

Medical Consultation Assistant

Microsoft AI School

# Medly

Team 3



Listen Back  
Heal Better

# Contents

01 프로젝트 개요

02 Market Research

03 Product Design

04 AI 모델 개발

05 APP 개발

06 인공지능 윤리

07 Outro : 자체 평가

08 현장 시연

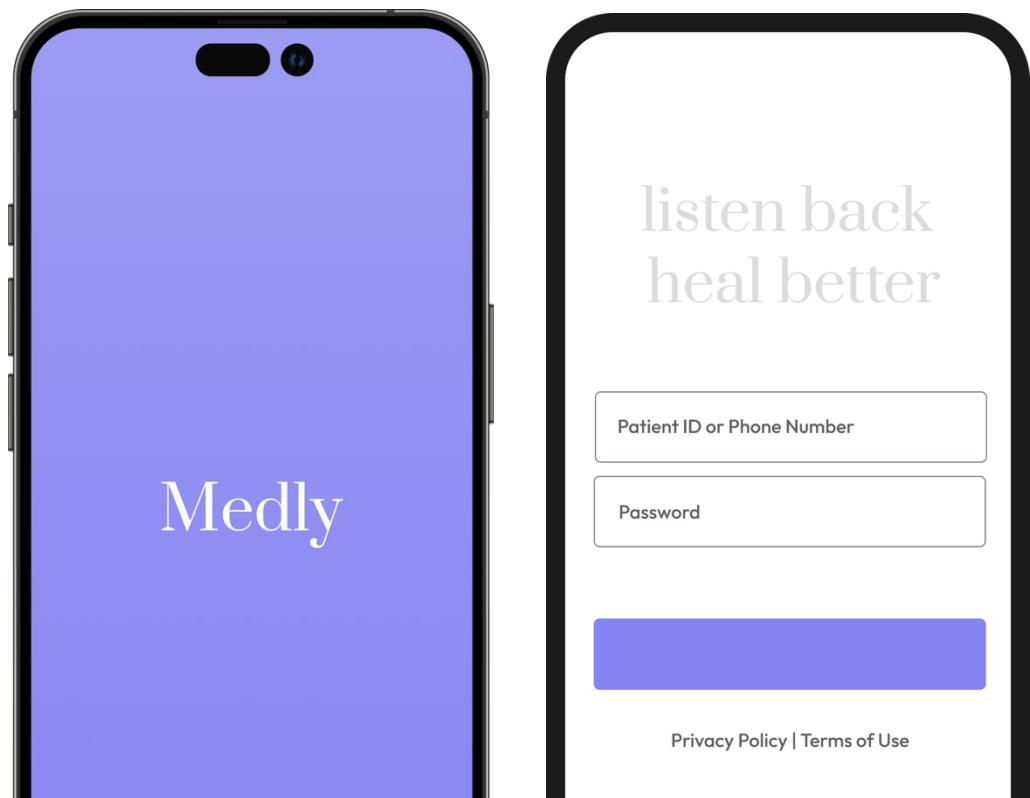
## 01 프로젝트 개요

서비스 Medly

"수술 전 환자와 의료진 모두에게 정확하고 신뢰성 있는 정보를 제공함으로써 의사소통의 질을 향상시키는 온인원 서비스"

# Medly?

**Medical + Melody**



의료(Medical)와 멜로디(Melody)를 결합한 단어로,  
의료 정보를 듣는 기능(녹음 및 재생)을 강조함

### 1. 수술 동의서 상담 관리 및 보조

- 환자 목록 조회
- 의료 동의서 열람
- 구두 설명 녹음 및 재생과 요약 기능

### 2. 의료 정보 검색 및 의료 지식 챗봇

- 문헌 검색
- PatientCareGPT
- MedSurgGPT
- MedBioGPT

## 01 프로젝트 개요

시연 영상

Medly

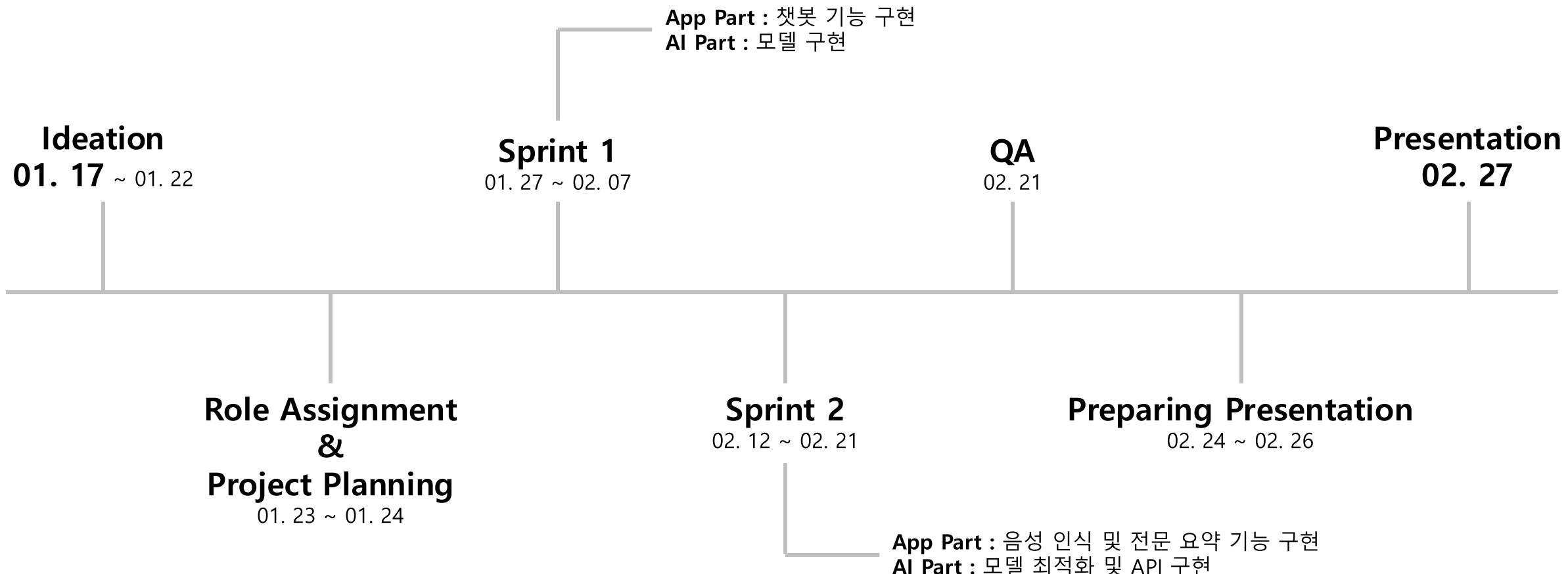
# 01 프로젝트 개요

## 팀 구성 및 역할

이름	역할	담당 업무
선우진성	팀장, 프론트엔드	클라이언트(RN) 개발, 구조 설계, 프로젝트 매니징, 서버(Nest) 개발 지원
이진형	백엔드	서버(Nest) 개발 서브, 인프라 구축
정지현	머신러닝	MedSurgGPT(수술 가이드라인 서칭 챗봇) 개발 및 API 구현
김재빈	머신러닝	논문 검색 Model, 환자용 FAQ Model 개발 및 API 구현
현수민	머신러닝	환자용 PatientCare Model(Azure OpenAI Fine-tuing, RAG) 개발 및 API 구현
이현영	머신러닝	서비스 기획, STT Model, Summarization Model 개발 및 API 구현
김선옥	머신러닝	서비스 기획, MedBioGPT 개발 (Azure OpenAI Fine-tuing, AI Search), API 구현
안의진	백엔드	서버(Nest) 개발 메인

# 01 프로젝트 개요

## 타임라인



## 02 Market Research

### 시장 내 이슈

이슈사항

- ① 의료소송과 의사의 설명의무  
(헬스조선 기사) 환자가 이해 못했다면 수술동의 받았어도 '유죄'  
(코메디닷컴 기사) 수술안내서 쭉 읽고 동의 받으면 설명의무

#### 기존 판례에서의 설명의무 기준

설명의무 위반으로 인한 배상 판결 사례

- 2010년 사례: 안압강하제(메타졸아마이드) 처방 후 희귀한 부작용 발생 사실 인정 x, 그러나 설명의무 위반으로 1,600만 원 배상
- 2019년 사례: 70대 환자가 아들에게 설명했다는 이유로 2,000만 원 배상 판결
- 2020년 사례: 12세 모야모야병 의심 환자 → 보호자에게만 설명했다는 이유로 2,000만 원 배상

#### 최근 대법원 판결: 설명의무의 시간적 요소 추가

대법원 판결 (2022년 1월 27일)

- 설명 후 충분한 숙고·결정 시간을 줘야 한다는 판례 확립
- 40분 만에 수술 진행 → 환자가 숙고할 시간이 부족했다고 판단
- 병원의 설명의무 위반 인정

현실적 한계

- 환자 본인 vs 보호자 설명 기준 불분명 (예: 미성년자, 고령 환자)
- 환자의 이해도 고려 부족:
  - ✓ 의료 지식이 부족한 환자에게 어떤 방식으로 설명해야 할지 모호
  - ✓ 현실적으로 모든 가족에게 개별 설명이 어려움
- 법적 소송 증가 → 방어진료 확산 및 고위험 시술 기피 증가

#### ② 의료 정보의 이해와 환자의 자기결정권

- (연합뉴스 기사) 의사소통 어려운 병원... '수술동의서 이해도 70점대'  
(메디컬옵저버 기사) 수술 전 의사 말 한마디의 효과

#### 수술동의서에 대한 환자 이해도

부산대 연구 결과 (2023년)

- 정형외과 병원에서 수술을 앞둔 환자 243명 대상 조사수술동의서 이해도: 100점 만점에 73.9점
- 의료진 설명 후에도 기억하는 의학 용어는 10개 중 6.24개
- 특히 50대 이상, 학력이 낮을수록 이해도가 더욱 저조

수술동의서 용어 인지율 예시

- '색전증' (혈전으로 혈관이 막히는 질환) 인지율 32.5%
- 수술동의서에 대한 설명만족도 평균 68.04점

연구 결론 분석

- 동의서 설명이 형식적 절차에 그치고 있음
- 단순 서명이 아니라 환자의 실질적 이해도 확인과 보충 설명 필요

• 의료 전문 용어를 완전히 배제하기 어려움

✓ 진료 시간이 짧아 충분한 설명을 듣기 어려운 경우가 많음

• 보호자가 원격으로 서명을 하더라도, 모든 보호자가 충분한 정보를 공유 받고 동의했는지 확인하기 어려움

• 전자 서명을 했다고 해도, 법적으로 충분한 보호를 받을 수 있을지 불확실

## 02 Market Research

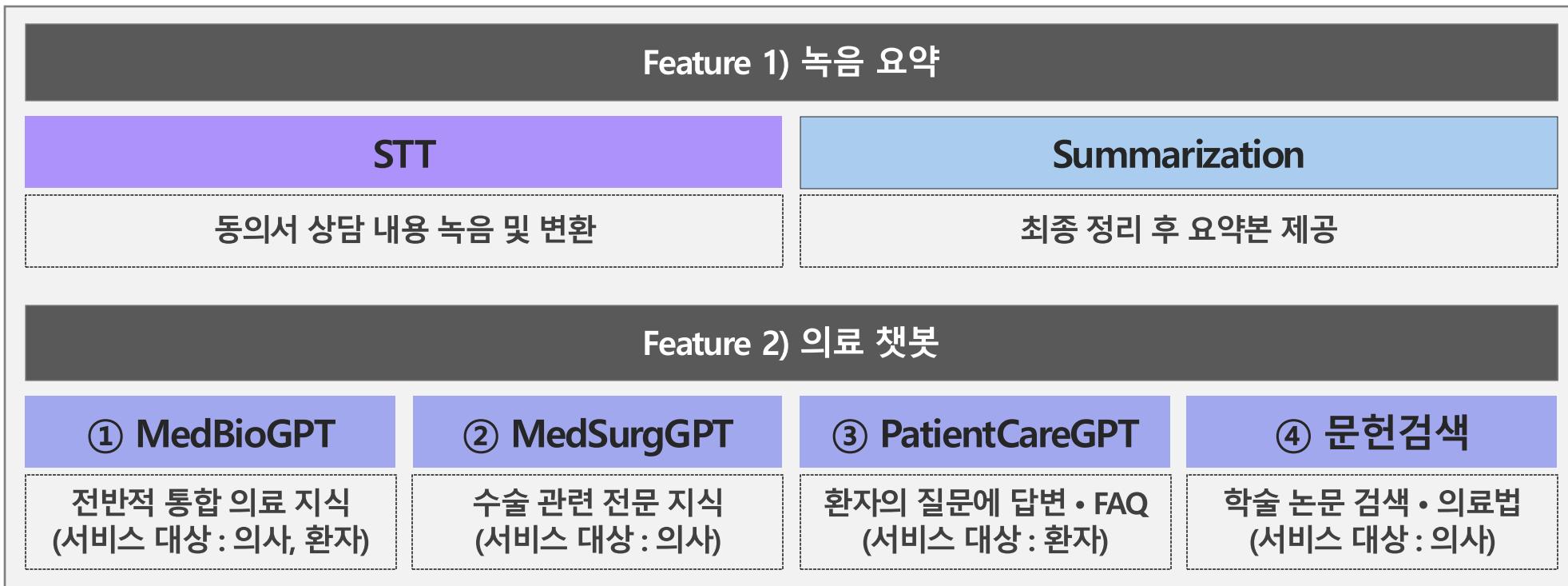
### 시장 니즈

의료 STT	의료진	Pain Point	시장 니즈	기대 효과
	환자			
의료 챗봇	의료진	<ul style="list-style-type: none"><li>• 환자가 반복적으로 문의하여 상담 시간이 과다</li><li>• 수술 동의서 오해로 인해 법적 분쟁 발생 가능</li></ul>	<ul style="list-style-type: none"><li>• STT 기반 수술 동의서 자동 기록<ul style="list-style-type: none"><li>✓ 수술 전 환자와 의사의 대화 과정을 녹음한 후, 텍스트로 변환하여 정확하고 객관적인 기록</li></ul></li></ul>	<ul style="list-style-type: none"><li>• 의사</li></ul> <ul style="list-style-type: none"><li>▶ 최신 연구 및 통계를 기반으로 한 의사결정 지원</li><li>▶ 문헌 탐색 시간 절약</li><li>▶ 반복적인 설명 부담 감소</li><li>▶ 정확한 기록 확보로 법적 리스크 감소</li><li>▶ 환자와의 신뢰 구축 및 원활한 소통</li></ul>
	환자	<ul style="list-style-type: none"><li>• 수술 전 불안 및 정보 부족으로 인한 환자 불만족, 의료 분쟁 증가</li><li>• 의료 용어가 난해하여 내용을 숙지하기 힘듦</li><li>• 보호자와 함께 다시 설명을 듣기 어려움</li><li>• 환자의 모든 질문을 개별 응대해야 함</li><li>• 반복적인 기본적인 질문들이 업무 부담을 가중</li><li>• 최신 의료 법, 논문을 실시간으로 반영하기 어려움</li><li>• 질문이 있음에도 시간적 제약으로 모든 궁금증을 해결하기 어려움</li><li>• 병원 방문 및 전화상담을 통해 질문해야 해서 불편함</li></ul>	<ul style="list-style-type: none"><li>• AI 기반 요약 모델 제공<ul style="list-style-type: none"><li>✓ 변환된 텍스트를 요약하여 전달 함으로써, 환자의 편의 향상</li></ul></li><li>• 의사용 최신 의학 문헌 탐색 기능<ul style="list-style-type: none"><li>✓ 환자와 상담시 최신 정보를 기반으로 자세한 통계 수치 등을 제공 가능</li></ul></li><li>• 환자용 의료정보 및 의료법 검색 기능<ul style="list-style-type: none"><li>✓ 시간과 장소에 구애받지 않고 환자의 궁금증을 해소하여 수술 전 의사결정에 도움 및 불안 해소</li></ul></li></ul>	<ul style="list-style-type: none"><li>• 환자</li></ul> <ul style="list-style-type: none"><li>▶ 어려운 의료 정보를 쉽게 이해</li><li>▶ 수술 및 치료 과정에 대한 불안감 감소</li><li>▶ 충분히 이해한 상태에서 의사결정 가능</li><li>▶ 궁금한 정보를 신속하게 얻을 수 있는 접근성 향상</li></ul>

기존의 단순한 전자 동의서 기록 시스템을 넘어 환자와 의료진 모두의 입장에서  
실질적인 가치를 제공하는 최초의 통합형 의료 커뮤니케이션 솔루션으로, 현재 시장에 존재하지 않는 독보적인 서비스.  
→ 이를 통해 의료 현장에서의 소통 문제를 근본적으로 해결하고, 환자 중심의 안전한 진료 환경을 조성하는 데 기여.

