

휴튼 AI

in 휴튼 질문 패키지

AI 엔지니어 담당자 임세진

heuton

목차

1. 휴튼 질문 패키지에서 휴튼 AI의 역할 ----- 10분 소요
2. 휴튼 AI에 사용된 기술 소개 ----- 20분 소요
3. 휴튼 AI가 앞으로 해결해야 할 문제들 ----- 5분 소요
4. 휴튼 AI에 대해 궁금한 점 ----- 15분 소요

휴튼 질문 패키지에서 휴튼 AI의 역할

The screenshot shows the Hutton Question Package interface. At the top, it displays '세진 질문 패키지' (Sejin Question Package) with a creation date of '2023.08.23'. Below this, there's a summary section with '나에게 집중한 시간' (Focused time on me) at '02:03:15' and '글자 수' (Number of characters) at '10399'. To the right, there's a '추가 메뉴' (Additional menu) with options like '답하지 않은 질문 (0)', '다른 생각 엿보기', and '휴키피디아'. A red box highlights the '이번 주의 질문이에요 (0)' (This week's question is 0) section, which shows a success message: '이번 주의 질문 모두 성공! 🎉' (All this week's questions were successful!). Below this, another red box highlights the '세진의 답변이에요 (70)' (Sejin's answer is 70) section, which contains a question from '임세진님은 AI 엔지니어로서의 공부를 즐기나요?' (Do you enjoy studying as an AI engineer?) and a detailed response about AI engineering. A button in this section says '답변 70개 모두 보러 가기' (View all 70 answers). A red arrow points from the '답변 70개 모두 보러 가기' button to a callout box on the right.

일주일에 한번 씩
3~5개의 질문 생성

후속 질문 생성 가능!

휴튼 질문 패키지에서 휴튼 AI의 역할

사람들이 왜 하필 휴튼 질문 패키지를 쓸까?

사람의 추가 질문 vs. 휴튼 AI의 추가 질문

- 휴튼 AI의 경우, 질문 검수 과정만 더 면밀하게 손 보면 확실한 익명화가 가능하다.
- 사람이 모든 사용자의 질문을 퀄리티 높게 생성하는 데에는 물리적인 한계가 존재한다.
- 질문을 생성하는 사람이 모르는 내용의 답변일 경우에 추가 질문을 생성하기 어려워진다.

휴튼-AI의 존재 의의

휴튼 질문 패키지에서 휴튼 AI의 역할

그렇다면, 질문을 어떻게 생성해야 하는가?

- 최대한 간결하고 알아 보기 쉽게
- 담백하고 의미가 명확하게 전달 되게
- 정중한 듯 하지만 너무 딱딱하지는 않게 (해요 체)
- 번역체가 아닌 최대한 자연스러운 대화체로
- 서술형이 아닌, 질문의 형태로
- 사용자의 답변에서 **Heuton이 의도하는 바에** 따라

휴튼 질문 패키지에서 휴튼 AI의 역할

나쁜 예시

1. 앞으로 도전해보고 싶은 일에 대해 어떤 것들이 있는지 파악하고 싶습니다.
2. 어떤 상황에서 불행한 기분을 느끼는지 파악하기 위해 어떤 접근 방식을 사용하시나요?
3. 실생활에서 어떤 일을 할 때 소소하게 행복한지 파악하기

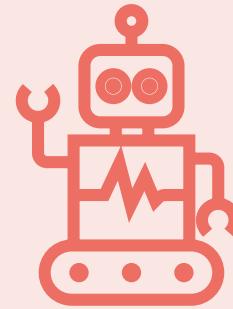
좋은 예시

1. 앞으로 도전해보고 싶은 일로 어떤 것이 있나요?
2. 불행한 기분을 파악하기 위해 어떤 방법을 이용하나요?
3. 실생활에서 어떤 상황에서 소소하게 행복한가요?

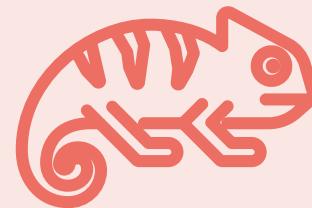
휴튼 AI에서 사용된 기술 소개



Q&A Network



OpenAI
(GPT API)



Polymorph
(Fine-tuning)

