
ASK 2024 논문집
Annual Symposium of KIPS 2024
www.kips.or.kr

신진학자 워크숍

# 의료 분야에서 대형 언어 모델의 활용

**도재영 교수**
(서울대학교)

# ASK2024 신진학자 워크숍

서울대 전기·정보공학부 도재영 교수

# 발표자 소개

| | |
|---|---|
| 2007.09 ~ 2012.12 | 학사: 한국 과학 기술원 (KAIST), 전산학 전공 |
| 2007.09 ~ 2012.12 | 석사, 박사: UW-Madison, Computer Sciences |
| 2012.12 ~ 2014.02 | Research Engineer: Microsoft Data Lab |
| 2014.03 ~ 2016.09 | Senior Scientist: Microsoft Jim Gray Systems Lab (GSL) |
| 2016.10 ~ 2021.02 | Senior Researcher: Microsoft Research (MSR) |
| 2021.02 ~ 2023.11 | Senior Applied Scientist (Manager): Amazon Alexa AI |
| 2024.02 ~ 현재 | 서울대학교 전기·정보공학부 조교수 |

# AIDAS (AI, Big Data, and System)

Microsoft

인공지능을 활용한 마이크로소프트 클라우드 빅 데이터 시스템 최적화 연구 및 개발

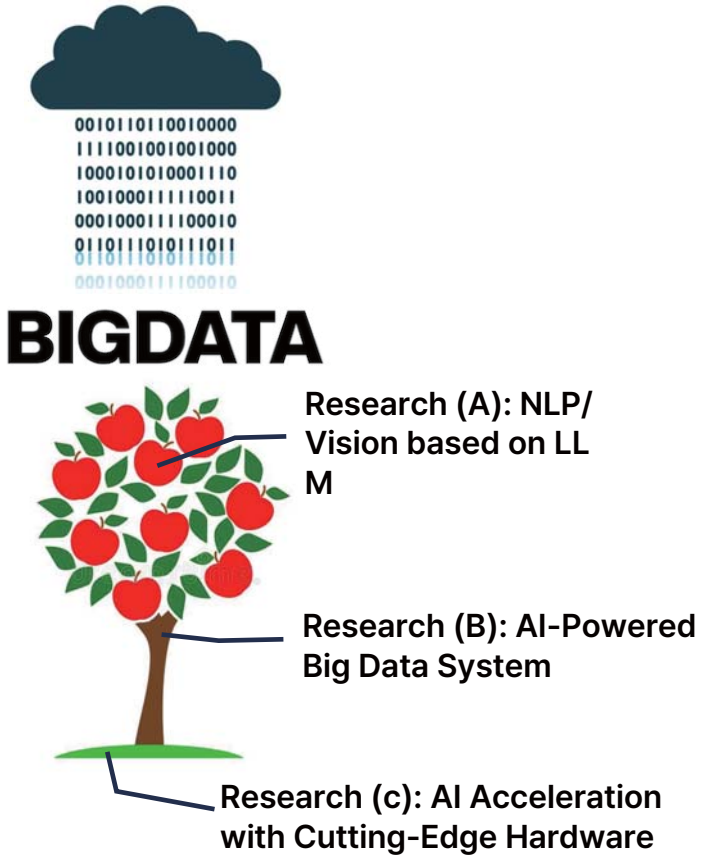
고객 개인화를 고려한 Alexa (아마존 스마트 스피커) 핵심 딥러닝 모델 (LLM) 연구 및 개발

amazon alexa

# 주 연구 분야

## 대용량 데이터 기반 인공지능 알고리즘, 그리고 시스템 최적화 연구

- [Research A] 대형 언어 모델 (LLM) 기반 생성형 AI. 대규모 딥러닝 훈련 및 추론
- [Research A] 멀티 모달 AI을 통한 자연어/비전 처리
- [Research B] 인공지능 활용 빅데이터 관리 시스템 개발
- [Research C] 차세대 메모리를 활용한 고성능 인공지능 데이터 분석 및 처리
- [Research A, B, C] ML/AI 응용을 위한 알고리즘/시스템 통합 설계




연구에 재미를 느끼고 싶은 열정 있는 학생을 찾고 있습니다.
관심 있으신 분은 연락 주세요!