## 03 Product Design

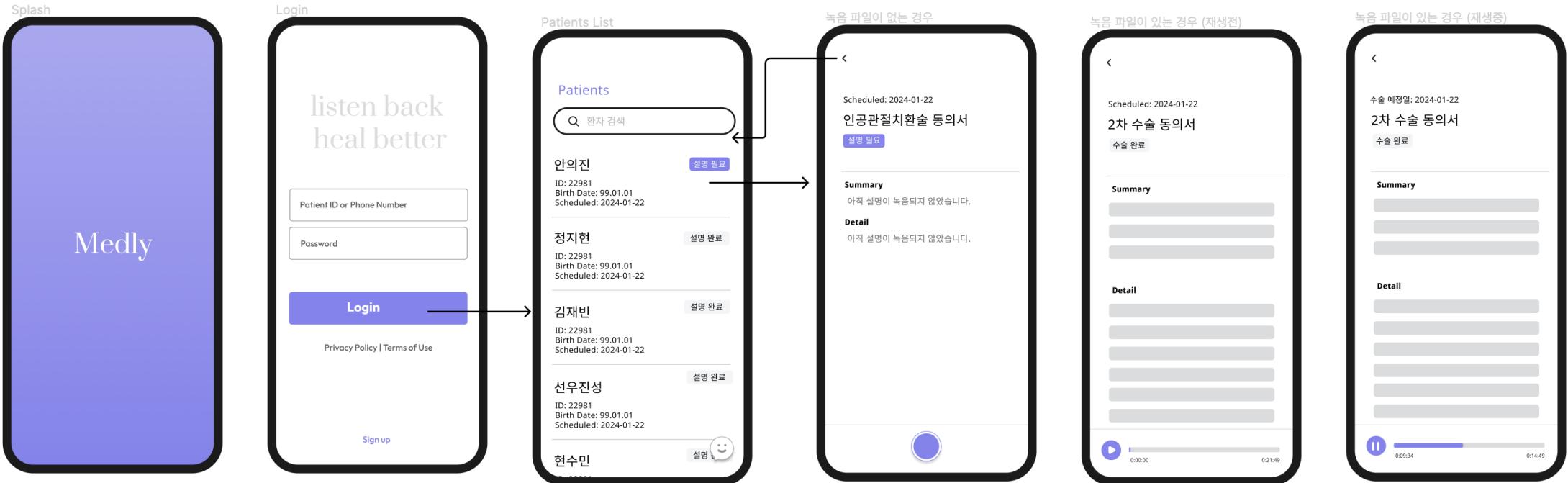
### Key Features



## 03 Product Design

### 녹음 요약 파트 화면 기획

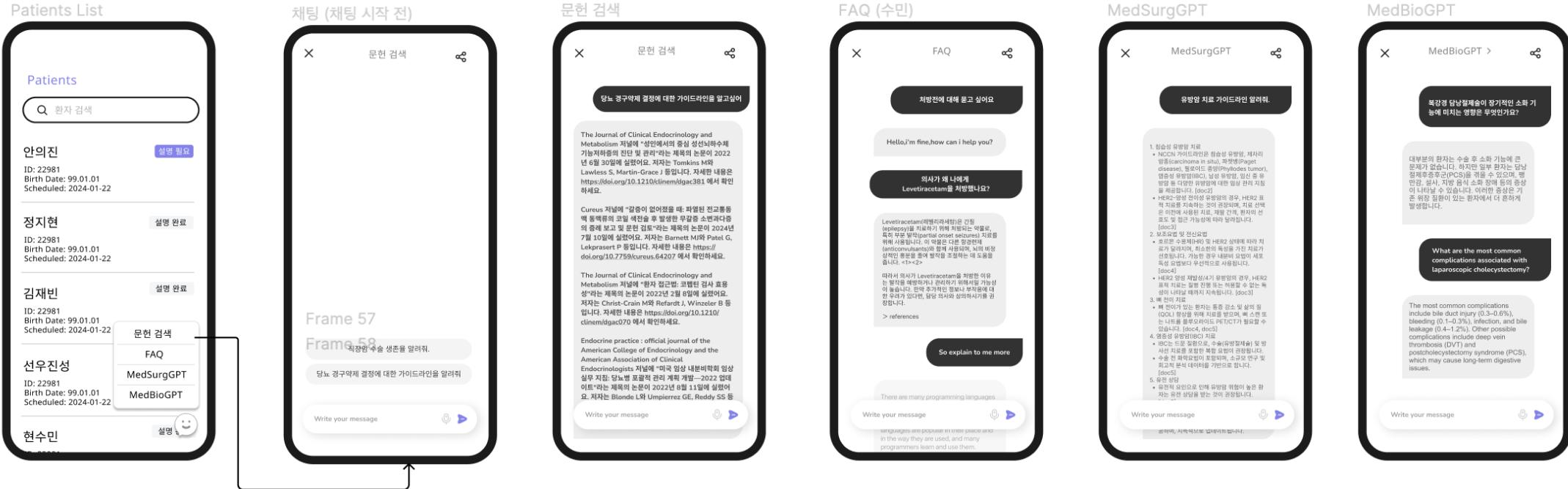
사용자 로그인 — 환자 목록 — 환자 선택 — 내용 열람 — 동의서 설명 녹음 — 재생



## 03 Product Design

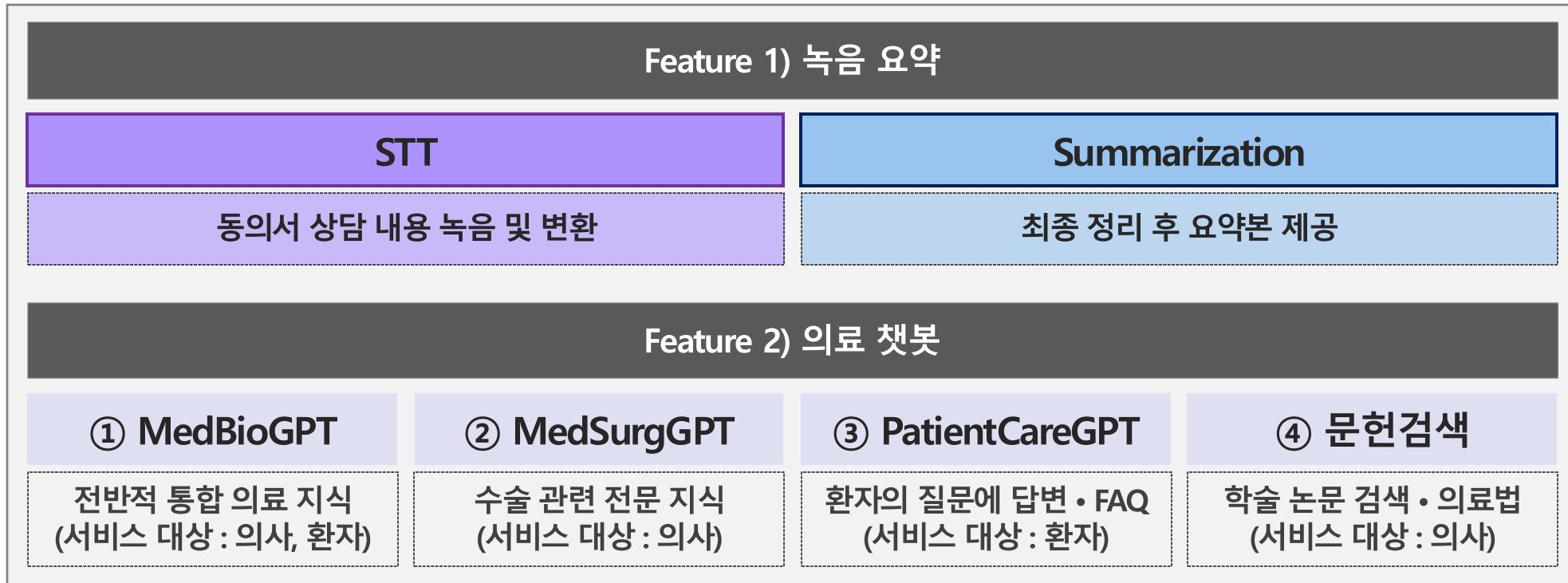
### 챗봇 파트 화면 기획

#### 채팅 플로팅 버튼 — 챗봇 : MedBioGPT / PatientCareGPT / MedSurgGPT / 문헌 검색 — 선택 — 채팅



## 04 AI 모델 개발

### STT Model Fine-tuning & Summarization Model



# 04 AI 모델 개발 (1) : STT Model Fine-tuning

## Dataset 생성

- 복강경 담낭절제술로 범위 한정하여 데이터 탐색 및 생성
- 동의서 및 관련 문헌 pdf 및 image를 Azure computer vision Read API를 이용하여 text로 변환
- 생성된 텍스트를 정제한 후 녹음하여, 오디오-대본 쌍을 생성
- 일반 텍스트 데이터
  - 수술 정보
  - 의학 용어
  - FAQ

**복강경 담낭절제술 동의서**

동 룩 번 호	환 자 명
나 이 / 성 별 /	시행예정일
진 단 명	

참여 의료진 (집도의가 다수인 경우 모두 기재해 주시기 바랍니다.)

집도의1	<input type="checkbox"/> 전문의 <input type="checkbox"/> 일반의	진료과목 _____
집도의2	<input type="checkbox"/> 전문의 <input type="checkbox"/> 일반의	진료과목 _____

다음의 설명 내용은 의료진이 환자 본인이나 대리인(보호자)에게 환자가 시행 받을 수술에 대한 정보를 제공하여 환자 본인이나 대리인(보호자)이 자의에 따라 수술여부를 결정할 수 있도록 하기 위한 것입니다. 설명 내용 중 이해가 되지 않는 부분이 있다면 설명하는 의료진에게 추가 설명을 요청할 수 있습니다. 환자(또는 대리인)는 이 동의서 또는 별지 사본에 대한 교부를 의무기록 사본 발급 절차에 따라 요청할 수 있으며, 이 요청이 있을 경우 지체 없이 교부하도록 합니다. 단, 동의서 또는 별지 사본 교부 시 소용되는 비용을 청구할 수 있습니다.

1. 담낭의 기능

간에서 합성된 담즙은 담관이라는 통로를 통해 심이지장으로 흘러 들어가게 되는데 담관의 중간에 담낭이 연결되어 있습니다. 담낭의 기능은 식간에 담즙을 저장 및 농축하고 있다가 식사 후 음식물이 위장에서 심이지장으로 이동하면 담낭이 수축하여 농축된 담즙을 심이지장으로 배출하고 이렇게 배출된 농축 담즙이 소화된 음식물에서 지방의 흡수를 돋는 것입니다.

복강경 담낭절제술 동의서  
등록번호  
환자  
명  
나 이 / 성 별 /  
시행예정일  
진단  
명  
참여 의료진 (집도의가 다수인 경우 모두 기재해 주시기 바랍니다.)  
집도의1 \_\_\_\_\_  
 전문의  일반의  
진료과목 \_\_\_\_\_  
집도의2 \_\_\_\_\_  
 전문의  일반의  
진료과목 \_\_\_\_\_  
다음의 설명 내용은 의료진이 환자 본인이나 대리인(보호자)에게 환자가 시행 받을 수술에 대한 정보를 제공하여 환자 본인이나 대리인(보호자)이 자의에 따라 수술여부를 결정할 수 있도록 하기 위한 것입니다. 설명 내용 중 이해가 되지 않는 부분이 있다면 설명하는 의료진에게 추가 설명을 요청할 수 있습니다. 환자(또는 대리인)는 이 동의서 또는 별지 사본에 대한 교부를 의무기록 사본 발급 절차에 따라 요청할 수 있으며, 이 요청이 있을 경우 지체 없이 교부하도록 합니다. 단, 동의서 또는 별지 사본 교부 시 소용되는 비용을 청구할 수 있습니다.  
1.  
담낭의 기능  
간에서 합성된 담즙은 담관이라는 통로를 통해 심이지장으로 흘러 들어가게 되는데 담관의 중간에 담낭이 연결되어 있습니다. 담낭의 기능은 식간에 담즙을 저장 및 농축하고 있다가 식사 후 음식물이 위장에서 심이지장으로 이동하면 담낭이 수축하여 농축된 담즙을 심이지장으로 배출하고 이렇게 배출된 농축 담즙이 소화된 음식물에서 지방의 흡수를 돋는 것입니다.  
2.  
담낭절제술의 목적 및 방법  
이러한 기능을 하는 담낭에 결석(담석)이 생겨 소화불량, 복통 등의 증상이 발생하거나 담낭염이 발생할 경우, 혹은 일정 크기 이상의 담낭 용종이 발생하거나, 혼담관합류이상이 등반된 경우 등에서 담낭절제술을 하게 됩니다. 담낭 결석 혹은 담낭 용종이 있는 경우 담석이나 용종만 제거할

STT Model Fine-tuning을 위한 Dataset 마련

# 04 AI 모델 개발 (1) : STT Model Fine-tuning

Azure Speech Studio / Custom Speech

어떤 형식의 데이터를 업로드하시겠습니까?

일반 텍스트  학습  테스트  
대상 시나리오와 관련된 많은 양의 구조화되지 않은 문장(한줄 용어 포함)이 포함된 텍스트 파일(.txt)입니다.

구조화된 텍스트  학습  테스트  
Markdown(.md) 파일의 클래스 및 연락처 목록

오디오 + 휴먼 레이블 대본  학습  테스트  
오디오 파일(zip)은 오디오 파일 이름을 나타내는 형식화된 스크립트(txt)와 이를 이루어 악간의 액션, 말하기 스타일 및 배경 소음과 같은 음향적 측면을 개선합니다. 또는 오류를 숫자로 모델 품질을 평가하세요.

오디오  학습  테스트  
정확도 번호 없이 모델 품질을 빠르게 검사하기 위한 오디오 파일의 컬렉션(zip)입니다.

음성 텍스트(자동 오디오 합성)  미리 보기  학습  테스트  
실제 오디오가 없나요? 일부 테스트 문장이 포함된 텍스트(txt) 파일을 업로드하면 각 음성 문장에 대한 오디오 쌍이 차동으로 합성되어 최신 인공신경망 텍스트 음성 변환 기술을 활용합니다.

출력 형식  미리 보기  학습  테스트  
최종 표시 출력을 정의하는 텍스트(txt) 파일의 어휘-표시 형식 규칙(예: 재구성, 유클리드, 액 텍스트 정규화 폴더) 목록입니다.

모델 ▾	단어 오류 비율 (WER) ⓘ ▾	삽입 (WER) ▾	대체 (WER) ▾	삭제 (WER) ▾	토큰 오류 비율 (TER) ▾	삽입 (TER) ▾	대체 (TER) ▾	삭제 (TER) ▾
모델 1: mixed12	6.63%	42 (0.80%)	183 (3.51%)	121 (2.32%)	13.32%	43 (0.80%)	369 (6.84%)	307 (5.69%)
모델 2: 20241203	7.49%	75 (1.44%)	206 (3.95%)	110 (2.11%)	14.02%	43 (0.80%)	388 (7.19%)	297 (5.50%)

모델 ▾	단어 오류 비율 (WER) ⓘ ▾	삽입 (WER) ▾	대체 (WER) ▾	삭제 (WER) ▾	토큰 오류 비율 (TER) ▾	삽입 (TER) ▾	대체 (TER) ▾	삭제 (TER) ▾
모델 1: mixed12	2.23%	18 (0.28%)	85 (1.33%)	39 (0.61%)	10.51%	41 (0.62%)	314 (4.73%)	343 (5.17%)
모델 2: 20241203	2.86%	29 (0.46%)	105 (1.65%)	48 (0.75%)	11.04%	41 (0.62%)	334 (5.03%)	350 (5.27%)

- Medly Model(mixed12)이 단어 오류 비율 (WER)과 토큰 오류 비율 (TER)에서 Base Model(20241203)보다 더 나은 성적을 보여주었음
- Medly Model에서 문맥에 큰 영향을 미치지 않는 일반 단어(조사 등)의 오류가 더 많았기 때문에 의학 전문 용어의 측면에서 비교시 더 우수함

준비한 데이터셋으로 Custom STT Model 학습 및 테스트

# 04 AI 모델 개발 (1) : STT Model Fine-tuning

Azure Speech Studio / Custom Speech

휴면 레이블 대화 기록(정규화됨) ① ▾

일 받으실 수술이 복강경 **담낭절제술** 이신데 알고계시죠  
음 김땡땡님의 담낭 담낭이 이제 급성으로 염증이 생겨서

복 수술에 비해서 회복 속도도 빠르고 입원 기간도 짧아서  
좀 어 더 이득이 있고 통증도 적은 편이에요 그래서 **상처**  
부작용하고 호흡기 합

그리고 검은 변이 나오거나 심하게 어지러운 경우 출혈의  
정후일 수 있으니 바로 알려주셔야 하고 **배액관이** 있을 경  
우에 지속적으로 담즙이 나오는지 잘 관찰해야 합니다 제

증이 발생할 수 있는데 어 **담관이** 손상되거나 담즙이 이제  
손상된 부위로 **누출되는** 경우가 있어요 그리고 이제 담관  
이 협착이 돼서 어 문제가 생길수도 있고 크흠 수술 중이

어요 그래서 이렇게 되는 경우에는 항생제 치료를 우선적  
으로 해야되고 어 그 **누출** 부위를 바깥으로 빼주는 ...

지 뭐 그런 그런 방법은 없을까요 어 이제 **급성 담낭염**의  
경우에는 어 일단은 가장 최선의 치료책이 담낭절제술을  
하는것으로 이제 되어있고요 수술을 받지 않으면 담낭염

모델 1: mixed12 (어휘) ① ▾

받으실 수술이 복강경 **담낭절제술** 이신데 알고 계시죠 @  
김땡땡님의 담낭 담낭이 이제 급성으로 염증이 생겨서 이

술에 비해서 회복 속도도 빠르고 입원 기간도 짧아서 좀  
어 더 이득이 있고 통증도 적은 편이에요 저 **상처** 부작용  
하고 호흡기 합

그리고 검은 변이 나오거나 심하게 어지러운 경우 출혈의  
정후일 수 있으니 바로 알려주셔야 하고 **배액관이** 있을 경  
우에 지속적으로 담즙이 나오는지 잘 관찰해야 합니다 제

증이 발생할 수 있는데 어 **담관이** 손상되거나 담즙이 이제  
손상된 부위로 **누출된** 경우가 있어요 그리고 담관이 이제  
협착이 돼서 어 문제가 생길 수도 있고 수술 중이나 수술

그래서 이렇게 되는 경우에는 항생제 치료로 우선적으로  
해야 되고 어 그 **누출** 부위를 바깥으로 빼주는 ...

든지 뭐 그런 그런 방법은 없을까요 어 이제 **급성 담낭염**  
의 경우에는 어 일단은 가장 최선의 치료책이 담낭절제술을  
하는 것으로 이제 되어 있고요 수술을 받지 않으면 담

모델 2: 20241203 (어휘) ① ▾

받으실 수술이 복강경 **담당 절제술** 이신데 알고 계시죠 그  
김땡땡님의 담낭 담낭이 이제 급성으로 염증이 생겨서 이

복 수술에 비해서 회복 속도도 빠르고 입원 기간도 짧아서  
좀 어 더 이득이 있고 통증도 적은 편이에요 그래서 **성적**  
부작용하고 호흡기 협

그리고 검은 변이 나오거나 심하게 어지러운 경우 출혈의  
정후일 수 있으니 바로 알려주셔야 하고 **베이관이** 있을 경  
우에 지속적으로 담즙이 나오는지 잘 관찰해야 합니다 제

병증이 발생할 수 있는데 어 **당분간** 손상되거나 담즙이 이  
제 손상된 부위로 **노출된** 경우가 있어요 그리고 담관이 이  
제 협착이 돼서 어 문제가 생길 수도 있고 수술 중이나 수술

래서 이렇게 되는 경우에는 항생제 치료로 우선적으로 해  
야 되고 어 그 **노출** 부위를 바깥으로 빼주는 ...

