



Ai 기반 언어 독성제거 프로젝트

7팀 - 코드를 달리다

목록



1 프로젝트 소개

2 핵심 아이디어

3 상세 프로세스

4 Demo Page 테스트

5 슬랙 프로세스

6 슬랙 시연영상

프로젝트 소개

1

최근 공격적인 말투로 인한 갈등이 사회적 문제로 대두되고 있습니다.

2

이 문제를 해결하기 위해 저희는 AI를 활용한 '언어 독성 제거' 프로그램을 개발하여 공격적인 말투를 순화된 표현으로 바꾸어 갈등을 줄이고 긍정적인 소통을 촉진하고자 합니다.



핵심 아이디어

LLM을 이용한 언어
순화



설정된 프롬프트에
따라 언어순화

다수의 LLM
교차검증



AI 의 답변을
AI 들이 투표해서 선출

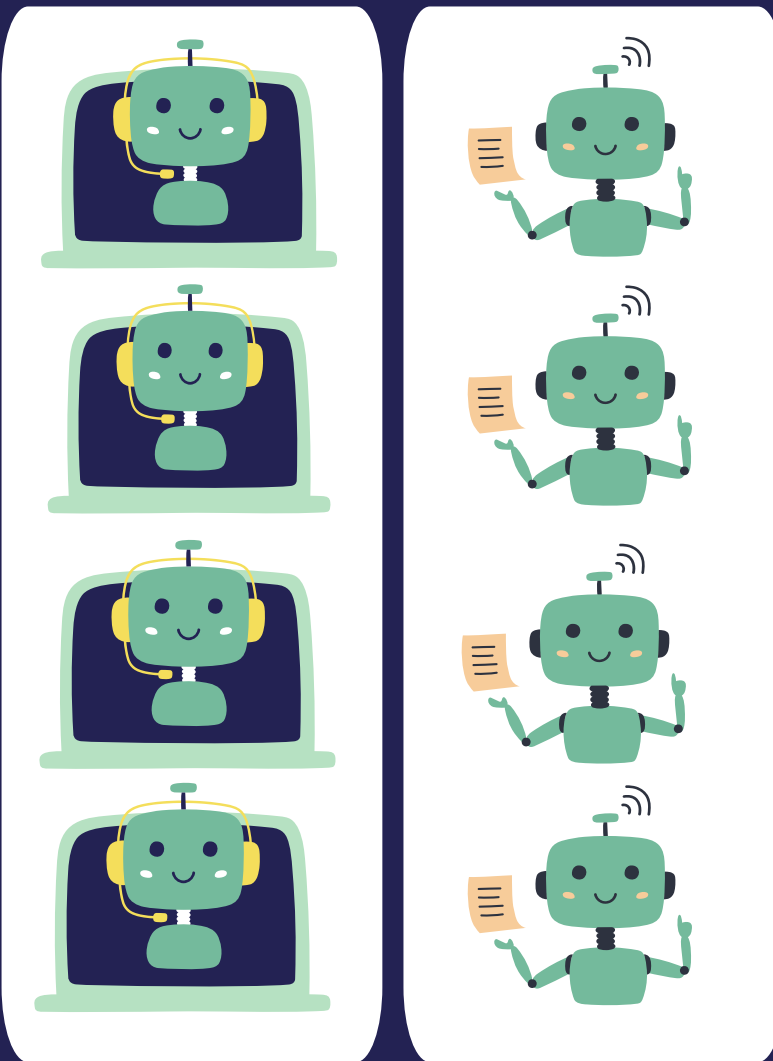
메신저 프로그램
적용 등 범용성



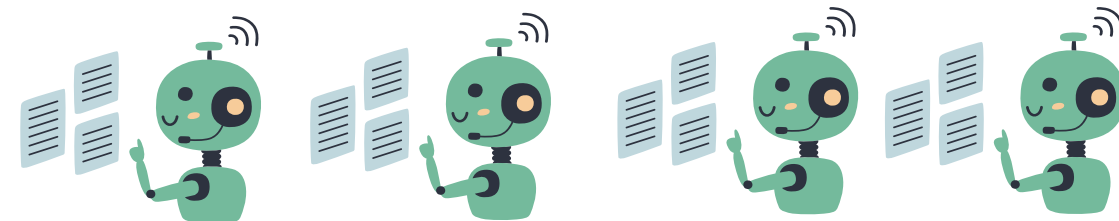
입력값 다른 채팅방으로
순화출력

상세 프로세스

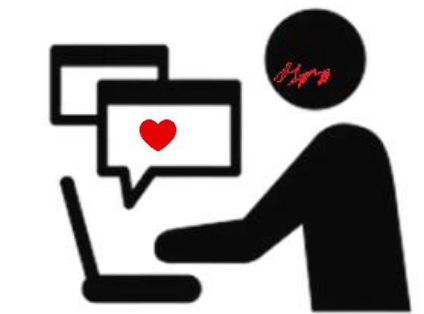
1.언어 전달,