Q&A Network



**Q&A Network
(User A)**



**Q&A Network
(User B)**

**Q&A Network
(User C)**



Rule

답변들 간의 유사도가 특정 수치를 넘어가면 Link 연결

활용 방안

- 질문 생성할 때 참고 용 이전 질문과 답변을 추출할 때 이용됨
- 질문 추천 시스템에 이용 가능

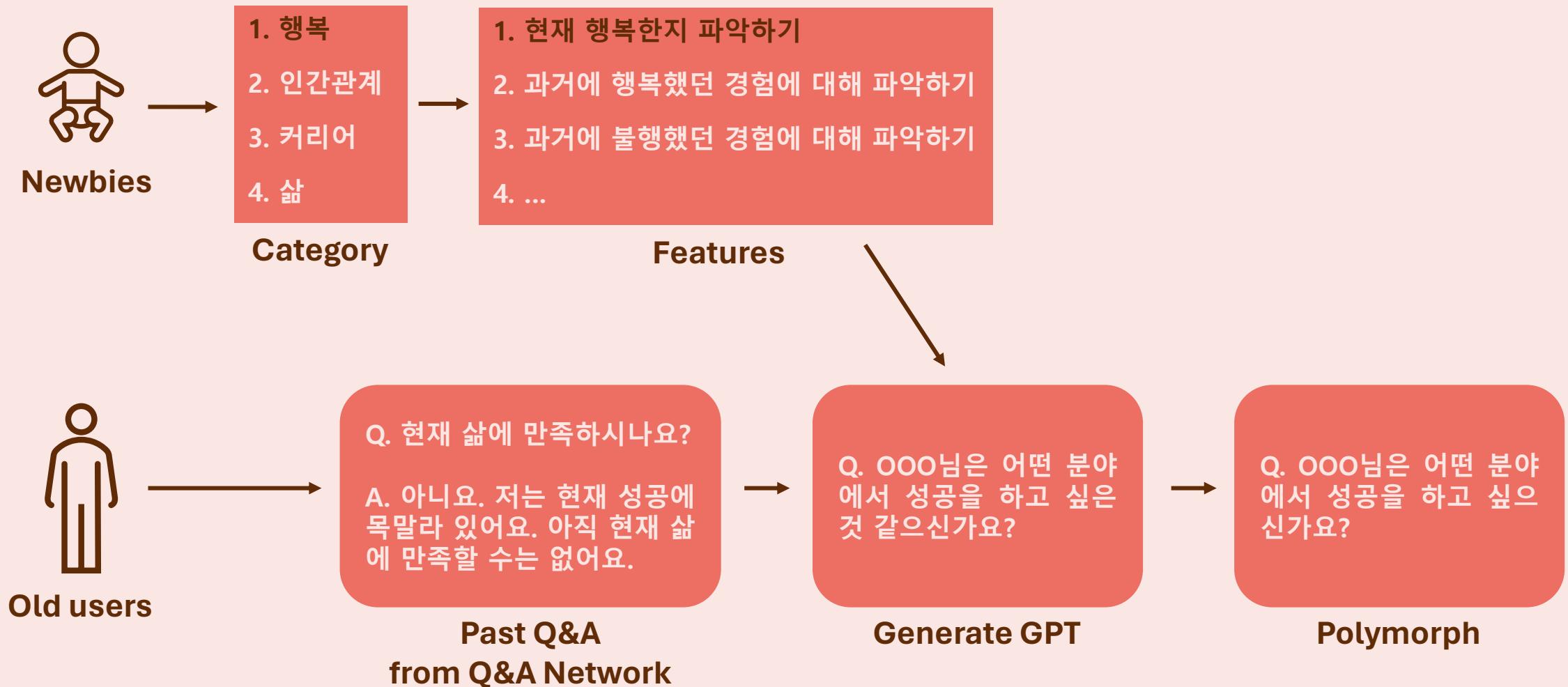
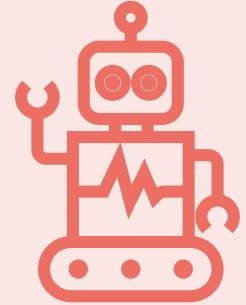
Q&A Network



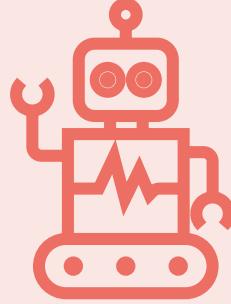
유사도 산정 방식

- Transformers Model 기반 Text Cosine Similarity 계산 (sentence-transformers Python 패키지 사용)
- 현재는 Cosine Similarity 값이 0.5 이상이면 연결선을 연결 (question id 기반으로 연결 – many to many Field 이용)
- 하루마다 질문 생성 이전에 Q&A Network를 초기화 하고, 위와 같은 방식으로 재구축 함.