다든지 뭐 그런 방법은 없을까요 어 이제 **급성장염의** 경우  
에는 어 일단은 가장 최선의 치료책이 담당 절제술을 하는  
것으로 이제 되어 있고요 수술을 받지 않으면 담낭염이 악

Custom STT Model과 Base Model 간 비교 : 수술과 관련된 용어 인식 — Medly 모델이 우수함

# 04 AI 모델 개발 (1) : STT Model Fine-tuning

Azure Speech Studio / Custom Speech

휴면 레이블 대화 기록(정규화됨) ① ▾

하시는 것이 훨씬 더 환자분에게 이득이라고 어 생각이 되고 그렇기 때문에 이 수술을 시행하시는 겁니다 어 제가

하실 수 있어요 그래서 이렇게 가쁜 다음에 담낭을 제거하는 수술이신데 이제 복강경으로 할 수 없는 정도로 염증이 심하거나 이렇게 막 예전에 수술 하셨어서 유착이 심하거나 뭐 아니면 뭐 암이 의심되거나 할 때는 개복수술로

수술 부위 감염은 어떻게 알 수 있나요 창상 감염은 수술

필요하다고 이제 생각이 된 상태시고 이제 급성 담낭염이더라도 항생제로 염증을 안정시키고 환자분 상태를 조절

기구에 넣어서 꺼내기 때문에 주변 장기가 뭐 손상되거나 감염되거나 하지도 않으실 거예요 그래서 이렇게 안전하

아 네 지금은 이제 병원에서 혈당을 조절해드리고 있는데 이제 금식이 풀리면 수술하고 나서 식사를 하시기 시작하

모델 1: mixed12 (어휘) ① ▾

하시는 것이 훨씬 더 환자분에게 이득이라고 어 생각이 되고 그렇기 때문에 이 수술을 시행하시는 겁니다 제가 설명

실 수 있어요 그래서 이렇게 가쁜 다음에 담낭을 제거하는 수술이신데 이제 복강경으로 할 수 없는 정도로 염증이 심하거나 이렇게 막 예전에 수술하셨어서 유착이 심하거나 뭐 아니면 뭐 암이 의심되거나 할 때는 개복수술로

수술 부위 감염은 어떻게 알 수 있나요 창상 감염은 수술

하다고 이제 생각이 된 상태시고 이제 급성 담낭염이더라도 항생제로 염증을 안정시키고 환자분 형태를 조절한 뒤

은 기부에 넣어서 꺼내기 때문에 주변 장기가 뭐 손상되거나 감염되거나 하지도 않으실 거예요 그래서 이렇게 안전

아 네 지금은 이제 병원에서 혈당을 조절해드리고 있는데 이제 금식이 풀리면 수술하고 나서 식사를 하시기 시작하

모델 2: 20241203 (어휘) ① ▾

하시는 것이 훨씬 더 환자분에게 이득이라고 어 생각이 되고 그렇기 때문에 이 수술 후에 시행하시는 겁니다 어 제

실 수 있어요 그래서 이렇게 가쁜 다음에 담낭을 제공하는 수술이신데 이제 복강경으로 할 수 없는 정도로 염증이 심하거나 이렇게 막 예전에 수술하셨어서 유착이 심하거나 뭐 아니면 뭐 암이 의심되거나 할 때는 계속 수술로

수술 부위 감염은 어떻게 알 수 있나요 청산 감염은 수술

하다고 이제 생각이 된 상태시고 이제 급성 담낭염이더라도 항생제로 염증을 안정시키고 환자분 형태를 조절한 뒤

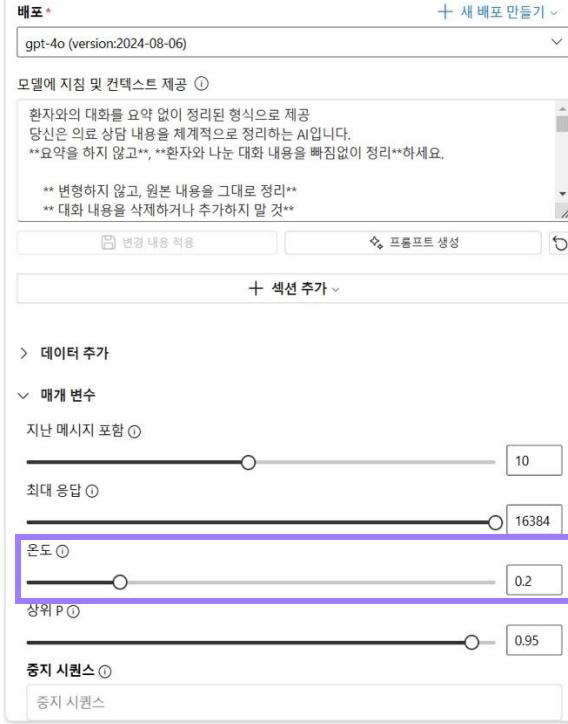
에 넣어서 꺼내기 때문에 주변 전기가 손상되거나 감염되거나 하지도 않으실 거예요 그래서 이렇게 안전하게 수술

아 네 지금은 이제 병원에서 혈당을 조절해드리고 있는데 이제 금식이 풀리면 조절하고 나서 식사를 하시기 시작하

Custom STT Model과 Base Model 간 비교 : 수술과 관련된 용어 인식 — Medly 모델이 우수함

# 04 AI 모델 개발 (2) : Summarization Model

Azure Open AI / gpt-4o



**prompt = f""**  
당신은 의료 상담 내용을 체계적으로 정리하는 AI입니다.  
\*\* 요약을 하지 않고\*\*, \*\*환자와 나눈 대화 내용을 빠짐없이 정리하세요.\*\*  
\*\* 변형하지 않고, 원본 내용을 그대로 정리\*\*  
\*\* 전체 대화 내용 중 의료에 관련된 내용은 빠짐없이 어딘가에 반드시 한 번은 등장해야 함\*\*  
\*\* 불필요한 압축 없이, 키워드 형식으로 정리\*\*  
\*\* 진단, 치료 계획, 부작용, 주의사항 등을 명확히 분류\*\*  
\*\* '진단' 섹션은 상담의 주제와 관련된 한 가지 질환과 그에 관한 내용만 포함하며, 이후 등장하는 기저질환 정보는 '기타 사항'으로 이동해야 함\*\*  
\*\* '진단' 섹션의 첫 줄에는 '진단명'만 명시, 현재 증상 및 진행 상태 자세히 설명\*\*  
\*\* '진단' 섹션의 '암의 가능성' 부분에 관련된 설명을 자세히 설명하고, 만약 관련된 내용이 없다면 해당 항목을 삭제하세요.\*\*  
\*\* '진단' 섹션의 '암의 가능성' 부분에 관련된 설명을 자세히 설명하고, 만약 관련된 내용이 없다면 해당 항목을 삭제하세요.\*\*

## 1번 프롬프트

**prompt = f""**  
다음은 복록별로 정리된 의료 상담 내용입니다.  
이를 최종적으로 하나의 문서로 깔끔하게 통합 정리하세요.  
\*\*요약을 하지 않고\*\*, \*\*환자와 나눈 대화 내용을 빠짐없이 정리\*\*, \*\*키워드 형식으로\*\* 출력하세요.  
  
\*\* 변형하지 않고, 원본 내용을 그대로 정리\*\*  
\*\* 기존 섹션을 중복 생성하지 말고, 기존에 존재하는 섹션에 병합\*\*  
\*\* 불필요한 내용 중복은 제거하고, 동일한 내용은 적절히 병합하세요.\*\*  
\*\* 질문과 답변은 정확히 매칭되도록 주의하세요.\*\*  
\*\* 모든 섹션 제목 앞에 '.' 사용\*\*  
\*\* 전체 대화 내용 중 의료에 관련된 내용은 빠짐없이 어딘가에 반드시 한 번은 등장해야 함\*\*  
\*\* [질문] → [답변]은 해당하는 내용만 그 형식으로 출력하고 '질문'과 '답변' 단어는 직접적으로 출력하지 마세요.\*\*  
\*\* [질문] → [답변]에 해당하는 내용이 없는 경우 해당 항목을 삭제하세요.\*\*  
\*\* '진단' 섹션은 상담의 주제와 관련된 한 가지 질환과 그에 관한 내용만 포함하며, 이후 등장하는 기저질환 정보는 '기타 사항'으로 이동해야 함\*\*  
\*\* '진단' 섹션의 '암의 가능성' 부분에 관련된 내용을 환자의 현재 상태와 더불어 자세히 설명하고, 만약 관련된 내용이 없다면 해당 항목을 삭제하세요.\*\*  
\*\* 출력 순서는 반드시 유지하세요.\*\*

## 2번 프롬프트

- \*\*출력 형식 (반드시 이 순서 유지할 것):\*\*
  - \*\*진단:\*\*
    - [진단명]
    - [진단 근거]
    - [현재 증상 및 진행 상태]
    - [암의 가능성]
  - \*\*수술/치료 계획:\*\*
    - [수술명]
    - [치료 선택 이유]
    - [수술 과정 및 방법]
    - [마취 방법]
    - [예상 효과]
    - [의료진의 설명]
    - [질문] → [답변]
  - \*\*부작용 및 위험:\*\*
    - [수술 중 예상 가능한 문제]
    - [수술 후 발생할 수 있는 부작용]
    - [대처 방법]
    - [질문] → [답변]
    - [질문] → [답변]
  - \*\*주의사항:\*\*
    - [수술 전후 지켜야 할 점]
    - [유의해야 할 증상]
    - [예방해야 할 문제]
    - [예방을 위한 방법]
    - [생활습관 관련]
    - [질문] → [답변]
  - \*\*기타 사항:\*\*
    - [기저질환 및 복용 중 약물:] - 모든 기저질환과 복용 중인 약물을 목록화하세요.
    - [과거력 및 수술력:]
    - [환자의 일반적인 건강 상태:]
    - [가족력:]
    - [기타:]
    - [질문] → [답변]

공통

Azure Open AI Prompting을 통해 Summarization Model 개발

04 AI 모델 개발 (2) : Summarization Model

# Open AI / gpt-4o

## STT 결과물

## Summarization Model

지나치게 긴 입력값으로 처리 실패

Text 블럭화 : 2000자 단위

문장이 중간에서 잘리지 않도록 문장 부호 기준으로 나누도록 하고, 문맥 유실을 방지하기 위해 앞 뒤 블럭의 100자 가량씩 중복되도록 설정

# 1번 프롬프트로 블럭별 정리 후 학습

## 2번 프롬프트로 최종 정리 및 요약

## 22분의 음성대화 : 약 5000 단어

# 04 AI 모델 개발 (2) : Summarization Model

Open AI / gpt-4o

## • 진단:

- 담낭 용종
- 최근 검사에서 용종 크기가 증가하는 경향이 있었고 크기가 8mm 정도임. 전형적인 양성 용종의 특징을 보임.
- 현재 증상 및 진행 상태: 담낭 용종은 담낭 안쪽벽에 혹처럼 자라는 돌기. 크기가 작고 변화가 없는 경우에는 지켜보기도 하지만, 크기가 1cm 이상이거나 빠르게 커지는 경우, 담석이 동반된 경우에는 담낭암 위험이 있어 예방적으로 절제 권장. 환자의 경우 최근 크기 증가로 수술 결정.
- 암의 가능성: 담낭 용종의 90% 이상은 양성이나마, 1cm 이상이거나 빠르게 자라는 경우 악성 가능성 5-10%. 초음파 및 CT 검사상 주변 조직 침범 소견 없음. 암일 경우 초기 단계일 가능성 높음.

## • 수술/치료 계획:

- 복강경 담낭절제술
- 치료 선택 이유: 용종 크기 증가로 인한 예방적 절제, 비교적 안전한 수술, 합병증 발생률 낮음
- 수술 과정 및 방법: 복강경으로 진행, 배에 작은 구멍 3-4개 뚫고 카메라와 기구 삽입하여 담낭 제거
- 마취 방법: 전신마취, 수술 전 마취과에서 평가 및 조절
- 예상 효과: 회복 빠르고 통증 적음, 담낭 제거로 인한 증상 완화
- 의료진의 설명: 수술 시간 1-2시간, 입원 2-3일. 특수한 경우 개복수술로 전환 가능, 담당 집도의가 직접 수술 예정, 부득이 한 사유로 변경될 가능성 있음
- 혹시 제 담낭 용종이 암일 가능성도 있나요? → 네, 일부 경우 악성 가능성 있음. 크기 크거나 빠르게 자라는 경우, 담석 동반 시 가능성 높아짐.
- 만약 수술 후에 조직 검사에서 암으로 나오면 어떻게 해요? → 암세포 발견 시 진행 정도 확인. 담낭벽 안쪽 국한 시 경과 관찰, 벽 넘어간 경우 추가 검사 필요. 현재 검사상 주변 조직 침범 소견 없어 초기 단계 가능성 높음.

## • 부작용 및 위험:

- 수술 중 예상 가능한 문제: 복강경에서 개복수술로 전환 가능성 (출혈 심하거나 염증 심한 경우, 복강 내 시야 확보 어려운 경우), 장기 손상 가능성
- 수술 후 발생할 수 있는 부작용: 기름진 음식 섭취 시 소화 불편, 더부룩함, 설사, 장유착, 장폐색, 혈전 형성, 폐색전증
- 대처 방법: 기름진 음식 섭취 피하고 천천히 늘려가기, 수술 후 가벼운 움직임, 다리 운동, 빠른 걷기
- 담낭 제거 후 건강 문제 여부 → 큰 문제 없음, 몸이 자연스럽게 적응
- 황달 발생 이유 문의 → 담관 손상, 협착 가능성 설명
- 수술 부위 감염 확인 방법 문의 → 상처 붓기, 염증, 고름 발생 시 의심

## • 주의사항:

- 수술 전후 지켜야 할 점: 기름진 음식 섭취 피하기, 가벼운 움직임, 다리 운동, 빠른 걷기
- 유의해야 할 증상: 소화 불편, 더부룩함, 설사, 38도 이상의 고열, 심한 복통, 황달, 지속적인 구토, 복부 부종, 검은 변, 어지러움
- 예방해야 할 문제: 담낭암, 장유착, 혈전 형성
- 예방을 위한 방법: 담낭 절제, 수술 중 지혈, 담즙누출 시 경피적 담즙 배액술 또는 내시경적 역행성 담췌관 조영술
- 생활습관 관련 조언: 무거운 짐 들기 및 복부에 힘이 들어가는 운동 1달간 피하기, 수술 부위 깨끗이 유지

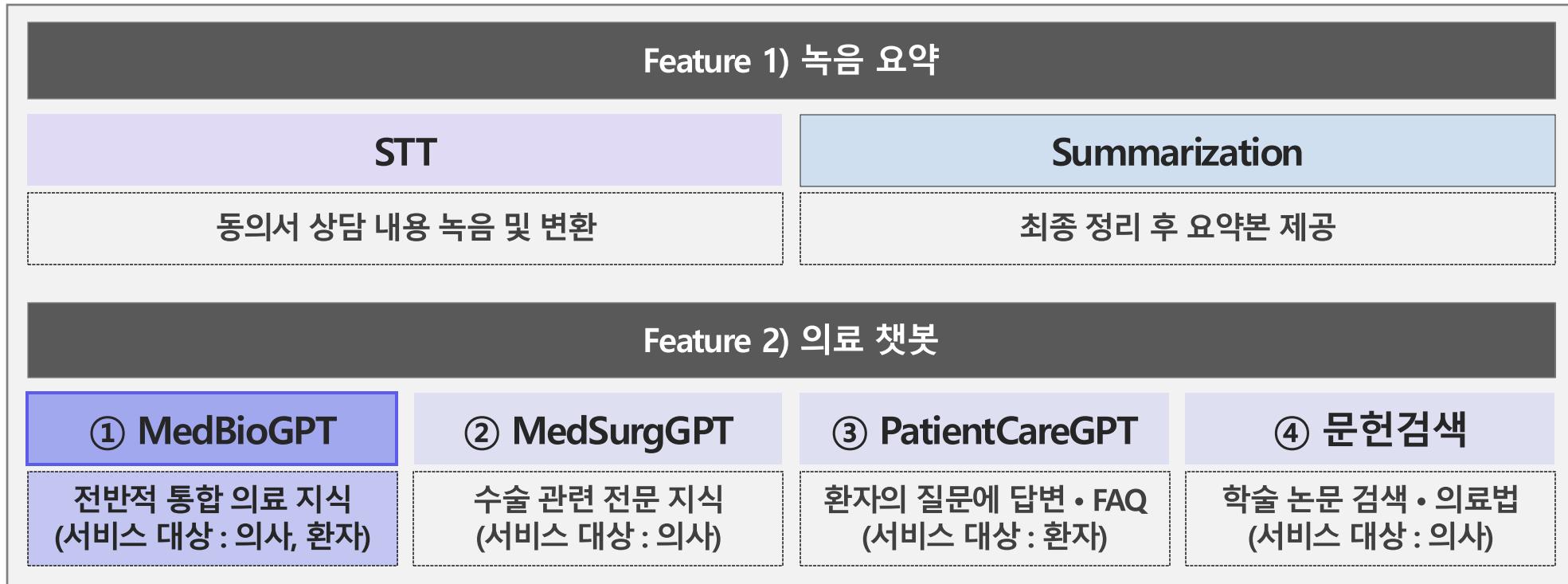
## • 기타 사항:

- 기저질환 및 복용 중 약물: 갑상선 기능저하증, 레보티록신 복용 중, 빈혈 및 혈액 응고 장애 없음
- 과거력 및 수술력: 제왕절개 수술력
- 환자의 일반적인 건강 상태: 갑상선 기능저하증 외 특별한 문제 없음, 검사 결과상 출혈 위험 낮음
- 가족력: 가족 중 암환자 없음
- 기타: 수술 후 회복기간 2-3일, 퇴원 후 일상 활동 가능, 무거운 짐 들기 및 복부에 힘이 들어가는 운동 1달간 피하기, 수술 부위 감염 예방 위해 목욕탕 및 수영장 1달간 삼가, 퇴원 후 약 1~2주 뒤 외래 방문하여 경과 확인, 상처 상태 및 불편한 점 확인 예정