jaeyoung.do@snu.ac.kr

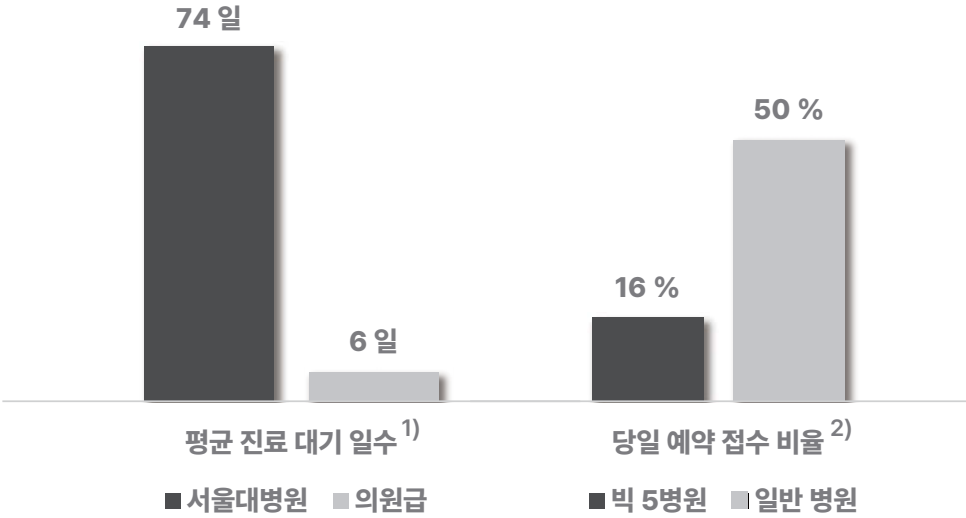
# 의료 분야에서 대형언어모델의 활용

# 누구나 최고의 의사를 만나고 싶어합니다

대학병원 가고 싶은데 너무 오래 기다려요

지방에서 가면 2박 3일은 잡아야 해요

집 근처에서는 빨리 해 주긴 하는데
치료받다가 문제생기면 어떻게 하죠?




## 환자들은 대학병원의 우수 의료진을 선호하는 현상이 뚜렷함

- 빅5 병원 급여 진료 수입 중 60%가 지방 환자 3)
- 빅5 병원 초진 환자 평균 1달 이상 대기 2)
- 국내 암 환자의 61.8%는 상급종합병원을 이용 3)

*1) 2021, 보건복지부, 의료서비스 경험 조사*
*2) 2020, 건강보험심사평가원, 의료이용 추이 모니터링 고도화 연구*
*3) 2014 ,지역별 의료이용 통계연보*
*4) 2020, 이용호 의원실 자료*

# 최고의 의사를 쉽게 만나는 것: ChatGPT가 해결할 수 있을까요?

**ChatGPT는 환자 증상에 대한 일반적인 정보를 제공하는 역할에 그침**

**You**

다음 증례를 보이는 환자에게 필요한 치료법을 알려줘.

"38세 된 남자환자로서 3일전 부터의 인후통, 기침, 가래, 호흡 곤란을 주소로 내원하였다. 환자는 12년 전부터 천식이 있었으며, 3일전부터 미열, 인후통이 생긴후 잦은 기침, 맑은 색 가래, 경한 호흡곤란이 같이 생겼다. 증세가 점점 심해지고 특히 밤에 호흡곤란이 심해져서 내원하였다. 환자는 내원 당시 열은 없었지만, 심한 기침과 그로 인한 인후통, 하루 소주잔 반잔 정동의 가래, 낮에는 생활에 지장이 없을 정도로 경하지만 밤에는 심해지는 호흡곤란을 호소 하였으나, 비강 폐쇄, 비루, 쉰 목소리, 객혈은 없었다. "

**ChatGPT**

환자의 급성 기관지염 치료법 요약:

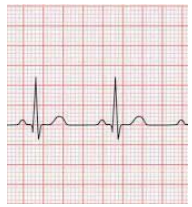
1. 기관지 확장제 및 스테로이드 치료: 호흡 곤란 및 염증을 완화하기 위해 처방될 수 있음.
2. 부종제 및 해열제: 염증 및 부종을 줄이기 위해 사용.
3. 천식 관리: 기존 천식 치료법을 조정하여 발작을 관리.
4. 안정과 휴식: 호흡 곤란 및 기침 시 충분한 휴식이 필요.
5. 주의사항: 의사의 지시에 따라 적절한 항생제 사용을 검토.