OpenAI (GPT API) – 질문 생성 로직



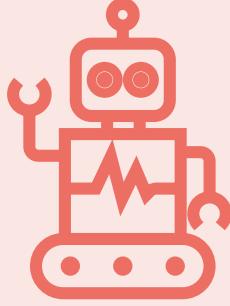
OpenAI (GPT API) – Prompt 적용방식



OpenAI Messages 방식과 Function Call 방식

- Messages 방식으로 role을 나누어 prompt 전달 (few shot learning 용이해짐) – user, assistant, system
- Function Call 기능으로 출력물을 json 형태로 변환하여 출력 → dictionary로 변환하여 처리
- Function Call에서 추가 질문과 부연 설명에 대한 description에 질문 생성을 명시함으로, 질문 생성을 보다 원활하게 하도록 함

OpenAI (GPT API) – Q-Type



선택	ID	질문	부연 설명	적합도
<input type="checkbox"/>	417	이사야 59장의 내용을 읽으시면서 구체적으로 어떤 부분이 눈에 띄셨나요? 어떤 동기를 느끼게 되었나요?	이전에 말해주신 것이 있나요?	35
<input type="checkbox"/>	419	이전에 말해주신 죄악이 관계를 갈라놓는다는 사실이 왜 중요한가요?	이 질문은 내가 죄악의 심각성을 인식하고 회개하고 돌아보는 동기를 부여하는 것입니다.	41
<input checked="" type="checkbox"/>	421	이전에 말해주신 것이 있나요? 그 구절이 왜 기억에 남았는지 좀 더 자세히 들려주실 수 있나요?	이전에 말해주신 것처럼, 이사야 59장 1절이 세진님에게 어떻게 관계를 갈라놓는다는 것이었나요? 이 구절이 세진님에게 왜 기억에 남았나요? 어떤 생각과 감정을 불러일으켰나요?	41
<input type="checkbox"/>	423	이사야 59장 1절-2절을 읽으면 어떤 동기를 느끼셨나요? 어떤 생각이 들었나요?	세진님은 이사야 59장 1절-2절이 여호와의 구원의 한계와 이유를 설명하며, 죄악이 관계를 갈라놓는다는 사실을 강조한다고 생각합니다. 이에 대해 세진님의 동기와 생각을 자세히 알고 싶어요.	38

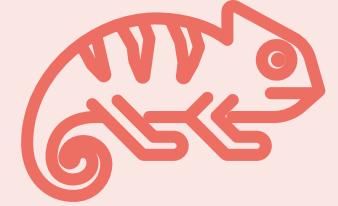
참고한 과거 답변 하나 당
6개의 Q-Type을 물어보는 질문 생성

이전 답변과 새로 생성된 질문 간의
유사도를 계산하여 적합도 산정

Question Type

1. 과거에 있었던 비슷한 경험을 물어보기 위해
2. 중요 키워드를 고르고, 그 중요 키워드가 사용자에게 왜 중요한지 물어보기 위해
3. 이유나 계기를 물어보기 위해
4. 향후 실행 계획을 물어보기 위해
5. 구체적인 사례 또는 부연 설명을 파악하기 위해
6. 핵심 사안을 다른 관점에서 바라보도록 하기 위해

Polymorph (Fine-tuning)