각각의 LLM 이 답변



2.답변 전달,
LLM 투표결과 반환



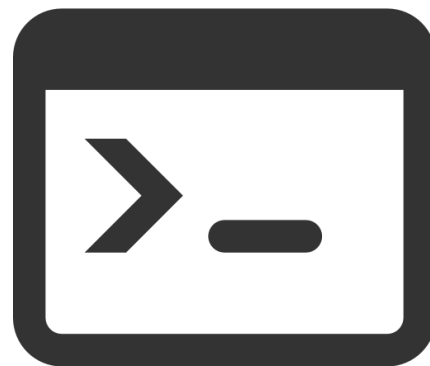
3. 순화언어 전달



대화상대 수신



사용자 입력



언어 독성 제거
프로그램

Demo Page 테스트



url :
<https://hanghae-psi.vercel.app/>

DEMO

Input

요즘 애들은 이런것도 못하나? 한심하구만.

Translate

Result

요즘 친구들은 이런 것도 어려워하는구나. 조금 더 익숙해지면 잘 할 수 있을 거야.

슬랙 프로세스

1. 사용자가
'언어 독성 제거 프로그램'
ChatBot 을 초대한 방에
대화를 입력한다.

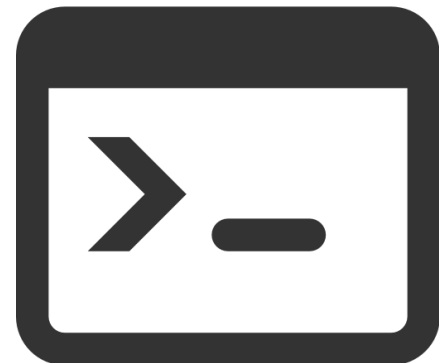


사용자 입력



2. ChatBot이 입력한 받은
내용을 프로그램에 전달한다.

언어 독성 제거
프로그램

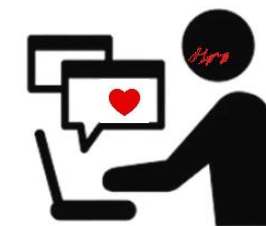
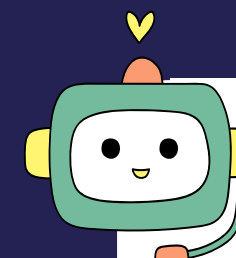