잘 정돈된 문서로 출력되며 STT에서 발생한 단어 오류도 Open AI의 Summarization 과정에서 교정됨

## 04 AI 모델 개발

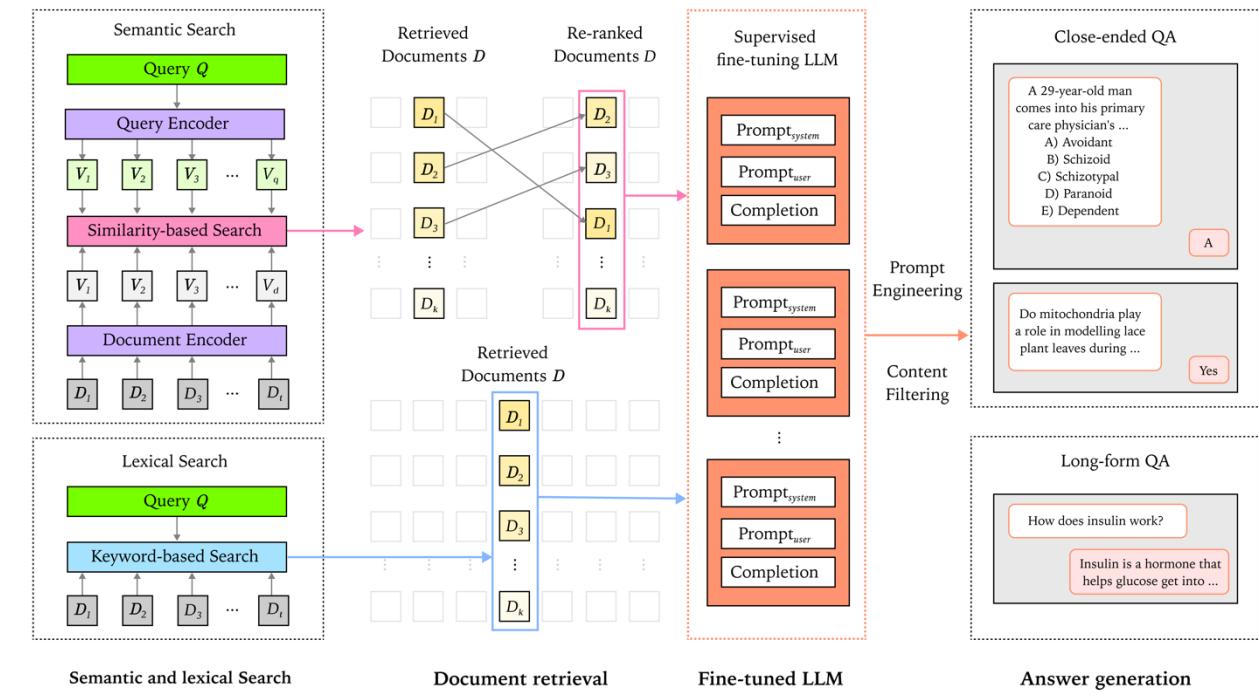
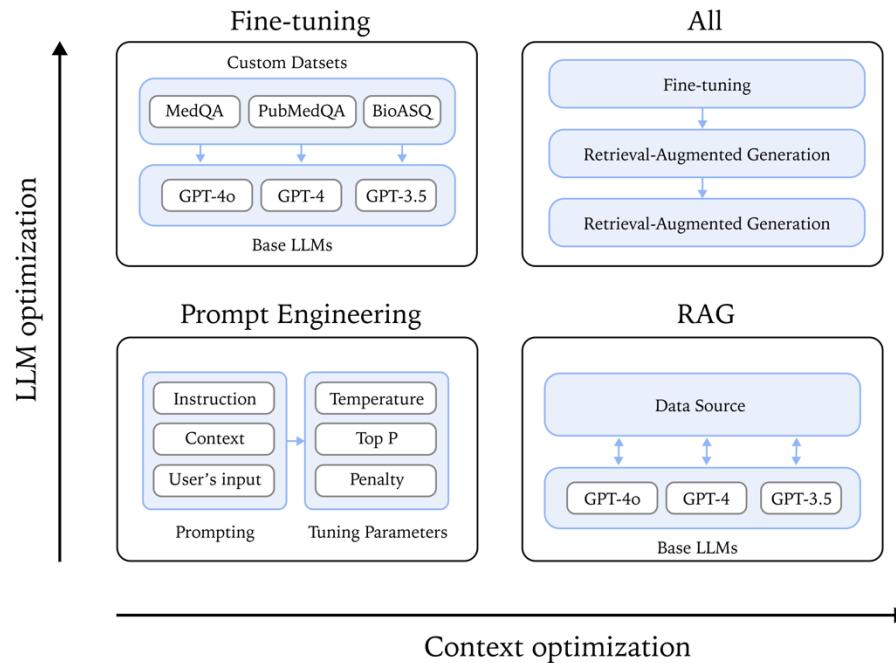
ChatBot : MedBioGPT



# 04 AI 모델 개발 (3) : MedBioGPT

## Overview

MedBioGPT는 Fine-tuning, Prompt Engineering, RAG 방식을 모두 활용하여 의료 및 생명과학 분야의 QA 시스템을 최적화한 모델



사용자는 Azure AI Search의 Semantic Search 기능을 중심으로 RAG(Retrieval-Augmented Generation) 시스템을 구성하고, Azure OpenAI의 Fine-tuning 기능을 활용하여 GPT 모델을 Fine-tuning 하여 AI Search와 결합함

## 04 AI 모델 개발 (3) : MedBioGPT

### Fine-tuning GPT4o

- 의료 QA 성능 향상을 위해 MedQA, BioSQA, PubMedQA 등의 데이터셋을 활용하여 Closed-ended, Long-form, Short-form QA 모델을 Azure OpenAI에서 Fine-tuning 진행했으며, 학습 결과, Validation Loss가 점진적으로 감소하며 모델이 안정적으로 수렴하는 것을 확인함

Task	Train Dataset	Train Samples	Training Duration	Epochs	Batch Size
Closed-ended QA	MedQA	10,178	3h 25m 33s	2	13
	PubMedQA (PQA-L)	552	7h 46m 44s	3	1
	BioSQA	5,049	3h 10m 7s	3	2
Long-form QA	PubMedQA (PQA-A)	196,144	1d 6h 29m 20s	1	64
	MedicationQA	551	1h 44m 18s	3	1
	LiveQA	500	1h 46m 17s	3	1
	BioSQA	5,049	2h 36m 49s	3	10
	Combined Custom Dataset	6,652	2h 6m 1s	3	13
Short-form QA	MedQA	10,178	1h 49m 44s	2	13

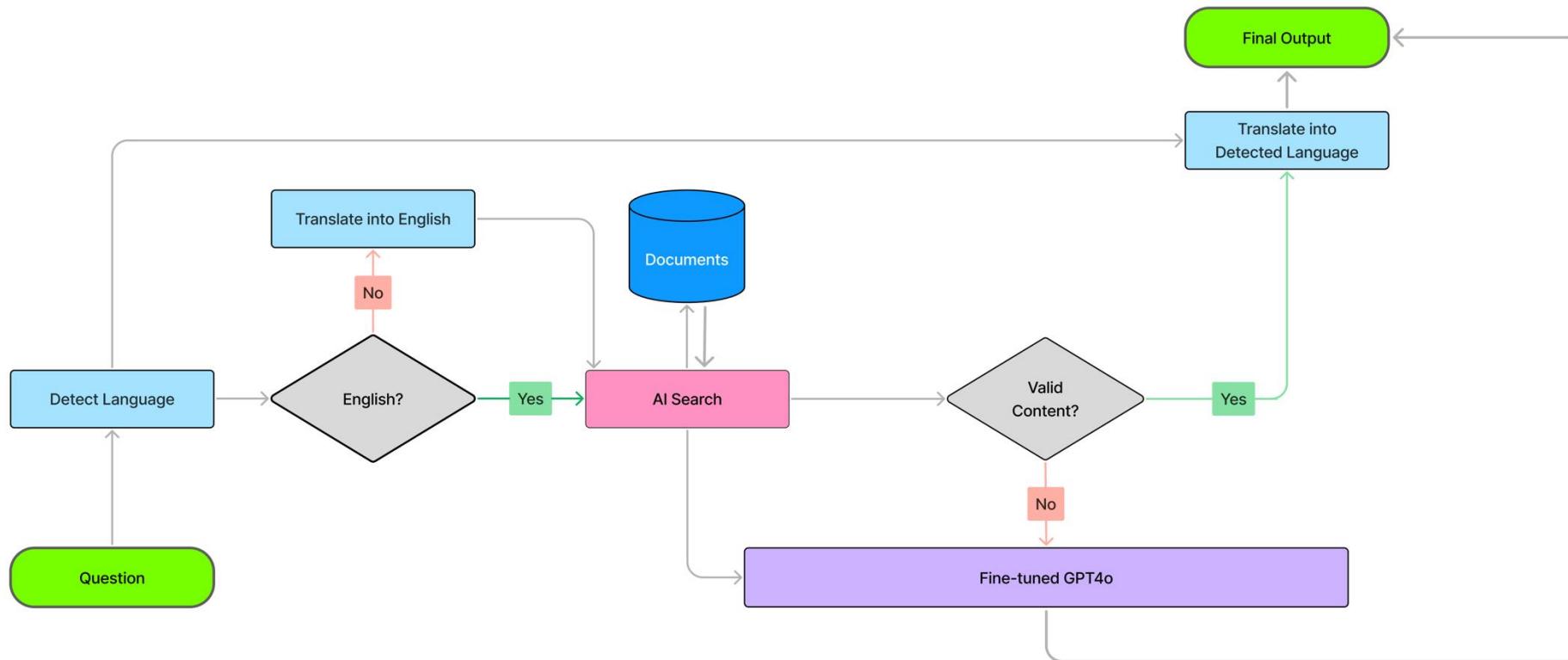
Experimental Details and Dataset Statistics



## 04 AI 모델 개발 (3) : MedBioGPT

### Search and Answer Generation Optimization

- 한국어로 검색 시 영어로 된 데이터 소스에서 검색을 잘 수행하지 못해서 언어 감지 후 영어가 아닐 경우 번역하여 AI Search를 수행하고, 검색된 내용이 유효하면 이를 기반으로 응답을 생성함
- 유효한 결과가 없을 경우, Fine-tuned GPT-4o를 활용하여 보완된 답변을 제공함



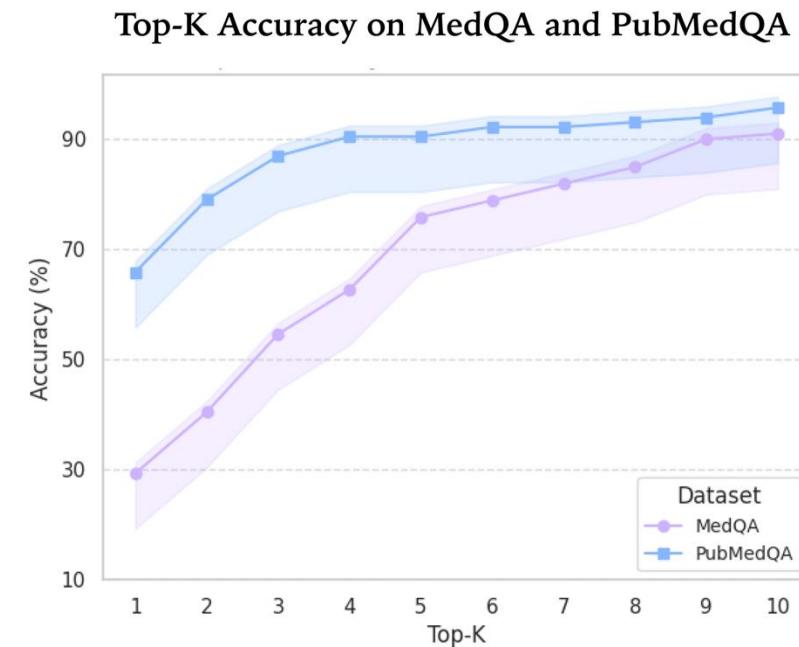
## 04 AI 모델 개발 (3) : MedBioGPT

### Azure AI Search Performance Evaluation

- Semantic Search가 Lexical Search보다 DCG@10, NDCG@10, MRR@10, Precision@10 등 대부분의 평가 지표에서 더 높은 성능을 보이며 검색 정확도를 향상시킴을 확인함
- Top-K Accuracy 분석 결과, PubMedQA와 MedQA 데이터셋 모두에서 K 값이 증가할수록 검색 정확도가 증가

Dataset	NFCorpus		TREC-COVID	
	Lexical	Semantic	Lexical	Semantic
Metric				
DCG@10	2.65	3.27	4.39	5.55
NDCG@10	31.34	37.91	48.35	61.02
MRR@10	51.63	64.29	82.50	89.17
Precision@10	23.04	27.88	49.60	64.20
Recall@10	15.95	18.70	0.43	0.54
F1-score@10	12.61	14.99	0.85	1.07
MAP@10	46.01	56.15	72.31	82.19

Text Retrieval Performance



Top-K Accuracy

# 04 AI 모델 개발 (3) : MedBioGPT

## Prompt Engineering

- 디코딩 과정에서는 각 태스크에 적합한 System Message 및 하이퍼파라미터(Temperature, Top-P 등)를 Task 별로 고정하여 비교 실험 진행

Tasks	Closed-Ended QA MedQA	Closed-Ended QA PubMedQA	Closed-Ended QA BioASQ	Long-Form QA MedicationQA, PubMedQA, BioASQ, LiveQA
<b>System Message w/o AI Search</b>	Answer with A, B, C, Answer only with 'yes' or 'no'. Do not add anything else.	Answer only with 'yes' or 'no'. Do not add anything else.	Answer only with 'yes' or 'no'. Do not add anything else.	You are a medical and biological expert. Answer the question accurately.
<b>System Message w/ AI Search</b>	Answer with A, B, C, Answer only with 'yes' or 'no'. Do not add anything else.	Answer only with 'yes' or 'no'. Do not add anything else.	Answer only with 'yes' or 'no'. Do not add anything else.	You are a medical and biological expert. Use the retrieved context to answer the question accurately.
<b>Prompt w/o AI Search</b>	{query}	{query}	{query}	{query}
<b>Prompt w/ AI Search</b>	Question: {query} Context: {retrieved documents}	Question: {query} Context: {retrieved documents}	Question: {query} Context: {retrieved documents}	Question: {query} Context: {retrieved documents}
<b>Max Tokens</b>	800	800	800	2999
<b>Temperature</b>	0.7	0.7	0.7	0.7
<b>Top P</b>	0.95	0.95	0.95	0.95
<b>Frequency Penalty</b>	0.0	0.0	0.0	0.0
<b>Presence Penalty</b>	0.0	0.0	0.0	0.0
<b>Stop Sequence</b>	None	None	None	None

Prompts and Decoding Parameters

The image shows two screenshots of the MedBioGPT interface. The top screenshot displays a question about laparoscopic cholecystectomy and its impact on long-term digestive function, followed by a response stating that lipid absorption is unaffected, a note about 1 reference, and a citation link. The bottom screenshot shows a question about the most common complication of gallbladder removal, followed by a response stating it's gallbladder removal itself, and another question about the overall mortality rate compared to gallbladder removal.

Long-form QA Examples

## 04 AI 모델 개발 (3) : MedBioGPT

### QA Performance Comparision

- Long-form QA 성능 평가 결과로, ROUGE, BLEU, BERTScore, BLEURT 등의 지표를 사용하여 모델의 답변 생성 품질을 측정하였으며, Fine-Tuned GPT-4o 모델에 AI Search를 결합하면 대부분의 데이터셋에서 ROUGE 및 BLEU 점수가 상승하여 더 나은 응답 품질을 제공함을 확인할 수 있었음
- Close-ended QA 성능 평가 결과에서 Accuracy(정확도)를 측정한 결과에서도 Fine-Tuned GPT-4o와 AI Search를 결합했을 때 성능이 가장 우수하였으며, 다른 모델 대비 높은 정확도를 기록함

Dataset	Model	ROUGE-1	ROUGE-2	ROUGE-L	BLEU	BERTScore	BLEURT
LiveQA	Fine-Tuned GPT-4o	24.12	6.18	13.31	1.63	1.10	-46.48
	+ AI Search	15.73	4.58	10.74	1.20	2.29	-86.99
	GPT-4o	26.96	5.80	13.42	1.41	-2.93	-34.79
	+ AI Search	<b>27.33</b>	<b>6.39</b>	<b>13.42</b>	<b>15.29</b>	<b>-1.60</b>	<b>-29.99</b>
MedicationQA	Fine-Tuned GPT-4o	24.69	8.80	17.61	2.49	<b>8.98</b>	-33.82
	+ AI Search	<b>27.73</b>	<b>15.09</b>	<b>22.72</b>	7.24	8.79	-33.63
	GPT-4o	22.92	13.69	18.70	<b>7.89</b>	8.55	<b>-6.92</b>
	+ AI Search	19.85	4.20	10.97	0.98	-7.63	-33.21
PubMedQA	Fine-Tuned GPT-4o	35.82	13.55	26.09	4.34	35.33	-9.23
	+ AI Search	<b>37.49</b>	<b>14.78</b>	<b>27.89</b>	<b>6.11</b>	<b>37.02</b>	<b>-3.89</b>
	GPT-4o	25.72	9.02	17.05	2.48	17.04	-9.04
	+ MedBioRAG	26.39	9.55	17.47	2.73	18.10	-7.86
BioASQ	Fine-Tuned GPT-4o	32.69	16.84	25.11	<b>6.52</b>	32.97	<b>-2.41</b>
	+ AI Search	<b>34.30</b>	<b>18.81</b>	<b>27.74</b>	6.12	<b>35.43</b>	-15.44
	GPT-4o	13.97	5.51	10.08	1.27	0.22	-24.84
	+ AI Search	22.29	8.21	15.64	2.27	11.60	-12.50

Long-form QA Performance

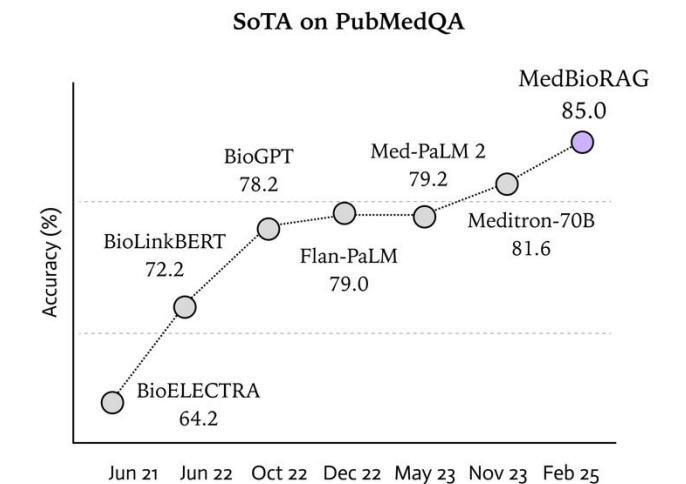
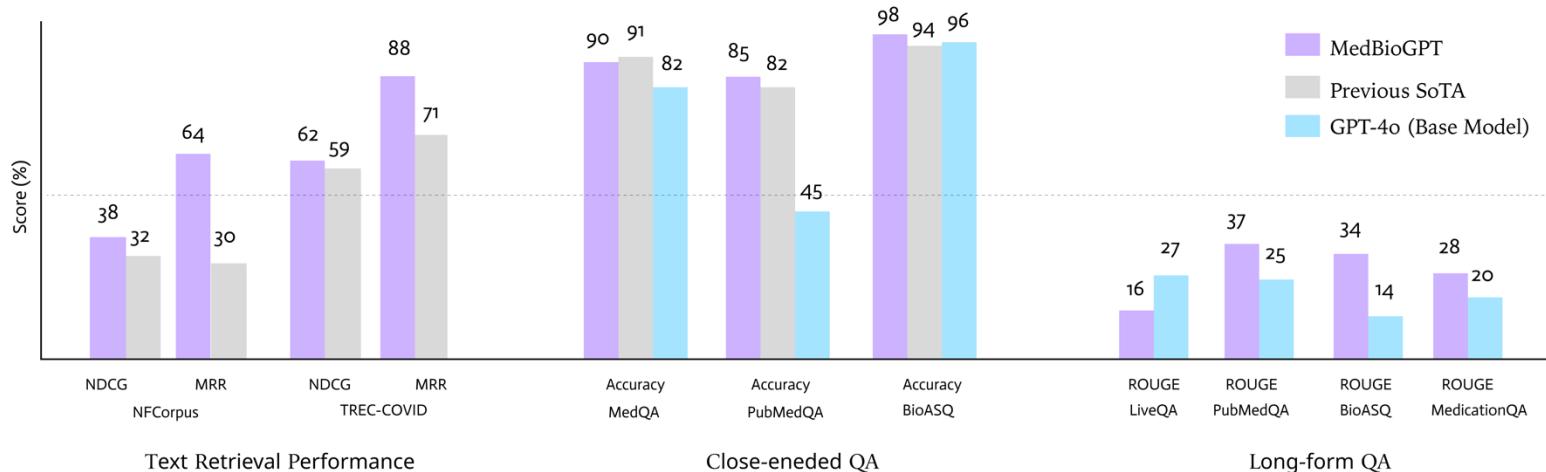
Method	MedQA	PubMedQA	BioASQ
Fine-Tuned GPT-4o	87.88	80.70	97.06
+ AI Search	<b>89.47</b>	<b>85.00</b>	<b>98.32</b>
GPT-4o	81.82	44.74	96.12
+ AI Search	<b>86.86</b>	<b>66.67</b>	<b>97.06</b>
GPT-4o-mini	67.68	77.55	96.32
+ AI Search	<b>70.71</b>	<b>76.32</b>	<b>97.06</b>
GPT-4	66.67	52.63	96.32
+ AI Search	<b>78.79</b>	<b>72.81</b>	<b>97.79</b>
GPT-3.5	51.52	19.30	88.24
+ AI Search	<b>45.36</b>	<b>38.60</b>	<b>66.91</b>

