 ~	조퇴
67	***	개인사유	2025-10-15 15:00 ~	조퇴
68	***	병원진료	2025-10-17 15:00-18:00	조퇴
69	***		2025-10-17	조퇴
70	***	개인사유	2025-10-17 14:30 ~	조퇴
71	***		2025-10-20 2:00~	지각
72	***	개인사유	2025-10-21 10:00 ~ 18:00	휴가
73	***		2025-10-22	조퇴
74	***		2025-10-24 10:00 ~ 18:00	공결
75	***		2025-10-27 14~16	외출
---				---

python

## 요약 ##

def create\_grouped\_documents(csv\_path):

# CSV 파일 로드

df = pd.read\_csv(csv\_path, encoding='cp949')

df = df[['이름', '사유', '날짜', '부재시간', '상태']].fillna()

# 이름 기준 그룹화

grouped = df.groupby('이름')

# 학생별 출결 기록 → 하나의 Document로 변환

documents = []

for name, group in grouped:

records = []

r['사유'] = {r['사유']}, 날짜: {r['날짜']},

상태: {r['상태']}, 부재시간: {r['부재시간']}

for \_, r in group.iterrows()

]

doc = Document({

page\_content=f'학생 이름: {name}\n\n' + "\n".join(records),

metadata={'학생이름': name, '총기록수': len(group)})

)

documents.append(doc)

return documents

## 7 데이터 전처리 (pdf)

1 160

SHARE VALUE, GROW TOGETHER



1 4

### [모두연 브랜딩 북 내용 정리]

#### 1. 핵심 슬로건

: '쉐어그루' - 쉐어 밸류 그로우 투게더(SHARE VALUE, GROW TOGETHER) 각  
영역 단어의 첫 한글 발음 모음

#### 2. '모두연'(모두의 연구소) 특성

이곳은 학교가 아니다.

그렇다고 대학원도 아니다.

우리가 알고 있는 보통의 연구소도 아니다.

그러나 이곳은 공부하는 곳이다.

함께 고민하여 연구하는 곳이다.

아울러 상생과 성장을 배우는 곳이다.

이곳은 경쟁하지 않는다.

그 대신 함께 협력하며 공통의 목표를 추구한다.

우리가 도달하고자 하는 곳은 더 좋은 세상이다.

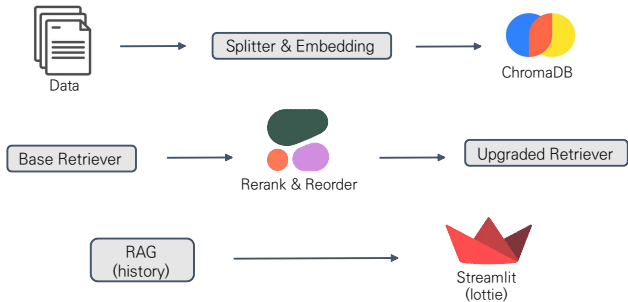
우리는 기존의 교목을 바꾸려 한다.

우리만의 방식으로 성장하고, 함께 행복해지고자 한다.

이 책은 그러한 모두의 노력을 담은 작은 기록이다.



## 8 RAG 구조



# 코드 이슈 & 해결방안

# 1 코드 이슈 & 해결방안

! 일정표에서 조퇴/휴가 등의 내역을 일부만 불러오는 문제 식별 !

	이름	사유	날짜	부재시간	상태
13	손호진	개인사유	2025-07-28	14:30 이후	조퇴
33	손호진	NaN	2025-09-05	14:30 이후	조퇴
44	손호진	NaN	2025-09-15	17시~	조퇴
65	손호진	개인사유	2025-10-15	15:00 ~	조퇴
67	손호진	NaN	2025-10-17	NaN	조퇴
71	손호진	NaN	2025-10-22	NaN	조퇴

```
1 answer = rag_chain.invoke('손호진님이 조퇴한 날짜를 알려줘')  
2 print(answer)
```

✓ 1.9s

손호진님이 조퇴한 날짜는 2025년 9월 5일, 2025년 9월 15일, 2025년 10월 15일, 2025년 10월 22일입니다.