# 현대 의료의 문제점: 의료 데이터 증가량

- 환자들을 기반으로 생성되는 의료 데이터는 최근 5년간 **연평균 18%의 증가율**을 보임 [1)]
- 입원 환자의 데이터는 **1인당 하루에 평균 86,400개**의 수치가 생성
- 의료기관당 생성되는 전체 의료 데이터의 크기는 **하루당 약 370TB에 달함** [1)]
- 수집되는 헬스케어 데이터의 80%는 텍스트나 이미지, 영상과 같은 **비구조화 데이터**임
- 의사는 환자와의 상담 및 검진 대신 컴퓨터를 보는데 거의 두 배 더 많은 시간을 할애함 [2)]
- 의료인들은 많은 의료 데이터의 분석에서 한계를 가지며, 미국에서만 **매년 25만명 이상이 의료 과실로 사망**함 [3)]

1) 2018, 삼정KPMG 경제연구원, Issue Monitor 제94호;
2) 2016, American Hospital Association;
3) 2016, Makary et al. , BMJ 353 (2016)

# 의료 데이터의 종류

| 형식 | 의미 | 예시 |
|---|---|---|
| 구조화 데이터<br>(Structured Data) | 체계화되어 데이터베이스에서 쉽게 검색 가능한 데이터 | • 전자 건강 기록(EHR): 환자 인구 통계, 진단 코드(예: ICD-10), 약물 코드, 실험실 테스트 결과 및 청구 정보 등<br>• 실험실 결과: 혈액 검사, 소변 검사, 기타 바이오마커 분석 등<br>• 영상 메타데이터: 영상 매개변수, 환자 위치 지정, 표준 영상 프로토콜과 같은 정보를 포함하는 의료 영상 |
| 반구조화 데이터<br>(Semi-Structured Data) | 관계형 데이터베이스에 상주하지 않지만 일부 조직적 속성을 가지는 데이터 | • 유전자 주석 파일: 원시 유전 데이터(구조화)와 데이터 포인트를 설명하는 주석 또는 메타데이터(비구조화)를 모두 포함하는 파일<br>• 영상 촬영용 DICOM 파일: 이미지 자체는 구조화되어 있지 않지만 영상을 저장하는 DICOM 파일에는 환자, 연구 및 영상 매개 변수에 대한 구조화된 메타데이터가 포함되어 있음 |
| 시계열 데이터<br>(Time-Series Data) | 연속적인 시간 간격으로 수집되거나 기록된 일련의 데이터 포인트. 추세, 주기 또는 예측을 이해하기 위해 사용되는 데이터 | • 환자 모니터링 데이터: 시간 경과에 따라 활력징후(예: ECG, 혈압, 산소 포화도)를 모니터링하는 연속적 데이터<br>• 웨어러블 장치 데이터: 웨어러블 건강 장치에서 수집된 데이터로 활동, 심박수, 수면 패턴 등의 데이터  |
| **비구조화 데이터**<br>**(Unstructured Data)** | **사전 정의된 데이터 모델이 없는 데이터** | • **임상 노트**<br>: 관찰, 치료 계획, 환자 진행 상황을 포함하여 의사, 간호사 및 기타 의료 서비스 제공자가 제공하는 차트 데이터, 영상 판독문 |