Close-ended QA Performance (Accuracy)

## 04 AI 모델 개발 (3) : MedBioGPT

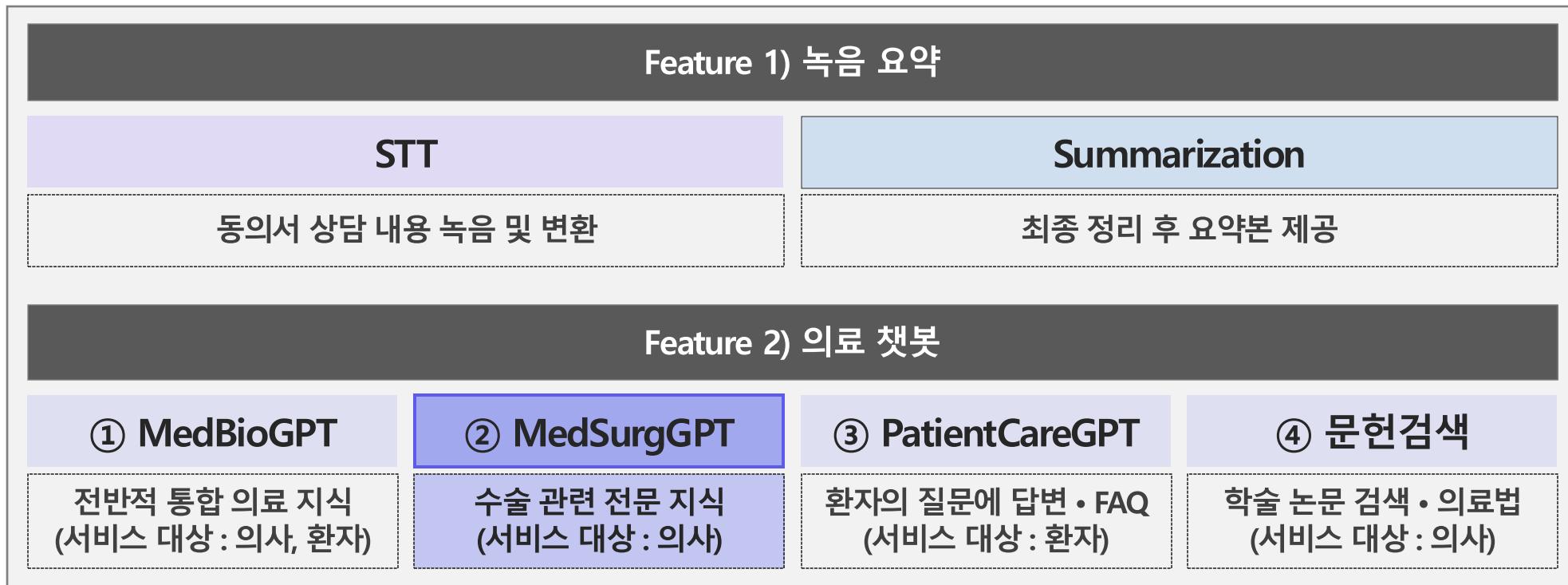
### SoTA Comparison

- MedBioGPT는 정보 검색(NDCG, MRR), Close-ended QA(정확도), Long-form QA(ROUGE)에서 기존 SOTA(State-of-the-Art) 및 GPT-4o(Base Model) 대비 우수한 성능을 보임
- 특히 PubMedQA 데이터셋에서 MedBioRAG는 85.0%의 정확도로 기존 모델(BioGPT, Med-PaLM 2, Meditron-70B 등)보다 최신 최고 성능을 달성했음을 보여줌



## 04 AI 모델 개발

ChatBot : MedSurgGPT



# 04 AI 모델 개발 (4) : MedSurgGPT

## 모델의 목표

- 수술의 가이드라인과 환자에게 설명 해줘야 할 수술의 생존율 등 관련정보 서칭을 위한 챗봇
- 신뢰할 수 있는 수술 가이드라인을 데이터 소스로 하는 RAG챗봇 구성
- Azure AI Search + Azure OpenAI GPT-4o

The screenshot shows a conversational interface with the following elements:

- Header:** MedSurgGPT > ☺ :
- Text Box (Top Left):** 유방암 치료 가이드라인 알려줘
- Text Box (Main Area):** 유방암 치료에 대한 NCCN 가이드라인은 다음과 같습니다:
  - 전신 치료 (Systemic Therapy):
    - 재발성 또는 4기 유방암(stage IV)의 전신 치료는 생존 기간을 연장하고 삶의 질(QOL)을 향상시키지만, 원치 목적은 아닙니다. 따라서 독성이 적은 치료법이 선호됩니다. 가능한 경우, 독성이 적은 내분비 요법(endocrine therapy)이 세포독성 요법(cytotoxic therapy)보다 우선적으로 고려됩니다 [doc1].
    - 뼈 전이(bone metastases)가 있는 경우, 통증 및 기능 저하와 같은 합병증을 관리하기 위한 치료가 필요합니다 [doc1].
  - 수술 및 재건:
    - 유방 보방 수술(lumpectomy) 후 미용적 결과에 불만족하는 환자는 성형외과 상담을 통해 유방 결손 복구를 고려할 수 있습니다.
    - 또한, 대칭성을 유지하기 위해 반대쪽 유방에 대한 수술이 병행될 수 있습니다 [doc2].
    - 유방 재건에 대한 결정은 종양 치료를 우선으로 하며, 치료가 손상되지 않도록 해야 합니다 [doc2].
  - 진단 및 영상 검사:
    - MRI, 초음파, CT 스캔 등을 통해 원발성 유방암 여부 및 전이 여부를 평가합니다. 특히, MRI 양성인 경우 초음파 또는 MRI 유도 생검을 통해 추가 평가를 진행합니다 [doc3]....
- Text Box (Bottom Left):** 내 환자는 ER-/PR-/HER2+ 인데 최적의 치료 옵션이 뭘까?
- Text Box (Bottom Right):** 이 환자의 표적치료가 가능한 경우와 불가능한 경우를 알려줘
- Input Field:** Write your message
- Buttons:** Microphone icon, Send icon

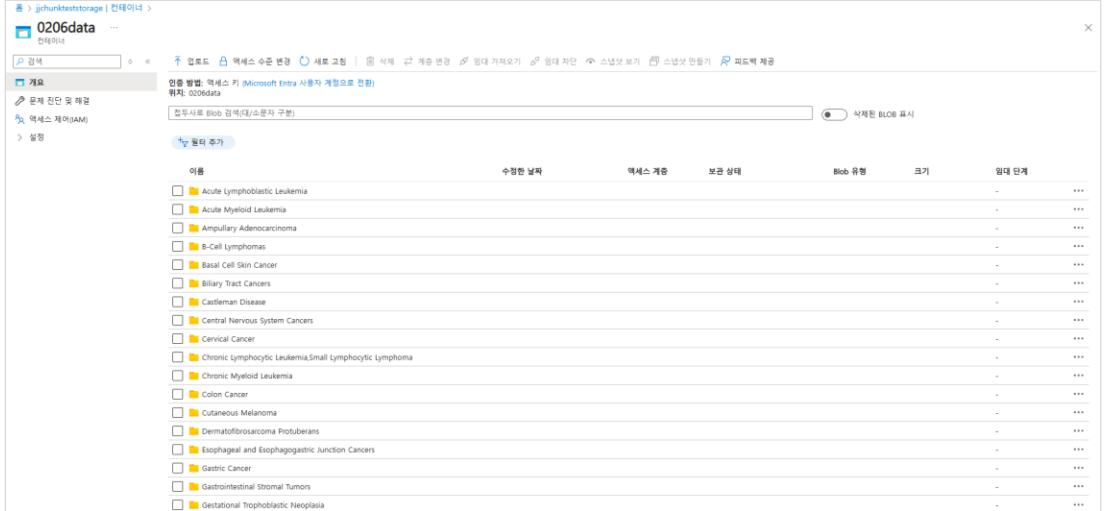
## 04 AI 모델 개발 (4) : MedSurgGPT

### 데이터 수집 및 전처리

- 공신력 있는 데이터 소스: NCCN, NCI, PubMed Central (PMC) 논문 등을 통해 수술 가이드라인 PDF 수집, 총 66종 질병
- 데이터 정제: 헤더와 푸터를 제외한 영역의 텍스트만 추출하여 불필요한 텍스트와 반복문자 제거, 그래도 남아있는 불필요 문자는 정규식으로 제거
- chunk 생성: 검색 성능 향상 및 맥락 유지를 위해 1000자 단위로 100자 겹침을 두고 chunk를 생성
- json 파일로 저장: 질병명, chunk ID, 데이터 출처를 포함, 신뢰도 확보 위함

```
1 {  
2   "chunk_id": "chunk_25",  
3   "chunk": "patients in the first decade after their diagnosis.[  
Effects of Retinoblastoma Therapy section. Trilateral retinob  
well-recognized syndrome that occurs in 5% to 15% of patients  
the development of an asynchronous intracranial mid line neur  
the ages of 20 and 36 months.[31 ] Because of the poor progn  
early detection and aggressive treatment of trilateral retino  
could potentially detect most cases within 2 years of the firs  
baseline brain MRI is recommended at diagnosis because it may  
subclinical stage. In a small series, the 5-year overall survi  
"disease": "Retinoblastoma",  
"source": "NCI-PDQ Treatment Information"  
5 }
```

[Json파일 데이터 구조]



The screenshot shows the Azure Blob Storage interface with a list of cancer types. The table has columns for Name, Last Modified, Access Tier, Blob Type, Size, and Block Count. The data includes various types of cancer such as Acute Lymphoblastic Leukemia, Acute Myeloid Leukemia, Ampullary Adenocarcinoma, B-Cell Lymphomas, Basal Cell Skin Cancer, Biliary Tract Cancer, Castleman Disease, Central Nervous System Cancers, Cervical Cancer, Chronic Lymphocytic Leukemia/Small Lymphocytic Lymphoma, Chronic Myeloid Leukemia, Colon Cancer, Cutaneous Melanoma, Dermatofibrosarcoma Protuberans, Esophageal and Esophagogastric Junction Cancers, Gastric Cancer, Gastrointestinal Stromal Tumors, and Gestational Trophoblastic Neoplasia.

이름	수정한 날짜	액세스 계층	보관 상태	Blob 유형	크기	임대 단계
Acute Lymphoblastic Leukemia					...	...
Acute Myeloid Leukemia					...	...
Ampullary Adenocarcinoma					...	...
B-Cell Lymphomas					...	...
Basal Cell Skin Cancer					...	...
Biliary Tract Cancer					...	...
Castleman Disease					...	...
Central Nervous System Cancers					...	...
Cervical Cancer					...	...
Chronic Lymphocytic Leukemia/Small Lymphocytic Lymphoma					...	...
Chronic Myeloid Leukemia					...	...
Colon Cancer					...	...
Cutaneous Melanoma					...	...
Dermatofibrosarcoma Protuberans					...	...
Esophageal and Esophagogastric Junction Cancers					...	...
Gastric Cancer					...	...
Gastrointestinal Stromal Tumors					...	...
Gestational Trophoblastic Neoplasia					...	...

[Azure Blob Storage]

## 04 AI 모델 개발 (4) : MedSurgGPT

### 인덱싱 및 챗봇 구성

#### 벡터화 및 인덱스 생성

- 데이터 벡터화: text-embedding-ada-002모델
- 인덱스와 인덱서 생성, semantic search 설정
- 인덱스에 disease 필드 추가: 질병별 구분

#### 챗봇 구성 및 개선

- 할루시네이션 최소화: strict=5, 온도 = 0.1
- 참조영역 표시: 원본 데이터와 데이터 출처 출력을 통해 신뢰성 확보
- 이어서 질문할 수 있도록 5개까지 답변을 기억하도록 설정
- 검색성능 향상 및 다국어 지원: 프롬프트를 영어로 번역하여 GPT모델로 요청한 뒤 질문에서 탐지된 언어로 대답

▼ Doc1

**Original Text**

engagement in the decision is encouraged. Guideline Stratification for Systemic Therapy for Stage IV/Recurrent Disease The systemic treatment of breast cancer recurrence or stage IV disease prolongs survival and enhances QOL but is not curative. Therefore, treatments associated with minimal toxicity are preferred. Thus, the use of the minimally toxic endocrine therapies is preferred to the use of cytotoxic therapy whenever reasonable.<sup>479</sup> Guidance for treatment of patients with breast cancer and brain metastases is included the NCCN Guidelines for Central Nervous System Cancers. Patients with recurrent or stage IV breast cancer at diagnosis are initially stratified according to whether bone metastases are present. These two patient subsets (those with and without bony metastases) are then further stratified by tumor HR and HER2 status. Therapy for Bone Metastases Complications from bone metastases include pain, decreased

**Data Sources**

disease: breast cancer, source: NCCN Guidelines

▼ Doc2

**Original Text**

stage IV disease whenever possible. FDG-PET/CT is a category 2B recommendation. The consensus of the Panel is that FDG-PET/CT can be performed at the same time as diagnostic CT. If FDG-PET and diagnostic CT are performed and both clearly indicate bone metastases, bone scan or sodium fluoride PET/CT may not be needed. Genetic counseling is recommended if the patient is considered to be at high risk for hereditary breast cancer as defined by the NCCN Guidelines for Genetic/Familial High-Risk Assessment: Breast, Ovarian and Pancreatic. Treatment The treatment of patients with IBC should involve a combined modality approach<sup>702</sup> comprising preoperative systemic therapy followed by surgery (mastectomy) and radiotherapy. Preoperative Chemotherapy There are no large randomized trials evaluating the optimal systemic treatment of IBC, since it is a rare disease. The systemic therapy recommendations are based on data from retrospective analyses, small

**Data Sources**

disease: breast cancer, source: NCCN Guidelines

[답변의 참조 영역 표시]

## 04 AI 모델 개발 (4) : MedSurgGPT

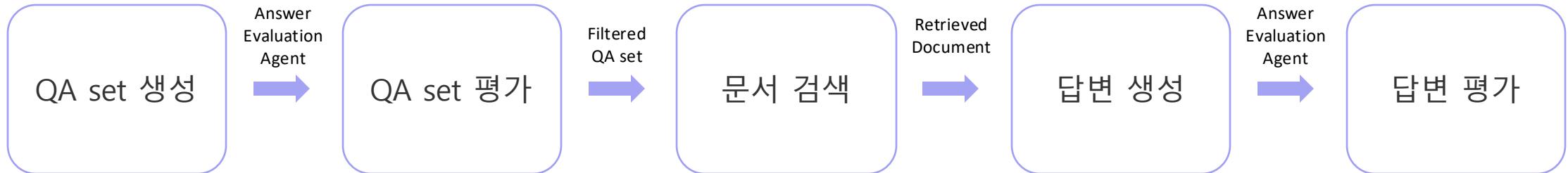
### 생성된 답변과 참조 데이터 매칭 테스트

- 데이터 전처리를 통해 사용자가 편하게 읽을 수 있도록 처리.
- 100회 이상 답변-참조 매칭 테스트를 통해 성능 검증.

The screenshot shows a user interface for a medical AI system. On the left, there's a 'Chat history' section with a message from the user asking about breast cancer treatment guidelines. Below it, a sidebar titled 'breast cancer 치료 가이드라인' lists three sections: 1. HER2-양성 유방암 치료, 2. 재발성 또는 4기 유방암 치료, and 3. 염증성 유방암(IBC) 치료. The first two sections are highlighted with red boxes, while the third is highlighted with a purple box. To the right, a vertical sidebar displays a list of documents: Doc1, Doc2, Doc3, Doc4, and Doc5. Document Doc3 is highlighted with a red box, and Doc4 is highlighted with a purple box. The content of Doc3 discusses treatment options for HER2-positive metastatic breast cancer, mentioning trastuzumab and additional lines of therapy. Doc4 discusses systemic treatment for breast cancer recurrence or stage IV disease, noting its palliative nature and preference for minimally toxic therapies like endocrine therapies over cytotoxic ones. A 'Submit' button is located at the bottom right of the interface.