## 2 코드 이슈 & 해결방안

[가정 1] 프롬프트를 구체화 하지 않아 발생했을 것이다.

### 목표

특정 인원의 모든 특이사항 (외출, 휴가, 조퇴, 지각 등) 날짜를 빠뜨리지 않고 정확하게 알림

### 지침

오직 주어진 데이터에서 찾을 수 있는 정보만 사용  
요청받은 인원의 모든 관련 내역 빠짐없이 제공

```
1 answer = rag_chain.invoke('손호진님이 조퇴한 날짜를 알려줘')  
2 print(answer)
```

✓ 10.4s

손호진님의 조퇴 내역은 다음과 같습니다:

1. [2025-07-28], [조퇴], [부재시간: 14:30 이후]
2. [2025-09-05], [조퇴], [부재시간: 14:30 이후]
3. [2025-09-08], [조퇴], []
4. [2025-09-15], [조퇴], [부재시간: 17시~]
5. [2025-09-22], [조퇴], []
6. [2025-10-02], [조퇴], [부재시간: 15:00 ~]
7. [2025-10-17], [조퇴], []
8. [2025-10-22], [조퇴], []

### 3 코드 이슈 & 해결방안

[가정 2] 임베딩 모델의 성능이 문제일 것이다.

#### 사용 모델

text-embedding-3-small

text-embedding-3-large

jhgan/ko-sroberta-multitask

[내용] 이름: 손호진 사유: 날짜: 2025-09-15 부재시간: 17시~ 상태: 조퇴...  
[유사도 점수] 0.8431

[내용] 이름: 손호진 사유: 날짜: 2025-09-05 부재시간: 14:30 이후 상태: 조퇴...  
[유사도 점수] 0.8443

[내용] 이름: 손호진 사유: 날짜: 2025-10-17 부재시간: 상태: 조퇴...  
[유사도 점수] 0.8669

[내용] 이름: 손호진 사유: 날짜: 2025-10-22 부재시간: 상태: 조퇴...  
[유사도 점수] 0.8723

[내용] 이름: 손호진 사유: 날짜: 2025-09-09 부재시간: 상태: 결석...  
[유사도 점수] 0.9287

[내용] 이름: 손호진 사유: 개인사유 날짜: 2025-10-15 부재시간: 15:00 ~ 상태: 조퇴...  
[유사도 점수] 0.9291

[내용] 이름: 손호진 사유: 개인사유 날짜: 2025-07-28 부재시간: 14:30 이후 상태: 조퇴...  
[유사도 점수] 0.9310

## 4 코드 이슈 & 해결방안

[가정 3] CSVLoader와 LLM의 작동 방식에 문제가 있을 것이다.

### CSVLoader

- 각 행을 하나의 chunk로 인식

### LLM

- 각 행을 독립적인 행으로 인식
- Lost in the Middle
- 추론 확장



### Solution

- 이름별로 Grouping
- 데이터 논리적으로 통합
- LLM의 인지적 부담 제거

## 5 코드 이슈 & 해결방안

[가정 3] CSVLoader와 LLM의 작동 방식에 문제가 있을 것이다.

	이름	사유	날짜	부재시간	상태
13	손호진	개인사유	2025-07-28	14:30 이후	조퇴
33	손호진	NaN	2025-09-05	14:30 이후	조퇴
44	손호진	NaN	2025-09-15	17시~	조퇴
65	손호진	개인사유	2025-10-15	15:00 ~	조퇴
67	손호진	NaN	2025-10-17	NaN	조퇴
71	손호진	NaN	2025-10-22	NaN	조퇴

```
1 answer = rag_chain.invoke('손호진님이 조퇴한 날짜를 알려줘')  
2 print(answer)
```

✓ 5.3s

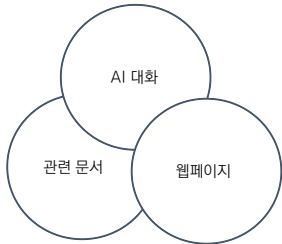
손호진님의 조퇴 내역은 다음과 같습니다:

1. [2025-07-28], [조퇴], [부재시간: 14:30 이후]
2. [2025-09-05], [조퇴], [부재시간: 14:30 이후]
3. [2025-09-15], [조퇴], [부재시간: 17시~]
4. [2025-10-15], [조퇴], [부재시간: 15:00 ~]
5. [2025-10-17], [조퇴], [부재시간: ]
6. [2025-10-22], [조퇴], [부재시간: ]

# 프로토타입 체험



## 1 UI 구성: 정보를 원하는 형태로 받아 보자



[ 정보의 취득 ]

- PDF, 웹문서, CSV 등 분산
- LMS, Zep, Notion 등 복잡

➡ 귀여운 AI (모두봇)

## 2 UI 구성과 기능배치

모두의 연구소

날짜 선택

원하는 날짜를 선택하세요:

2025/10/28

오늘은: 2025-10-28 입니다.

관련 사이트

모두의 연구소 홈페이지

대차 5기 노선 워크스페이스

대차 5기 ZEP

LMS 홈페이지

정부파일

휴가신청서 다운로드





안녕하세요! 저는 모두봇입니다.  
즐거운 모두연 생활을 위한 정보를 제공합니다. 🤖

무엇하든 물어보세요!

질문을 입력해주세요. >

### 3 시연 사례


 나는 데싸5기 손호진이야. 지금까지 내가 조퇴와 휴가를 언제 썼는지 궁금해.


 손호진 학생의 조퇴와 휴가는 다음과 같습니다:


- 조퇴: 2025-07-28, 2025-09-05, 2025-09-15, 2025-10-15, 2025-10-17, 2025-10-22
- 휴가: 2025-08-22, 2025-09-08

정보가 필요하시면 언제든지 말씀해 주세요!


 데이터 사이언스 5기 동료들 중 MBTI가 ISTJ 인 사람은 누구야?

 데이터 사이언스 5기 동료들 중 MBTI가 ISTJ인 사람은 추영재입니다.


 그 사람의 취미는 뭐야?

 추영재의 취미는 운동(풋살, 러닝, 헬스)과 독서입니다.

 예비군 훈련으로 출석을 못할 것 같은데 어떻게 해야해?

 부득이한 사유로 수업 참여가 어려운 경우, 운영진에게 사전에 알려야 합니다. 공가 신청은 발생일 +1일 까지 전산에 가능하며, 증빙 서류는 일주일 이내로 제출해야 합니다. 자세한 사항은 운영매니저님이나 파실님께 문의해주세요.

 단위기간 21일 중 14일만 출석했어. 나 훈련 장려금 받을 수 있을까?

 출석 일수가 80% 미만이므로 훈련장려금을 받을 수 없습니다.

 18일 출석했으면?

 단위 기간이 21일인 경우, 출석 일수 18일로 계산되어 훈련장려금을 받을 수 있습니다.

 얼마 받을 수 있어?

 훈련장려금은 1일 15,800원으로, 18일 출석 시 총 284,400원을 받을 수 있습니다. (15,800원 \* 18일)

# 기대효과 및 추후 개발 제언

## 1 기대효과

- 신입크루 상호간에 등록된 동료의 프로파일 등을 참조하여 상호 친근감 상승
- 개괄적 행정 절차를 실시간으로 이해를하고 세부사항은 대면 처리
- 교육진행 과정에서 필요하지만 익숙하지 않은 용어 개념 습득 (예, 쉼벨그투???)
- 시간과 장소 무관한 행정업무 내용을 쉽고 빠르게 인지

빠른 학습 환경 적응

## 2 추후 개발 제언

- 1. 현재 출석일을 고려하여 수려 기준 충족 일자 계산
- 1. 추가 문서와 웹페이지에 대한 접근
- 1. 교육 일정 반영
- 1. 오늘의 학습 내용 정리하여 제공