**• 전체 의료 데이터의 80%가 비구조화 데이터 → 생성AI의 공략 포인트**

# 비구조화된 의료 데이터의 문제점

- PI : 일을 하다가 갑자기 우측 제5수지 신전이 되지 않아 내원함.
- PEx

  Rt 5th finger Ext[illegible] Flexion은 G5
- X-ray: DA DRUJ wrist Rt
- Imp

  #1. Rupture EDC 5th finger hand Rt

  #2. Severe DA DRUJ wrist Rt
- Plan

  Exam: Sono

  OP: EDC transfer, Sauve kapandji procedure 시행예정

  Extension lag 남을 수 있음을 설명함

  기타: 환자가 비용 문제로 고민 중 → 산재신청 후 결과에 따라 수술

각 과마다 각자의 약어가 너무 많음

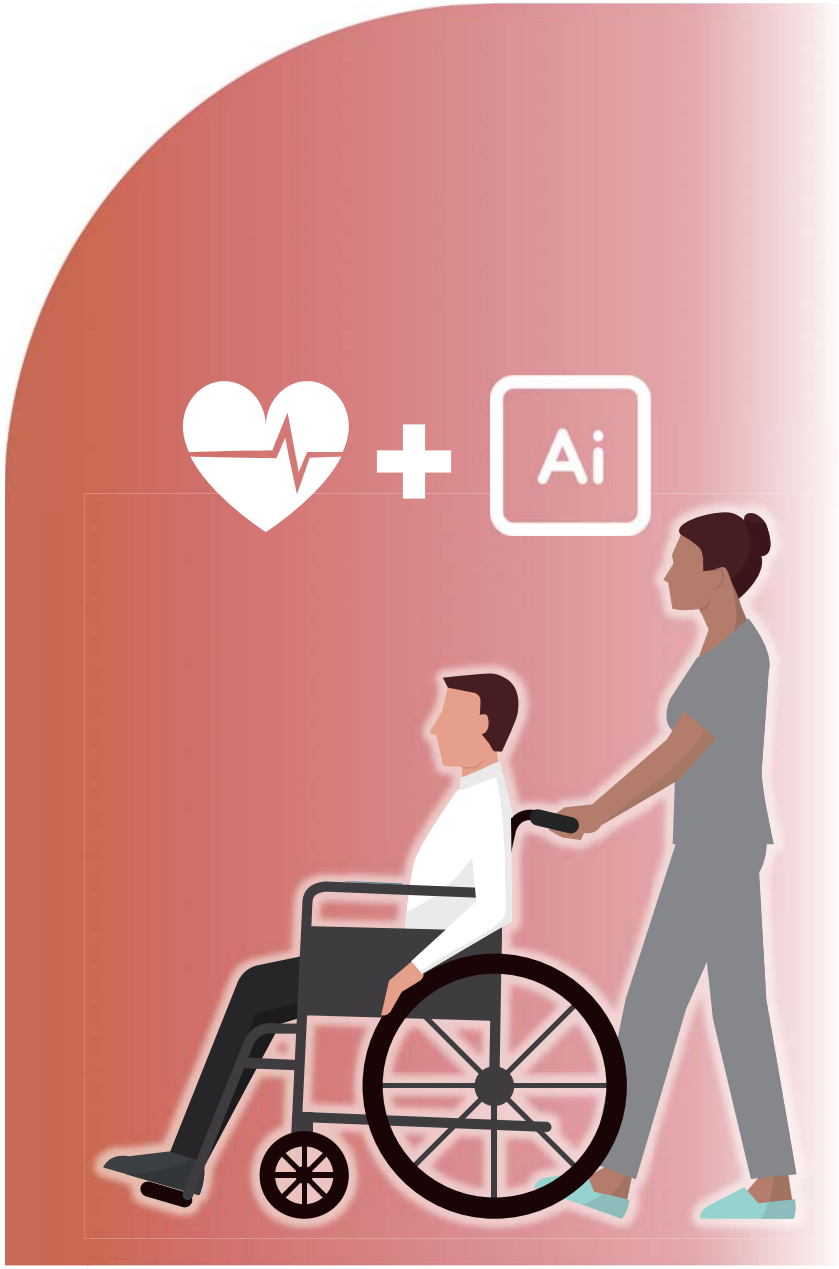
영어와 한글의 혼합