## 04 AI 모델 개발 (4) : MedSurgGPT

### 성능비교 및 평가



- 모든 생성모델과 평가 에이전트는 Azure OpenAI GPT모델을 이용해 구성
- 평가에 사용된 LLM에는 평가지표를 제공하여 구조화된 평가를 수행할 수 있도록 함(1~5점 사이 점수 부여).
- 각 LLM에 최종 점수를 부여하기 전 근거를 출력하도록 하여 판단을 공식화·정교화 하여 평가의 신뢰도 높임.

rubric representing an evaluation criteria are given.  
1. Write a detailed feedback that assesses the quality of the response strictly based on the given score rubric.  
2. After writing feedback, write a score that is an integer between 1 and 5.  
3. The output format should look as follows: \"Feedback: {[write a feedback for criteria]} [RESULT] {[an integer number between 1 and 5]}\"  
4. Please do not generate any other opening, closing, and explanations. Be sure to include [RESULT] in your output.

```
##The instruction to evaluate:  
{instruction}  
  
##Response to evaluate:  
{response}  
  
##Reference Answer (Score 5):  
{reference_answer}  
  
##Score Rubrics:  
[Is the response correct, accurate, and factual based on the reference answer?] Score 1: The response is completely incorrect, inaccurate, and/or not factual. Score 2: The response is mostly incorrect, inaccurate, and/or not factual. Score 3: The response is somewhat correct, accurate, and/or factual. Score 4: The response is mostly correct, accurate, and factual. Score 5: The response is completely correct, accurate, and factual.
```

[Answer Evaluation Agent 평가지표]

## 04 AI 모델 개발 (4) : MedSurgGPT

### 성능비교 및 평가

평가항목	
generation	응답을 생성한 질문 수
	Accuracy(정확도)
	3점 이하 답변 수(답변의 품질)
	5점 답변 수(답변의 품질)
Retrieval	Recall@5*
	Faithfulness(충실성)**

- 기존의 벤치마크 데이터셋 (Pubmed QA, Med QA 등) 과 일반적인 평가 지표(BLEU, ROUGE 등)를 사용하지 않은 이유
  - RAG 모델의 핵심: 외부 지식을 기반으로 정확한 답변을 생성하는 것
  - 기존 벤치마크셋은 제공된 데이터셋이 아닌, 일반적인 정보 기반으로 구성
  - 제공된 데이터가 없을 경우, RAG 모델은 임의의 답변을 생성하지 않도록 설계됨.
- 모델이 제공된 데이터 내에서 얼마나 적절한 문맥(context)을 검색(Recall@5)하고, 검색된 정보로부터 충실한 답변을 생성하는지(RAGAS의 Faithfulness 지표), 생성된 답변이 의도한 정답과 얼마나 일치하는지(정확도) 테스트

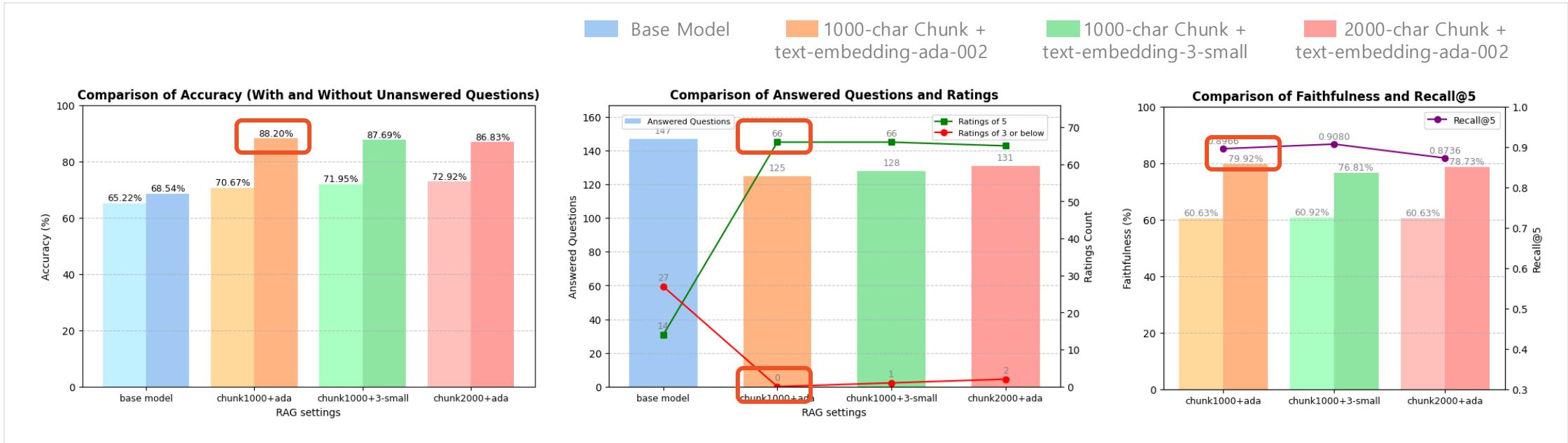
\*Recall@5: 정답 context가 검색해온 5개의 문서 안에 있는지 판단

\*\*Faithfulness(충실성) : 응답이 근거(retrieved context)에서 벗어나지 않고, 정확하게 정보를 반영하고 있는지 판단

$$\text{Faithfulness score} = \frac{\text{Number of claims in the generated answer that can be inferred from given context}}{\text{Total number of claims in the generated answer}}$$

# 04 AI 모델 개발 (4) : MedSurgGPT

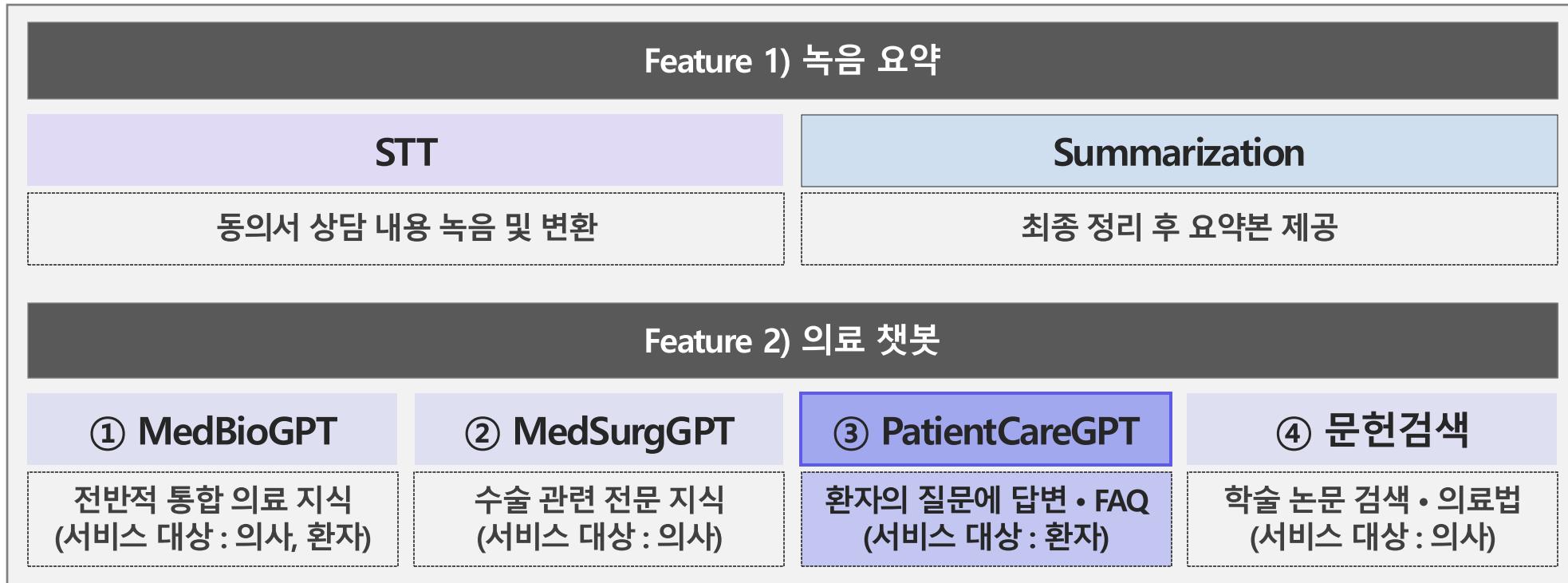
## 평가 결과



- 답변의 정확도와 응답의 품질면에서 베이스 모델보다 **RAG 모델의 점수가 높았음**.
- RAG모델의 목적에 비추어 응답하지 못한 답변을 제외, 대답을 한 경우의 점수에 주목(각 그래프의 오른쪽 바)
- 베이스모델: 응답을 많이 했지만 품질이 낮은 답변을 생성한 경우가 많았음(5점을 받은 답변이 적고 3점 이하 답변 수 많음)
- chunk 1000자 분할, text-embedding-ada-002 임베딩 모델조합**: 답변의 정확도, 대답의 품질(5점을 받은 답변 수 최다, 3점 이하답변수 0), 검색성능 모두 좋은 결과를 보임

## 04 AI 모델 개발

ChatBot : PatientCareGPT



## 04 AI 모델 개발 (5) : PatientCareGPT

### 모델 목적 및 구현 아키텍처

PatientCareGPT는 환자에게 수술 및 의료 법률 정보를 제공하기 위해 RAG(Azure AI Search) 및 Fine-tuning을 기반으로 개발된 챗봇

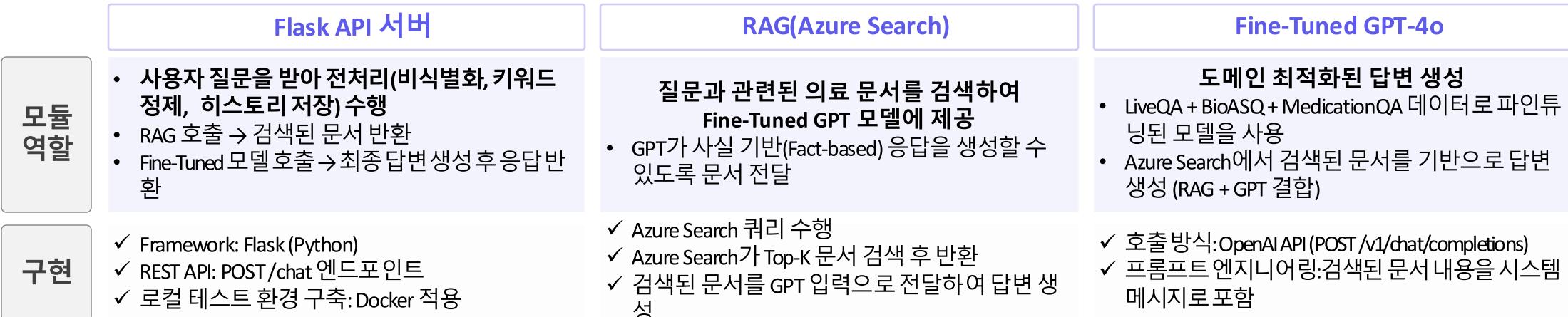
#### 목적

- 환자가 쉽게 이해할 수 있도록 의료 전문 용어를 풀어서 설명
- 수술 정보, 부작용, 회복 과정, 약물 정보 등 다양한 의료 주제에 대해 답변
- 수술 동의서 및 의료 법률 관련 문의에 대해 신뢰성 있는 정보를 제공
- 의료진의 업무 부담 경감 : 환자의 이해도 기반으로 반복적인 설명 시간 절약

#### 질문 예시

- (수술 정보, 후유증, 합병증 등)
  - ✓ 복강경담낭절제술에서 투관침 삽입에 따른 합병증이 무엇이죠?
  - ✓ 녹내장 시술 후의 후유증이 무엇인가요?
- (수술 관련 법률)
  - ✓ 환자에 관한 기록의 열람은 누가 할 수 있나요?

User → Flask API 서버 → RAG(Azure Search) → Fine-Tuned GPT → Response



## 04 AI 모델 개발 (5) : PatientCareGPT

### 데이터셋 구성 및 전처리 (Data Pipeline & Processing)

	수집 (Collection)	크기 (Size)	전처리 (Processing)
RAG	업데이트 필요 수술 동의서 의료법 관련 문서	PDF 193건 • 총 인덱싱된 문서 수 : 474개 • 전체 인덱스 크기 : 12.04MB	PDF→텍스트 추출→Embedding→Azure Search 인덱스 • 모델: OpenAI의 text-embedding-ada-002 • 기능: 각 텍스트 청크를 1536차원의 벡터로 변환 • 저장/인덱싱: 벡터 인덱스를 Azure Cognitive Search에 저장 (Semantic search 기능과 함께 사용)
Fine-Tuning	데이터 학습 완료 LiveQA (환자, 소비자건강) BioASQ (생의학, Pubmed) MedicationQA (약물, 복용법)	Train/Validation Split • Train: 6100개 • Validation: 163개	JSON 파일 → VS Code/스크립트 변환 → .jsonl 파일 OpenAI Fine-tuning 형식 • 형식: OpenAI 파인튜닝용 .jsonl(JSON Lines) ◦ 각 라인이 { "role": "...", "content": "..." } 구조를 갖춤 • 변환 도구/방법: ◦ VS Code+ Python 스크립트로 JSON에서 JSONL로 변환

#### 시스템 메세지

```
{"role": "system",
"content": "당신은 환자를 위한 의료 챗봇입니다.
사용자의 질문에 대해 RAG 검색된 문서를 반드시 활용하여 답변을 생성하세요."},  

{"role": "system",
"content": f"다음은 검색된 문서들입니다.
\n{retrieved_docs}\n\n사용자 질문: {prompt}"}
```

#### 프롬프트

```
{"role": "user", "content": f"{prompt}\n\n친절한 말투로 설명해 주세요."}
```

## 04 AI 모델 개발 (5) : PatientCareGPT

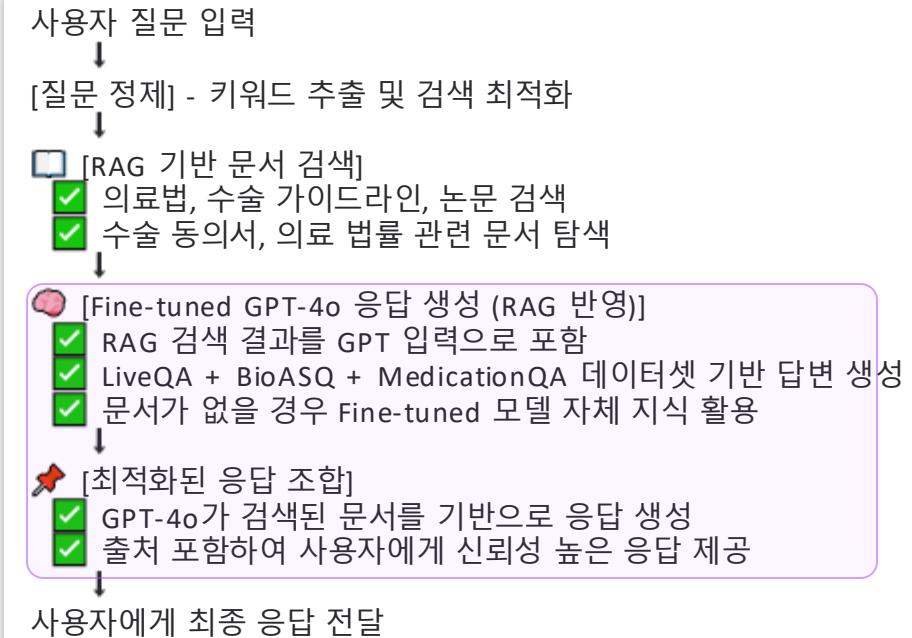
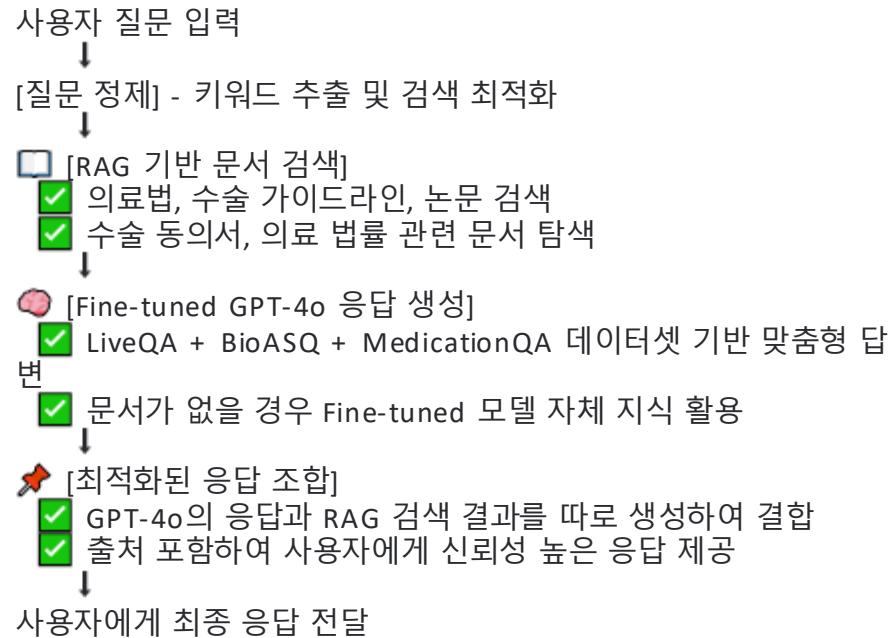
결합 방식: RAG + Fine-tuned GPT 모델

하이브리드 구조 : 최신 정보(RAG) + 도메인 전문성(Fine-Tuned GPT)