3. LLM이 언어를 순화한다.

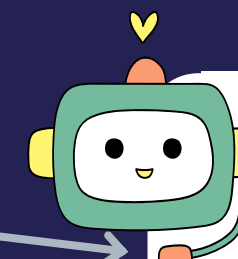


Gemini

4. ChatBot 이 대화상대에게
전달한다.

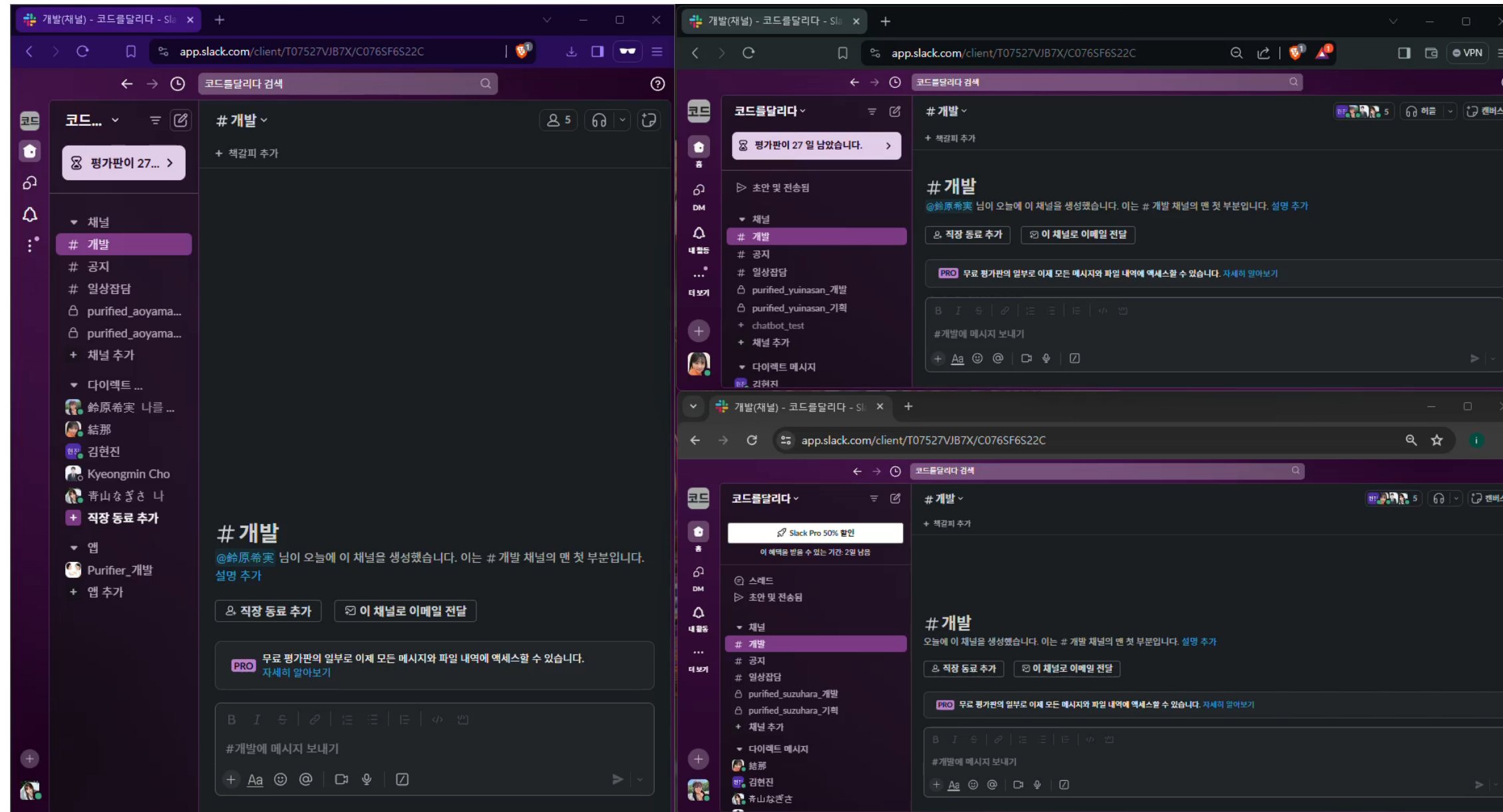