치료 방침에 환자의 개인적 상황이 반영

영상 판독이 어려움

# 의료 AI가 가져야 할 요건

1. 정확성과 신뢰성, 재현성
2. 증거기반 추론
3. 임상 검증
4. 의료윤리 준수
5. 데이터 보안 및 개인 정보 보호



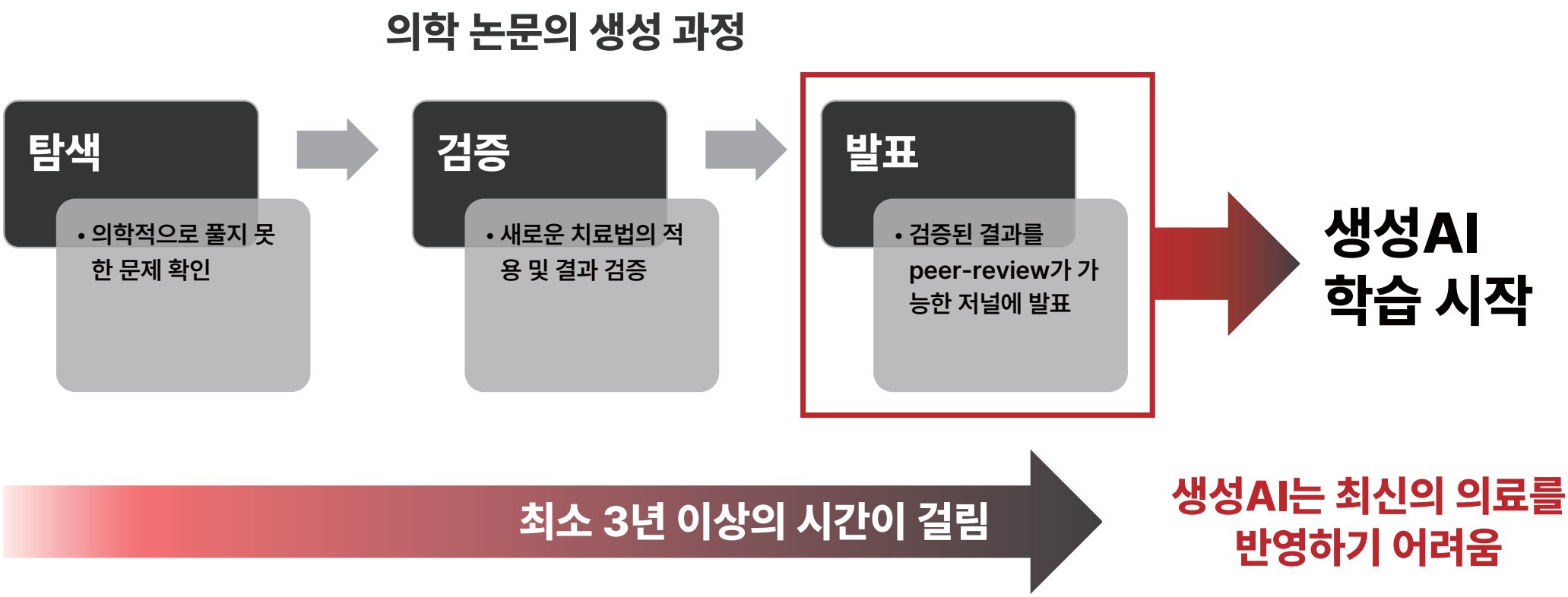

**의료 AI는 일반 AI에 비하여 까다로움**

# 생성 AI는 의료에 적합할까?

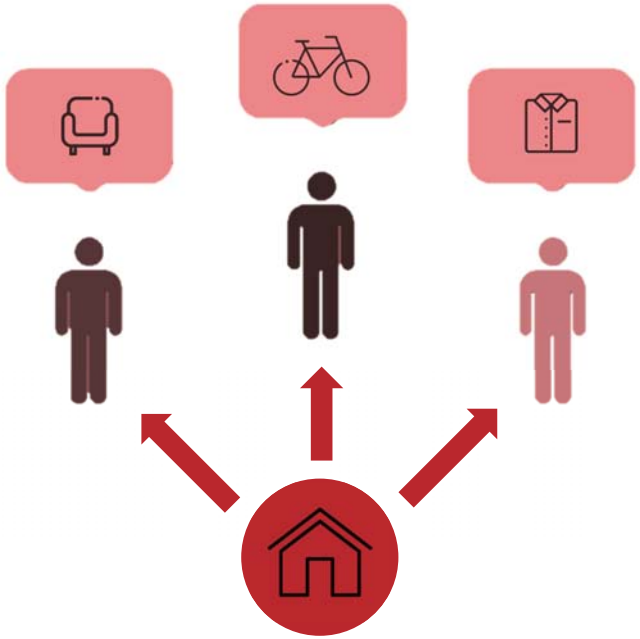
| | 생성AI의 한계점 |
|---|---|
| 1 정확성과 신뢰성, 재현성 | Hallucination 문제, 이미지는 부정확함 |
| 2 증거기반 추론 | Reference가 명확하지 않음 |
| 3 임상 검증 | 임상 검증에 시간이 걸림 |
| 4 의료윤리 준수 | AI는 윤리를 판단하지 못함 |
| 5 데이터 보안 및 개인 정보 보호 | 서버에 정보 업로드시 데이터 유출 |