RAG와 Fine-tuned GPT-4o와 독립적으로 동작

RAG 검색 결과를 Fine-tuned GPT-4o의 입력에 포함

최종 파이프라  
인  
다이어그램

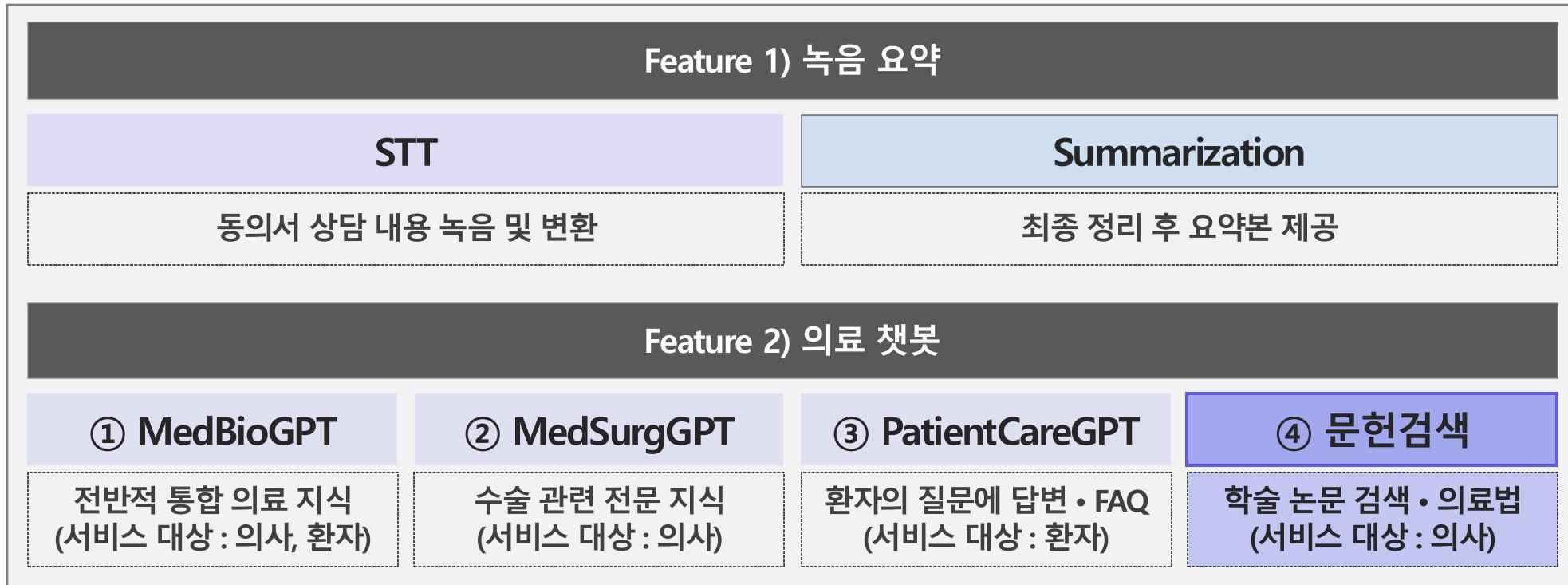


Fine-tuned GPT-4o가 RAG의 검색 결과를 입력으로 받아 답변을 생성하기 때문에 검색된 문서를 반영하여 더 신뢰도 높은 응답 제공

→ RAG 검색 결과와 모델의 응답을 별도로 결합하는 것이 아니기 때문에 일관성 있는 답변 가능

## 04 AI 모델 개발

ChatBot : 문헌검색



## 04 AI 모델 개발 (6) : 문헌 검색

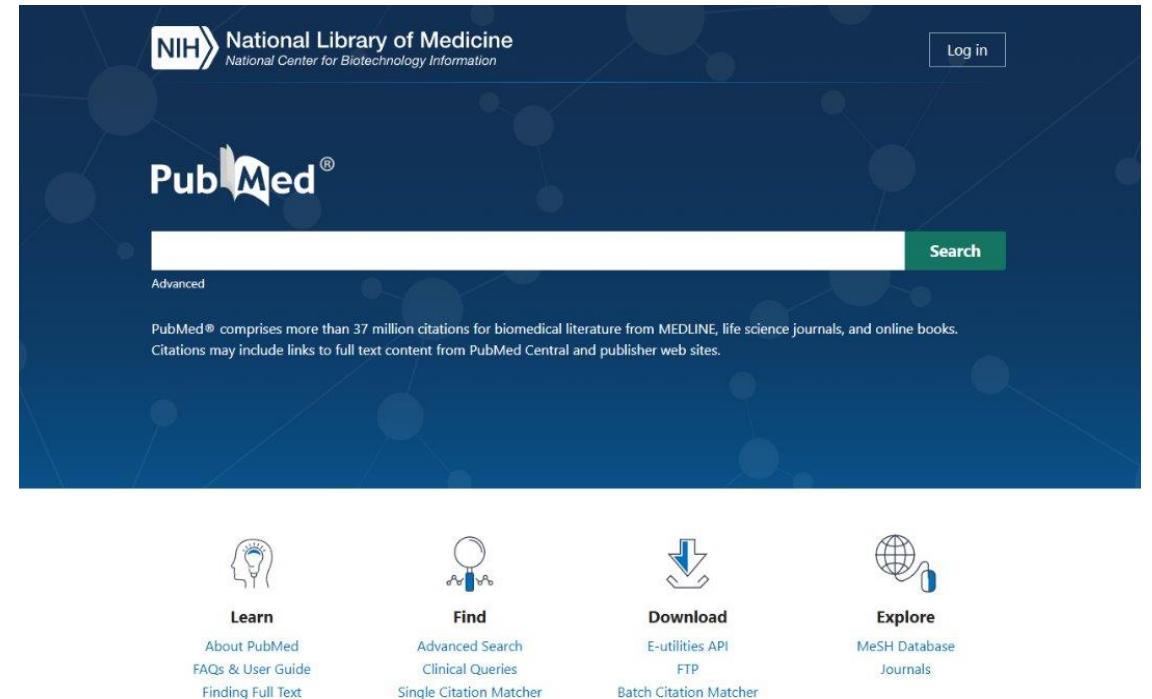
### 모델의 목적

문헌 검색 모델은 의료진이 원하는 논문을 제공해야 함

이를 위해 의학 논문 검색 데이터베이스 PubMed를 이용

의약학 분야의 연구자들은 논문을 검색할 때, 어떤 고전 논문  
이든, 최신 논문이든 가리지 않고 일단 PubMed에 접속함

기관의 논문은 내용을 확인하려면 돈이 드는 경우도 있지  
만 PubMed의 PMC(PubMed Central)에서는 내용을 확인  
하는 것이 무료이고, 구글 학술 검색보다도 사용 편의성이  
뛰어난 장점이 있음



## 04 AI 모델 개발 (6) : 문헌 검색

### RAG 구현을 위한 PubMed API의 사용

Azure Search AI로 RAG를 구성하는 것은 계속해서 새로운 논문을 업데이트 해줘야 하지만 매번 업데이트를 하기는 어려움이 있음

그래서 PubMed에서 제공하는 API를 활용해서 관련성 있는 최신 논문을 검색하고 의료진에게 제공할 수 있게 만듬  
PubMed API로 논문의 정보를 받아오는 방식으로 Retrieval (문서 검색) 기능을 구현

E-utilities라는 PubMed의 API 도구가 있는데 여러 API 중에서 ESearch, ESummary를 사용



Azure OpenAI



#### □ [Sample Applications of the E-utilities](#)

Created: April 24, 2009; Last Update: November 1, 2017.

[Introduction](#)

[Basic Pipelines](#)

[ESearch – ESummary/EFetch](#)

[EPost – ESummary/EFetch](#)

[ELink – ESummary/Efetch](#)

[ESearch – ELink – ESummary/EFetch](#)

[EPost – ELink – ESummary/EFetch](#)

[EPost – ESearch](#)

[ELink – ESearch](#)

## 04 AI 모델 개발 (6) : 문헌 검색

PubMed API의 종류

### ESearch

검색어를 입력하면 해당하는 UID(고유 식별자) 리스트를 반환하는 API

예를 들어, 'AI in healthcare'라는 검색어로 논문을 찾고 그 논문의 고유 ID를 얻을 수 있음

- ['9516129', '11239235', '9113794', '10200071', '6850413']

### ESummary

입력한 UID에 해당하는 논문의 요약 정보를 가져오는 API

ESearch로 관련성 높은 순으로 5개의 논문의 UID를 가져온 후 ESummary로 각 논문의 정보를 가져옴

- Title, Authors, Pubdate(출판일), Journal, DOI(Digital Object Identifier)라는 논문의 주소를 추출

-----  
Title: The utilisation of biliary organoids for biomedical applications  
Authors: Lei Z, Yang Y, Xiang Y  
Publication Date: 2025 Jan 7  
Journal: Frontiers in Bioengineering and Biotechnology  
DOI (논문 확인하기): <https://doi.org/10.3389/fbioe.2024.1501829>



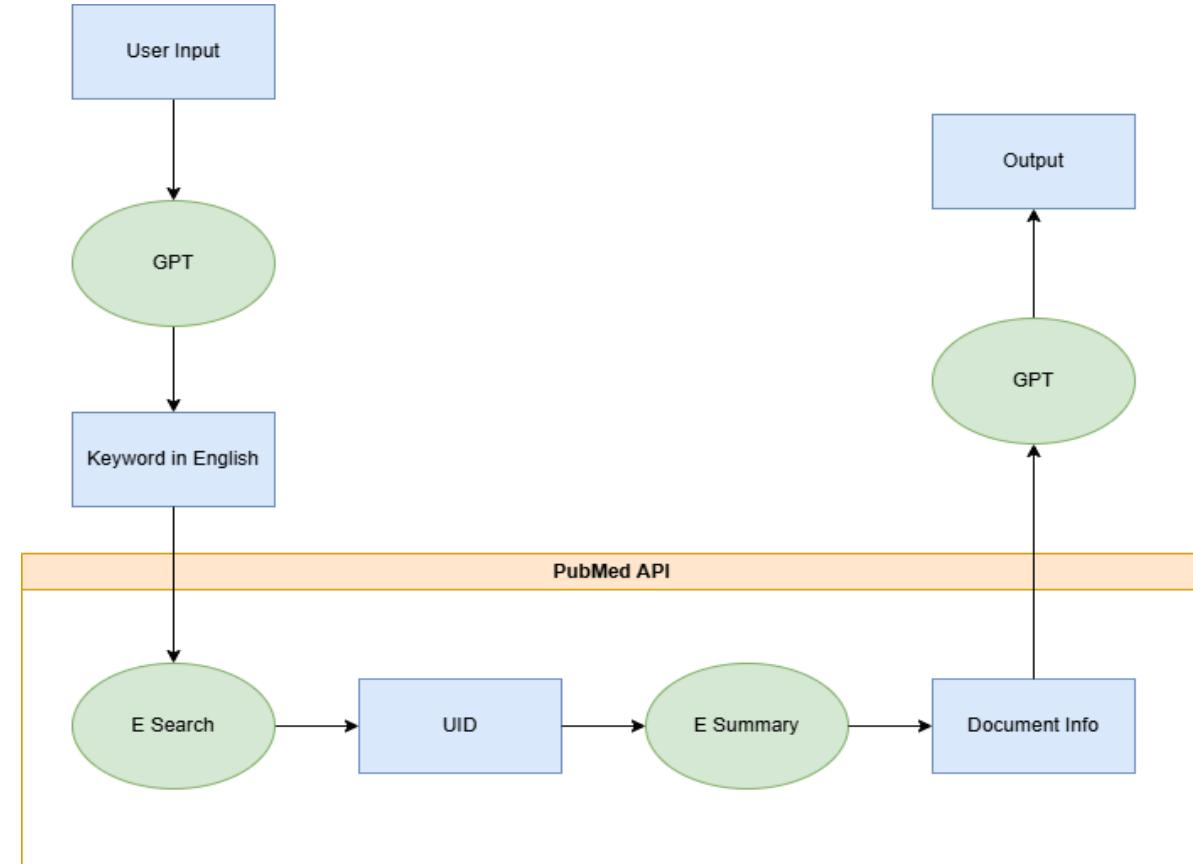
ESummary를 통해 나온 논문의 정보를 GPT로 가공해서 의료진에게 제공

'Frontiers in Psychology' 저널에 "인공지능이 대학생의 정신 건강 발전에 상당한 도움을 줌: 계량서적 분석"라는 제목의 논문이 2024년 3월 7일에 실렸어요. 저자는 Chen J와 Yuan D, Dong R, Cai J 등입니다. 자세한 내용은 <https://doi.org/10.3389/fpsyg.2024.1375294>에서 확인하세요

## 04 AI 모델 개발 (6) : 문헌 검색

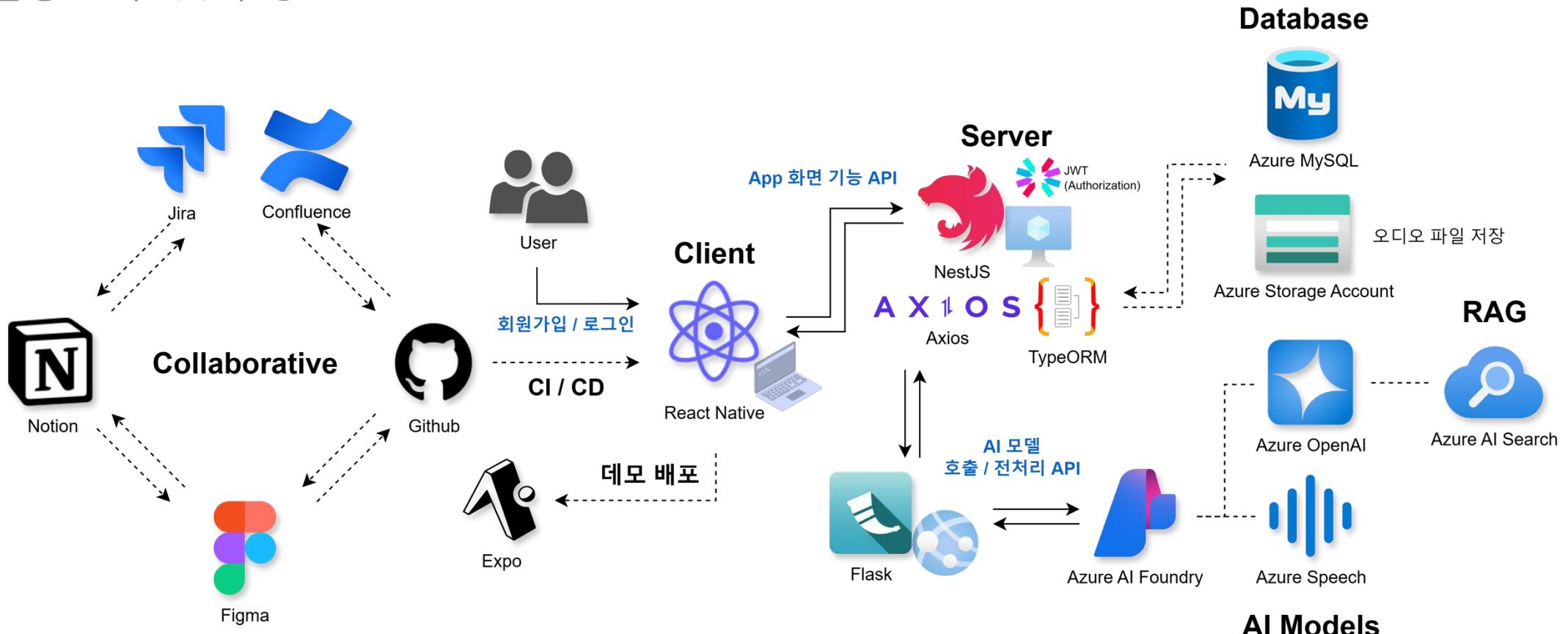
### 모델의 전체 작동 방식

1. API에 일관성 있게 입력을 하기 위해 의료진의 질문을 GPT로 키워드 추출 (영어)
2. 키워드를 PubMed API의 입력으로 제공
  - 1) ESearch로 논문의 UID 추출
  - 2) ESummary로 논문의 정보 추출
3. 논문의 정보를 다시 GPT로 가공하여 결과로 제공



## 05 APP 개발

### 활용 도구 및 구성



- React Native, NestJS : 코드 일관성, 생산성 등 이유로 같은 언어(JavaScript)에서 제공하는 프레임워크 사용
- Flask : AI 담당 팀원들의 코드 숙련도(친숙도) 고려

# 05 APP 개발

## 환경 구축 & DB 스키마



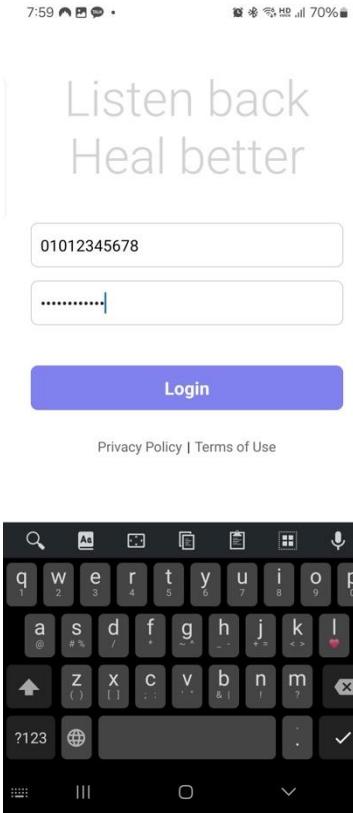
테스트 앱 설치

A terminal window titled 'TS BaseURL.ts M' showing the following TypeScript code:

```
src > constants > TS BaseURL.ts > ...
1 import { Platform } from "react-native";
2
3 const isPlatformAndroid = Platform.OS === "android";
4
5 // const ANDROID_URL = "https://medly-backend.eastus2.cloudapp.azure.com/";
6 const ANDROID_URL = "http://172.30.1.29:3000";
7 const IOS_URL = "http://192.168.35.150:3000/";
8
9 export const baseURL = isPlatformAndroid ? ANDROID_URL : IOS_URL;
10
```

Below the code, a QR code is displayed with the text 'Metro waiting on exp://192.168.0.12:8081'.

로컬 환경 실행



프론트엔드  
테스트 및 배포

## 05 APP 개발

### 환경 구축 & DB 스키마



Web Server  
(Reverse Proxy)