대화상대 수신

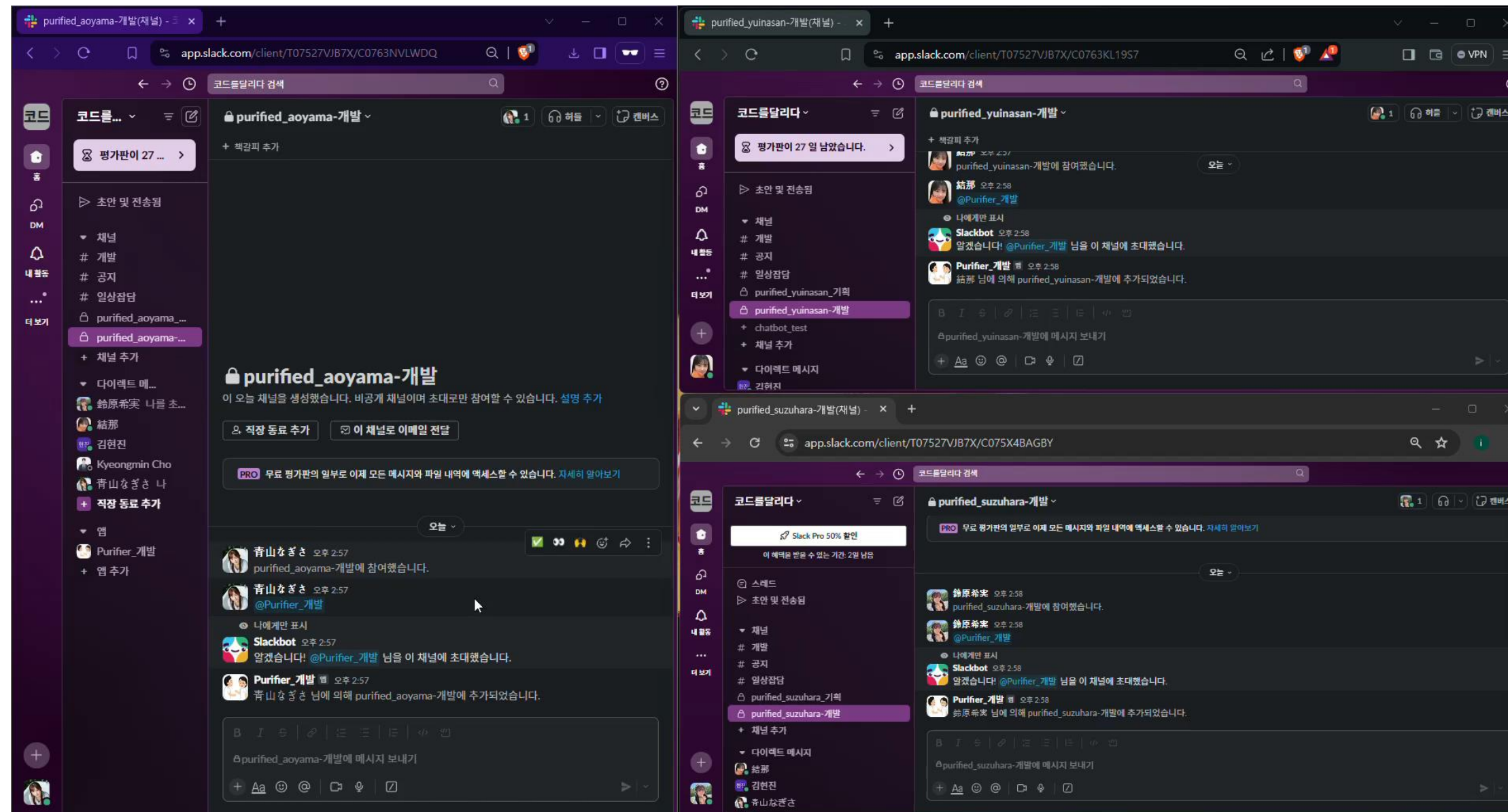


대화상대 수신

슬랙 시연영상 before



슬랙 시연영상 After





감사합니다

Thank you

7팀 코드를 달리다