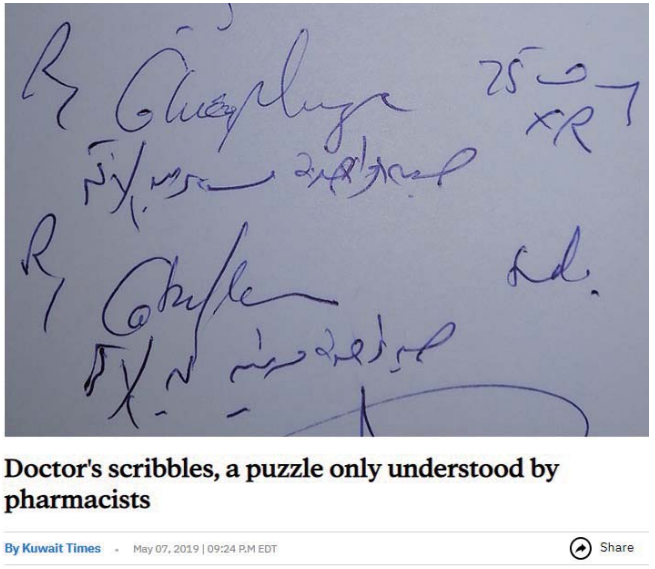
**생성AI를 의료에 바로 적용하는 것에는 한계점이 있음**

# 생성AI는 최신의 의료를 대변할 수 있을까?



# Personalization과 Fine-tuning





Doctor's scribbles, a puzzle only understood by pharmacists

By Kuwait Times · May 07, 2019 | 09:24 P.M EDT Share

- 교과서적 지식을 기반으로 학습한 생성 AI는 의료인 각각의 치료방식을 대변하지 못함
  → **Chatbot 으로서 역할 제한**

- 의무기록의 해석이 어렵고 LLM을 의료용으로 사용하기 위한 조정에 시간이 걸림

https://kuwaittimes.com/doctors-scribbles-a-puzzle-only-understood-by-pharmacists/

# 의료 생성AI가 가야할 방향

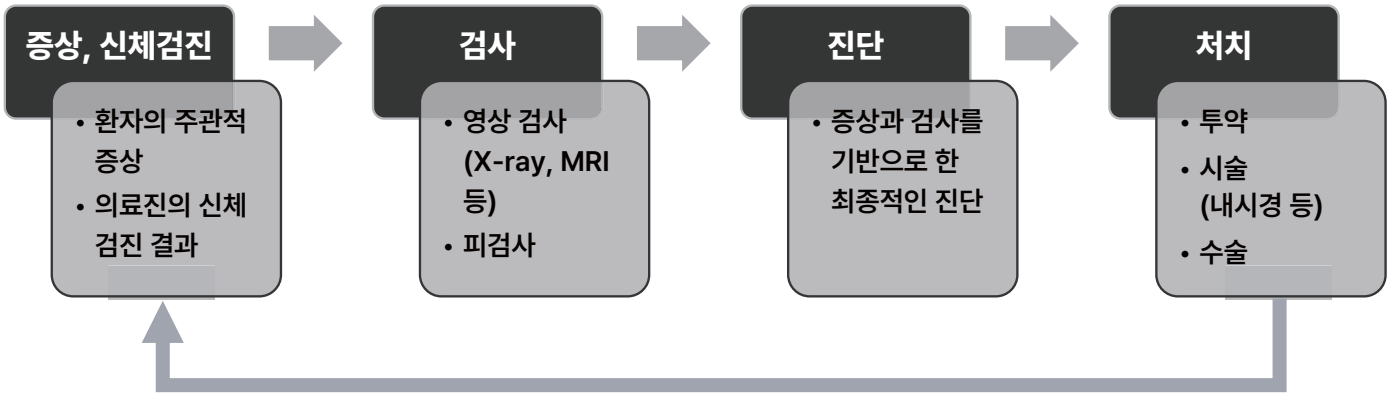
- Text 기반의 생성AI 개발
- 의료인들의 관리가 가능한 분야
- 아직 공개되지 않은 정보의 취득
- 특정 분야에 특화된 서비스
- Personalization