Node.js Version Manager



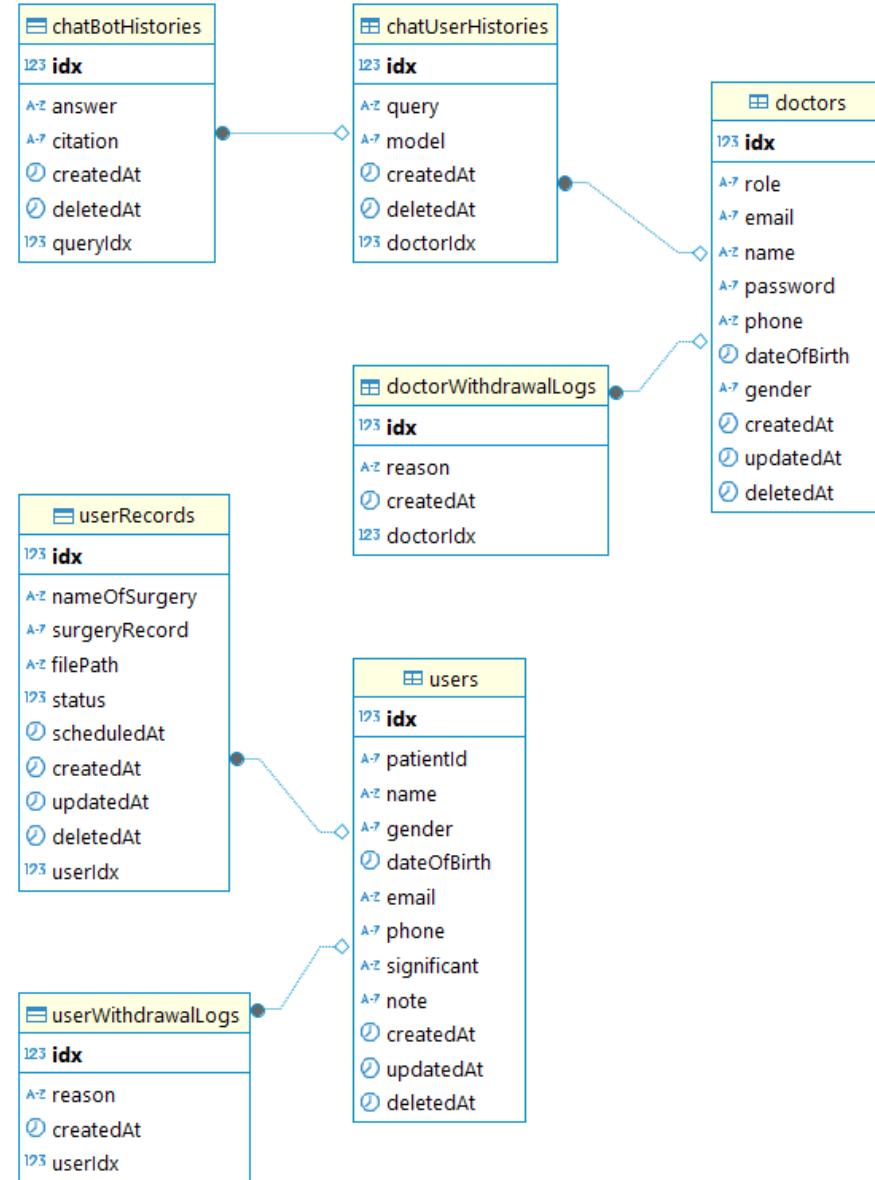
Runtime Environment



Package Manager

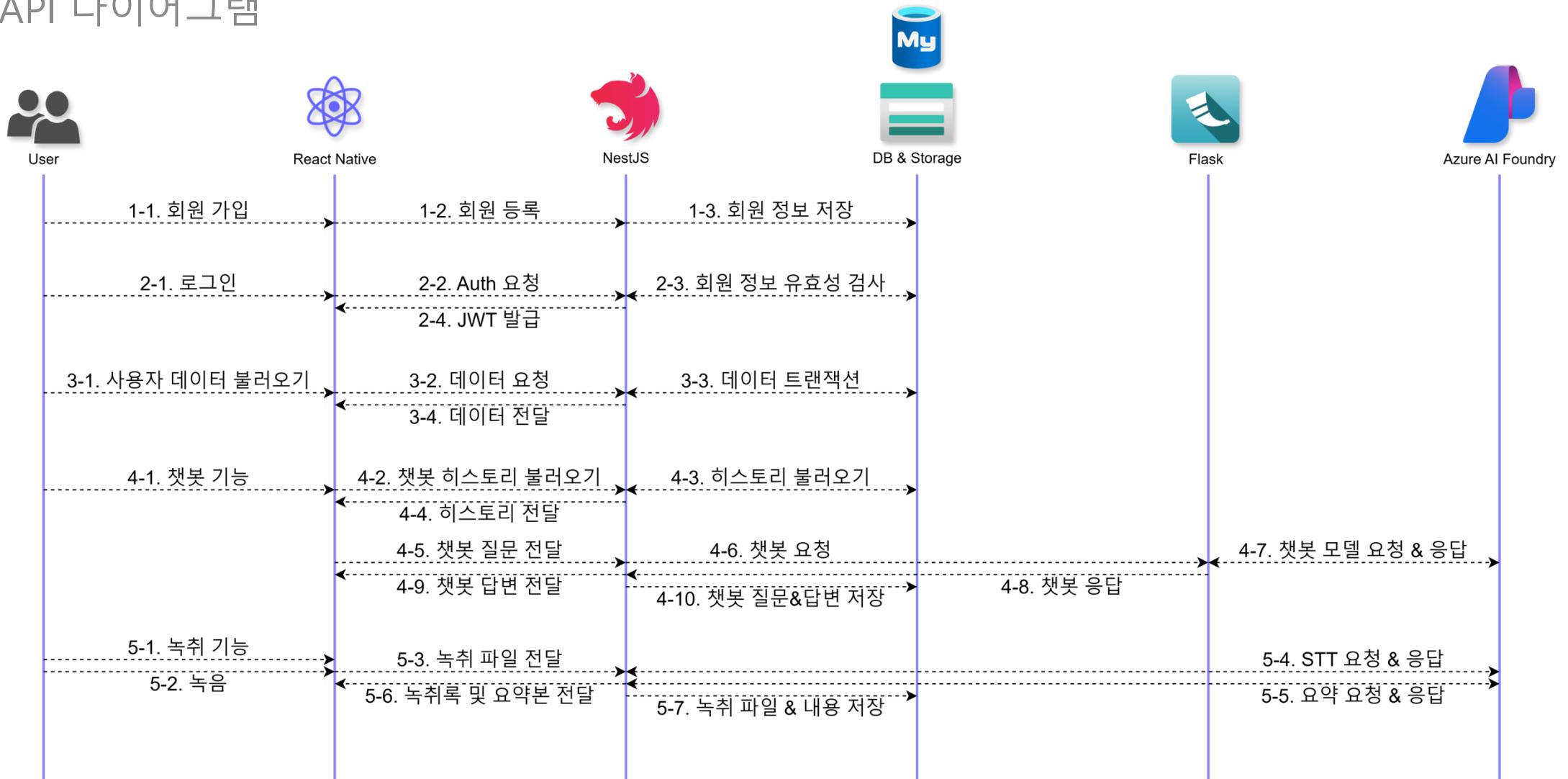
+

Linux Daemon (Prod)



# 05 APP 개발

## API 다이어그램



## 05 APP 개발

NestJS 배포 : Azure App Service 대신 VM을 사용한 이유?

- NestJS 서버를 통해 Database 연결 (Flask는 Database와 직접적으로 엔드포인트가 존재하지 않음)
- Azure App Service 환경 변수상에서 Azure MySQL 연동 문제 발생

```
graph TD; MySQL((MySQL)) --> Axios((Axios)); Axios --> NestJS((NestJS)); NestJS --> TypeORM[TypeORM]; TypeORM --> JWT[JWT (Authorization)];
```

홈 > 3rd\_prj\_team3\_eastus2 > 3rdprjteam3eastus2

[x] 3rdprjteam3eastus2 | 환경 변수 ☆ ...

검색 ◇ << 앱 설정 연결 문자열

개요 활동 로그 액세스 제어(IAM) 태그 문제 진단 및 해결 클라우드용 Microsoft Defender 이벤트(미리 보기) 권장 서비스(미리 보기) 배포 배포 슬롯 설정 원본 삭제

+ 추가 🔍 새로 고침 🔍 값 표시 🖊️ 고급 편집 🔍 끌어오기 참조 값

이름	값	배포 슬롯 설정	원본	삭제
APPLICATIONINSIGHTS_EXTENSION_VERSION	~	App Service		
ApplicationInsightsAgent_EXTENSION_VERSION	~			
AZURE_DB_SSL	-----BEGIN CERTIFICATE-----			
DB_DATABASE	-----END CERTIFICATE-----	값 표시		
DB_HOST	-----BEGIN CERTIFICATE-----	값 표시		
DB_PASSWORD	-----END CERTIFICATE-----	값 표시		
DB_PORT	-----BEGIN CERTIFICATE-----	값 표시		

| Invalid Certificate |

```
-----BEGIN CERTIFICATE-----\nMIIDzCCAp...-----END CERTIFICATE-----
```

# 06 인공지능 윤리

## 투명성

- 챗봇이 제공하는 정보의 출처를 명확히 표시함으로써 응답이 **생성된 근거 확인 가능**하도록 함
- STT 및 요약본 제공 과정에서 원본 녹음파일과 요약본을 제공함으로써 전체 내용 확인할 있도록 함

## 책임성

- 서비스가 제공하는 정보가 의료진의 최종 결정을 대체하지 않고 보조 역할을 한다는 점을 명확히 함.
- 정보의 **출처 제공**을 통해 올바른 출처를 통해 답변을 생성한 것인지 확인.
- 챗봇과 AI가 제공하는 정보에 대한 **피드백** 시스템 운영

## 신뢰성 및 안전성

- 의료 가이드라인 챗봇이 **최신 의료 지침을 반영**할 수 있도록 정기적인 업데이트 수행.
- 검색된 데이터가 없을 경우 임의로 답변을 생성하지 않고 검색된 데이터가 없다고 답변하도록 설정.
- 피드백 시스템 운영 및 모니터링을 통해 앱 운영 상황을 감시 및 대응

## 개인정보 보호 및 보안

- 환자의 **의료 정보 및 대화 데이터를 철저히 보호하고 암호화**하여 저장
- AI가 수집하는 데이터의 범위 고지 및 가입 시 사용자의 동의를 받는 절차 수행.
- 환자와 의사의 데이터 접근 권한을 명확히 구분하여 필요 최소한의 정보 제공.
- 환자 정보 (전화번호, 이메일 등) 입력될 시 비식별화하고 주의 안내 멘트 전달

### 5. 추적 관찰:

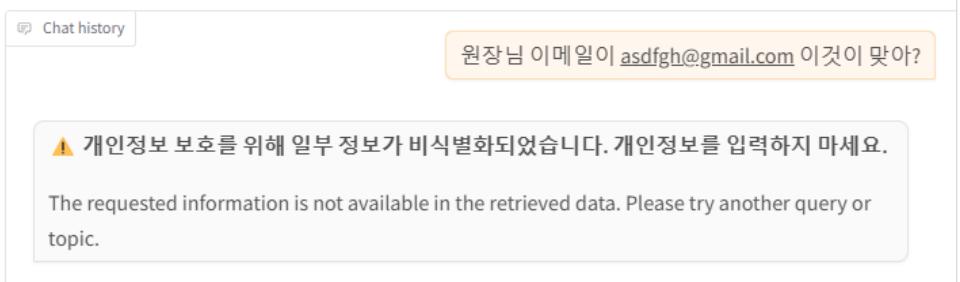
- 치료 후 정기적인 추적 관찰이 필요합니다. 5년 동안 4~6개월마다 병력 및 신체검사를 시행하며, 이후 매년 검사를 권장합니다. 또한, 매년 유방촬영술을 시행해야 합니다 [doc4].

치료 계획은 환자의 상태와 병기, 전신 건강 상태에 따라 조정되어야 하며, 다학제적 접근이 필요합니다.

[챗봇의 보조적 역할 명시]

"response": "The requested information is not found in the retrieved data. Please try another query or topic."

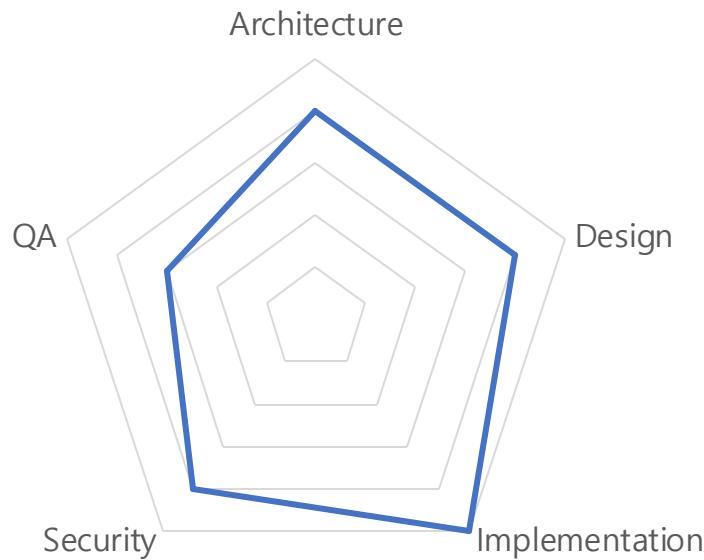
[검색된 데이터가 없는 경우 답변을 생성하지 않도록 설정]



[개인정보 입력 시 주의 안내 멘트 출력]

## 07 Outro

### 자체평가



#### 잘한 점

- 기획했던 대부분을 구현했음
- 분업을 잘 해서 원활하게 진행됨
- 서로 가지고 있는 지식을 나누면서 많이 배움
- 배운 내용에 국한되지 않고 적극적으로 사용 가능한 리소스를 발굴하여 사용함

#### 아쉬운 점

- 시간여유가 있어 QA를 좀 더 꼼꼼히 한 뒤 실제 앱스토어나 플레이스토어에 배포까지 진행 하였으면 좋았을 것
- 코드리뷰 시간이 있었으면 좋았을 것
- 시행착오 공유가 더 활발하게 이루어졌으면 좋았을 것

## 07 Outro

### 자체평가 : 서비스에 대한 개선점

- **STT-Summarization 모델**
  - 한글 STT 정확도 이슈: fine tuning, summarization 기능 통합으로 전문용어에 대한 인식을 개선했으나 여전히 오류 존재(치명적인 문제 발생을 방지하기 위해 기획상 의사의 최종 확인 후 환자측에 전송되도록 설정) - 데이터가 수집되고 학습 시키면서 개선될 것으로 기대
  - open AI gpt-4o 모델 prompting으로 최종 정리본을 출력하고 있는데, 같은 데이터라도 결과물의 형식이 약간씩 기대와 다른 경우가 있음, 향후 데이터를 쌓으면서 프롬프트 최적화 개선
  - 처리 속도 : 현재 22분의 대화를 최종 처리하기까지 약 4-5분 소요되는데, 바쁜 의료 현장에서 효율성을 높이기 위해 녹음과 STT를 실시간 처리와 컴퓨팅 파워 업그레이드 필요
- **MedSurgGPT 모델**
  - 현재 66종 질병에 대한 가이드라인 데이터가 들어가 있는데, 더 많은 질병을 다룰 수 있도록 확장
  - 추후에 가이드라인이 업데이트되면 자동으로 이를 반영할 수 있는 파이프라인 구성
- **MedBioGPT 모델**
  - 법적 책임 문제 대비: 의료 AI 모델의 특성상, 부정확한 답변이 환자 진료나 의사 결정에 영향을 미칠 경우 법적 책임 문제가 발생할 수 있음. 의료 전문가 검증 절차 및 휴먼 인더 루프(Human-in-the-loop) 시스템 도입, 모델이 불확실한 답변을 줄 경우 자동으로 신뢰할 수 있는 출처를 제시하는 기능 추가 등 개선 필요
- **전체 서비스**
  - **App 기능:** ChatGPT처럼 챗봇이 실시간으로 생성되는 것 같은 UX를 제공하기 위해 Websocket을 이용한 Real-time Chatbot Response Streaming 구현, 여러 개의 Chatbot 모델을 사용자에게는 단일 모델처럼 보이게 하고, 백엔드에서 Multi Model 개념 구현, Chatbot STT 기능 추가
  - **CI/CD 관련:** Test Environment 보완 및 구축 (ex. Devel Server, Staging Server 구축, Jest를 활용한 Test Code 작성)

## 07 Outro

### 자체평가 : 팀원별 소감

이름	소감
선우진성	장기간의 프로젝트 동안 각자의 결과물을 조합하여 하나의 실용적인 제품으로 완성해가는 과정이 보람차고 뜻깊었습니다.
이진형	프론트면 프론트, 백엔드면 백엔드, AI면 AI, 각자의 분야에 통달하신 분들이 계셔서 매끄럽게 진행된 프로젝트가 아니었나 싶습니다. 개인적으로는 처음 다뤄보는 도구가 많았고, 수박 겉핥기로만 백엔드를 알고 있었지만, 팀원분들 덕분에 한층 더 스텝 업 해서 좋은 결과물을 만들었다고 생각합니다. 모두에게 감사하다는 말씀을 드리고 싶습니다.
정지현	데이터 수집부터 정제, 서치와 결합하여 챗봇을 만들고 평가까지 모든 과정을 연구하고 적용해보는 과정이 정말 의미 있었습니다. 시간 여유가 있어 문제점을 파악하고 해결방안을 적용하며 발전시킬 수 있어서 즐거웠고, 유능한 팀원분들께 배우면서 더 성장한 것 같습니다.
김재빈	앱의 기획부터 구현까지 모든 과정을 함께 해낼 수 있어서 좋은 경험이었습니다. 모두가 최선을 다해서 해준만큼 좋은 결과가 나온 것 같아서 뿌듯 합니다. 기간이 충분하지 않아 앱스토어에 배포를 하지 못한 점은 아쉽지만 개발 과정에서 여러 많은 점을 배워서 너무 좋았습니다.
현수민	하나의 모델을 개발하는 과정을 경험할 수 있어 매우 의미 있는 프로젝트였고, 하는 과정에서 많은 시행착오가 존재했으나 팀원들 덕분에 어려움 속에서도 좌절하지 않고 해결하며 많은 것을 배울 수 있었습니다.
이현영	현실적으로 효용이 있을 서비스를 기획해서 개발해볼 수 있어서 의미 있었고, Azure의 다양한 툴을 사용해서 여러 가지 시도를 해볼 수 있어서 좋았습니다. 구현하는 과정에서 그 동안 배운 파이썬 코딩을 적절히 활용하는 부분이 특히 재미있었고 다른 팀원 분들께서 적극적으로 도와주시고 많이 배워서 감사했습니다.
김선옥	Azure 리소스를 활용하여 다양한 실험을 시도할 수 있었고, 원하는 만큼 연구를 진행할 수 있어 좋았습니다. 특히, Azure AI Search와 OpenAI Fine-tuning 기능을 다양한 세팅에서 실험하며 모델 성능을 최적화하는 과정에서 많은 것을 배울 수 있었습니다. 무엇보다도 유능한 팀원들과 크로스 러닝을 할 수 있어 더욱 의미 있는 시간이었습니다.
안의진	마지막 프로젝트를 진행하게 되면서 평소에 같이 일해보고 싶었던 분들이랑 함께 진행 할 수 있었던 것이 가장 좋았습니다. 다양한 분야에서 일하시던 분들과 협업하면서 많은 인사이트를 얻을 수 있었고, 원래 코딩을 접해보시지 않은 분들이 6개월만의 수업으로 대단한 결과물을 만드시는 것을 보고 이런 팀에서 마지막을 같이 할 수 있게 되어서 더 감사했습니다. 시간이 길었던 만큼 하나의 프러덕트의 제품을 만들어 각자 분야에서 하신 부분들을 시각적으로 볼 수 있었던 부분도 좋았습니다.

## 07 Outro

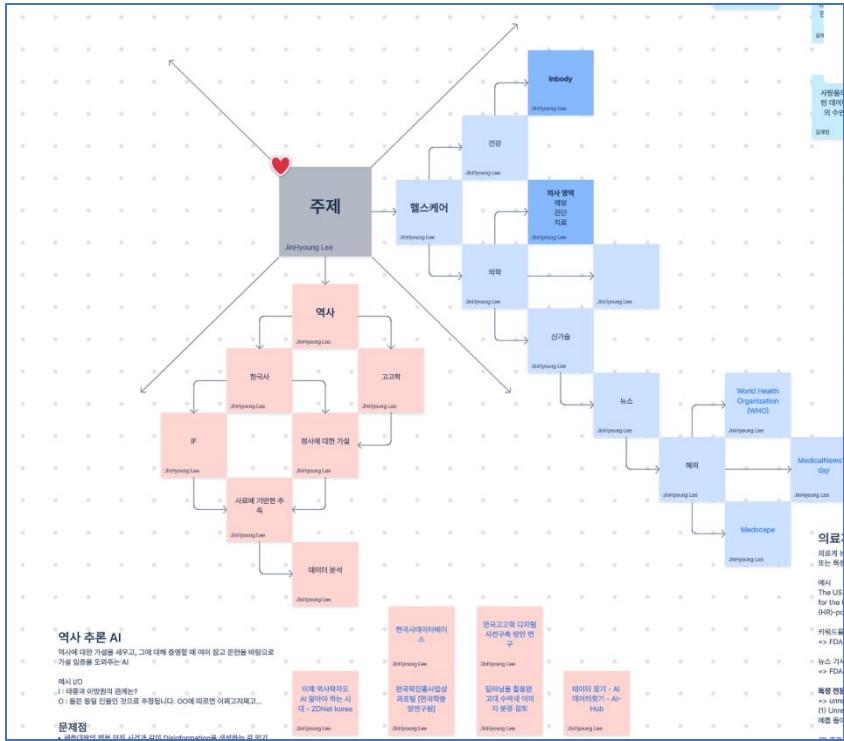
현장 시연

감사합니다.  
지금부터 Medly 시연을 시작하겠습니다.

# Appendix

## BrainStrom and Retro

### Brainstorming



### Retro