1	HeutonAI	student
2	업무 강도와 연봉을 어떤 기준으로 판단하시나요?	여러분에게 있어서 업무 강도와 연봉은 어떤 기준으로 판단하시나요?
3	회사에서 업무 강도가 센 것은 많은 일 처리를 해야 한다고 볼 수 회사에서 일하는 동안 업무 강도가 세다면 많은 일을 처리해야 하는 것으로 볼 수 있습니다.	
4	앞으로 도전해보고 싶은 일로 어떤 것이 있나요?	앞으로 도전해보고 싶은 일에 대해 어떤 것들이 있는지 파악하고 싶습니다.
5	황혜성님의 개인적인 관심사와 열정을 알아보는 데 도움이 될 것 황혜성님은 현재의 상황에서 미래를 내다보며 도전하고 싶은 분이나 프로젝트에 대해 생각해보시는 것 같아요.	
6	어떤 상황에서 불행하다고 느끼시나요?	불행한 상황이 어떤 것인지 파악할 수 있는 방법은 무엇인가요?
7	평소에 불행하다고 느꼈던 적이 있나요? 어떤 상황에서 불행하다고 느끼는 상황이 있었던 적이 있으신가요? 불행하다고 느끼는 상황은 어떤 것인가요?	
8	불행한 기분을 파악하기 위해 어떤 방법을 이용하나요?	어떤 상황에서 불행한 기분을 느끼는지 파악하기 위해 어떤 접근 방식을 사용하시나요?
9	신관용님이 불행한 상황을 파악하기 위해 어떤 방법을 사용하시나요?	신관용 님은 자신의 감정에 대해 솔직하고 정직하게 이야기하는 경향이 있기 때문에, 불행한 상황을 파악하는 데 있어 신관용 님의 경험과 접근 방식을 참고하는 경우 좋을 것 같습니다.
10	어떤 종류의 도전이나 목표가 있을 때 특별히 즐거운가요?	어떤 종류의 도전이나 목표를 가지고 있을 때 특히 즐거움을 느끼는지 알려주세요.
11	이전에 말해주신 것처럼 도전적인 목표를 갖고 성공하거나 성장하신 것처럼 도전적인 목표를 가지고 성공하거나 성장한다는 것에 흥분하고 즐거워하는 경우 많았던 것 같아요.	
12	실생활에서 어떤 상황에서 소소하게 행복한가요?	실생활에서 어떤 일을 할 때 소소하게 행복한지 파악하기
13	때로는 큰 일보다 소소한 순간들이 행복을 주기도 해요. 일상 생활때로는 큰 일보다는 소소한 순간들이 행복을 주기도 합니다. 일상의 일들 중 어떤 일들이 행복한 순간인가요?	
14	예전에 불행하셨던 적이 있나요?	과거에 불행했던 경험에 대해 파악하기
15	김보영님이 겪었던 불행한 경험에 대해 알아보자 해요. 과거에 김보영님에게서 과거에 불행했던 경험들을 알아보자 합니다. 과거에 불행했던 경험들이 어떤 것인가요?	
16	일상 생활에서 어떤 일을 할 때 소소한 행복을 느끼나요?	일상 생활에서 어떤 일을 할 때 소소하게 행복한지 생각해보세요?
17	많은 사람들이 큰 성취를 이룰 때에만 행복을 느끼는 것이 아니지만 많은 사람들이 큰 성취를 이룰 때에만 행복을 느끼는 것은 아닙니다. 소소한 순간들에서도 행복을 느끼는 경우 많았던 것 같아요.	
18	힘들어하는 사람들을 도울 때 어떤 방식을 선호하나요? 상담이나 힘들어하는 사람들을 도울 때 어떤 방식을 선호하시나요? 상담이나 조언을 통해 문제를 해결하는 데 있어 어떤 방식이 더 효과적인가요?	
19	김민형님은 사회적으로 영감을 줄 수 있는 사람이 되어 사람들에게 김민형님은 사회적인 영감을 줄 수 있는 사람이 되어 사람들에게 도움을 전달하고 싶다고 말씀하셨던 것 같아요.	
20	본인이 좋아하는 일과 잘하는 일 중에 어떤 일이 성공적일 것 같아요? 자신이 좋아하는 일과 잘하는 일 중에서 어떤 일이 직업으로서 성공적일 것 같으신가요?	
21	어떤 일에 흥미를 갖고 즐거움을 느끼나요? 본인이 갖고 있는 재·지금까지의 경험과 관심사를 토대로 직업 선택에 대해 고민해보기 위해서는 첫 번째로 자신의 재능과 관심사를 살펴보는 게 좋을 것 같아요.	
22	왜 유기견보호소에서 일하는 것을 좋아하시나요?	유기견보호소에서 일하는 것에 대해 좋아하는 이유가 무엇인가요?

매 질문 생성 시마다 GPT generated 데이터(student)와 Polymorph 된 데이터(HeutonAI) 구축



잘못된 변환 내용이 있을 경우에 원하는 질문 내용으로 수정 (라벨링)



Hugging Face에서 개발한 python transformers 패키지를 이용하여 앞에서 구축한 데이터로 fine-tuning 한 모델을 Polymorph에 재 적용



휴튼 AI가 앞으로 해결해야 할 문제들

- 질문 생성 고도화
 - OpenAI API 관련된 신기술 숙지
 - 지속적으로 업데이트 되는 GPT API follow up
 - 신기술 공부 (논문, 아티클 베이스)
- 말투 변환 고도화
 - Fine-tuning을 위한 데이터 구축 중 (2023-11-10 기준 약 2,000개 데이터 구축되어 있음)
 - 구축된 데이터에서 잘못 Polymorph 된 데이터 수정 (수정하는 비율 산정 필요: Polymorph 정확도)
- 다른 생각 엿보기 기능 고도화
 - 무작위 엿보기가 아닌, 나의 답변과 유사한 다른 유저의 질문 답변 보기 (추천 기능)
 - 마음에 드는 엿보기 질문을 “도둑질”하기