**전자 의무기록**
**(EMR, Electronic Medical Record)**




"https://demigos.com/blog-post/ehr-emr-interoperability/"의 내용을 수정함

# 의료 생성AI가 가야할 방향


## EMR의 특징

- Text 및 시계열 기반 정보
- 실시간 의료 정보로 논문화 되지 않은 데이터
- 의료진의 진단, 치료 등 개인화된 정보

## EMR 기반 생성AI 서비스

- 의료진 개개인의 공개되지 않은 Treatment algorithm 을 학습할 수 있음

→ 개인화된 의료생성AI 서비스

: 의무기록 작성 및 처방 보조, Medical simulator, Clinical navigation, 의료 자문

"https://demigos.com/blog-post/ehr-emr-interoperability/"의 내용을 수정함

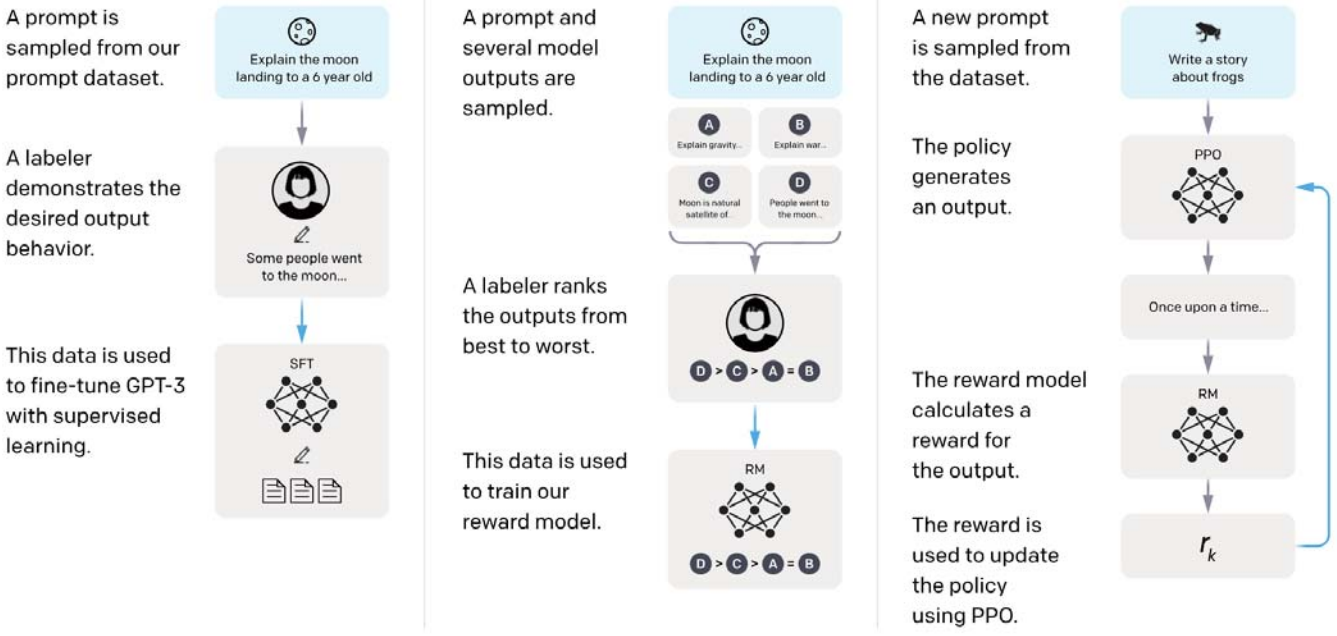
# 전문 의료진의 최종 검증이 필요한 의료 생성 AI

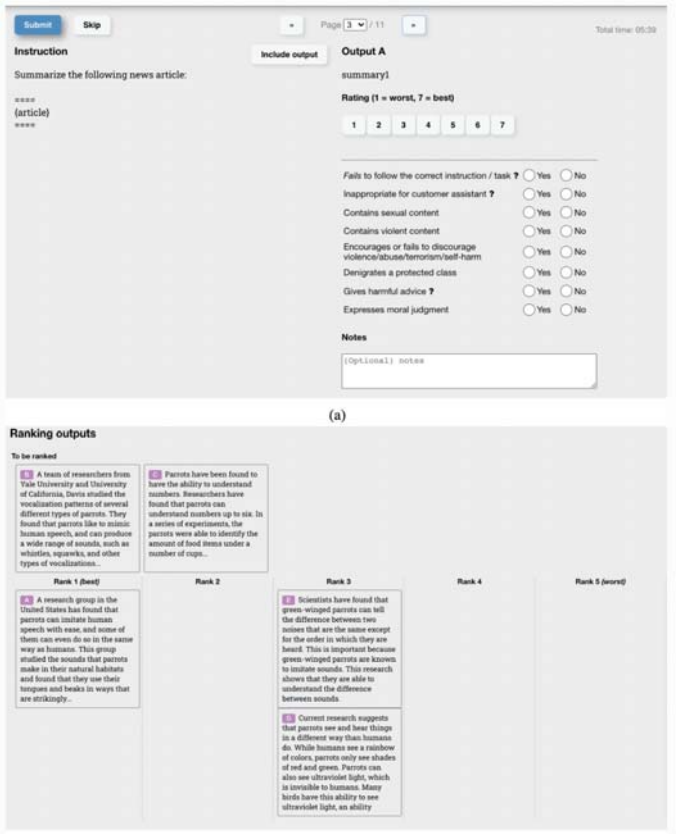
**환자-의사간 대화와 의무기록을 기반으로 차트 작성**

기존 의무기록과
환자-의사의 대화를
종합하여 새로운 차트 작성

- 일반적인 의료 상황에서는 의사가 환자를 상담 하며 동시에 차트 작성하거나, 전담 의료진이 차트 작성에 참여하는 경우가 일반적
  - 환자-의사간 대화와 이전 의무 기록을 바탕으로 차트를 작성하면 의사는 온전히 환자에 집중할 수 있을 뿐만 아니라 의료진의 인력을 절약할 수 있음
- 문제점: 생성 AI 가 만들어주는 차트가 의료진이 원하는 차트가 아닐 가능성이 있음
  - 잘못되거나 누락된 의료 정보
  - 개인화 되지 않은 형식
- 따라서 의료진의 수정된 차트를 바탕으로 의료 LLM을 정렬 (Alignment) 할 필요성 부각

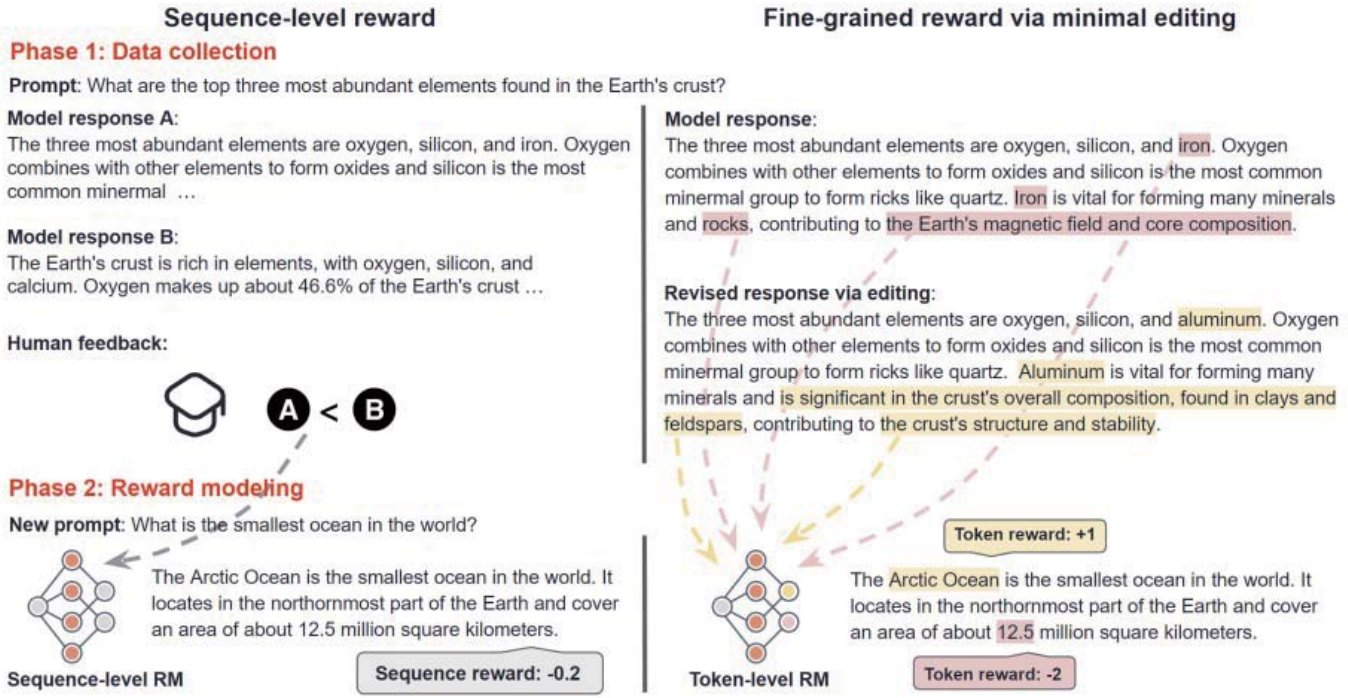
# LLM이 피드백을 학습 할 수 있는 방법: RLHF




**기존의 Sequence-Level Reward 보다 더 fine-grained 한 정렬 방법 필요**

# Token-Level Reinforcement Learning




| Model | Win rate (%) |
|---|---|
| **Fine-grained Token-level PPO** | **51.6 ± 1.8** |
| **Fine-grained PPO** | **51.2 ± 1.8** |
| Davinci003 (reference model) | 50.0 |
| PPO-RLHF | 46.5 ± 1.8 |

Table 1: Evaluation results by Claude. All results of other models are from (Dubois et al., 2023)

| Model | Accuracy (%) |
|---|---|
| **RM w/ Fine-grained dataset** | **85.2 ± 1.8** |
| RM w/o Fine-grained dataset | 58.2 ± 1.8 |

Table 2: Reward model accuracy. Leveraging the fine-grained dataset enhances the reward model's ability to assign correct rewards to responses.

Aligning Large Language Models via Fine-grained Supervision, *under submission*

# 감사합니다