

# Instruction Tuning Large Language Models to Understand Electronic Health Records

Zhenbang Wu<sup>1,2,3</sup>, Anant Dadu<sup>2,3</sup>, Mike Nalls<sup>2,3</sup>, Faraz Faghri<sup>2,3\*</sup>, Jimeng Sun<sup>1\*</sup>

<sup>1</sup>University of Illinois Urbana-Champaign, <sup>2</sup>National Institutes of Health, <sup>3</sup>Data Tecnica

\*Co-corresponding authors



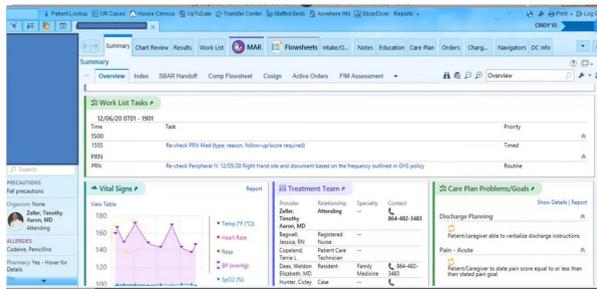
Paper



Code



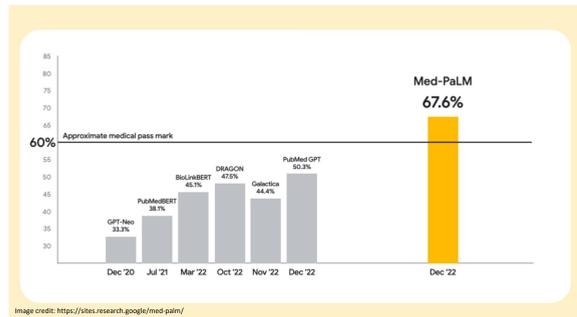
## Background: Complex EHR Systems



- 165.3 alerts/patient/day
- 4.5% were important
- Physicians spent 3.17 hours/day on EHR systems

## Idea: Streamline EHR with LLM

- LLMs can understand complex inputs and follow human instructions to solve diverse tasks
- LLMs also encode clinical knowledge



Example of USMLE-style question

A 32-year-old woman comes to the physician because of fatigue, breast tenderness, increased urinary frequency, and intermittent nausea for 2 weeks. Her last menstrual period was 7 weeks ago. She has a history of a seizure disorder treated with carbamazepine. Physical examination shows no abnormalities. A urine pregnancy test is positive. The child is at greatest risk of developing which of the following complications?

- Renal dysplasia
- Meningocele
- Sensorineural hearing loss
- Vaginal clear cell carcinoma

- Goal
- Develop a conversational AI assistant for EHR data

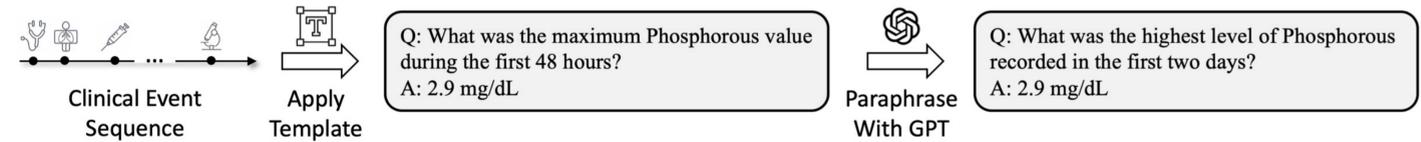
## Challenges

- 1. Lack of large-scale EHR-grounded instruction-following data

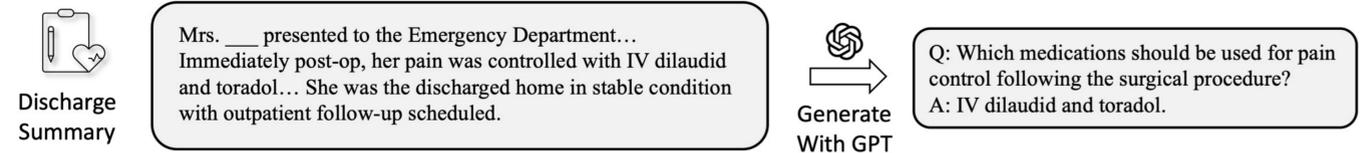
Dataset	Size	Source	Format	Answer Type
MedQA	13K	US medical licensing exam	Question + Answer	Multi Choice
MedMCQA	6K	AIIMS and NEET PG entrance exams	Question + Answer	Multi Choice
PubMedQA	0.5K	PubMed literature	Question + Context + Answer	Multi Choice
MMLU clinical	1K	US Medical Licensing Examination	Question + Answer	Multi Choice
EHRSQL	24K	MIMIC-III	Question + Answer	SQL
EHRNoteQA	0.9K	MIMIC-IV	Question + Note + Answer	Free Text
MedAlign	0.9K	EHRs (Stanford University)	Question + EHR + Answer	Free Text

- 2. Heterogeneous EHR Data
- E.g., Minute-level heart rate vs. weekly glucose tests
- ICD-9 vs. ICD-10
- Blood glucose in mg/dL vs. mmol/L
- Requires complex preprocessing

## MIMIC-Instr: A Dataset of 400K EHR-Grounded Instruction-Following Examples

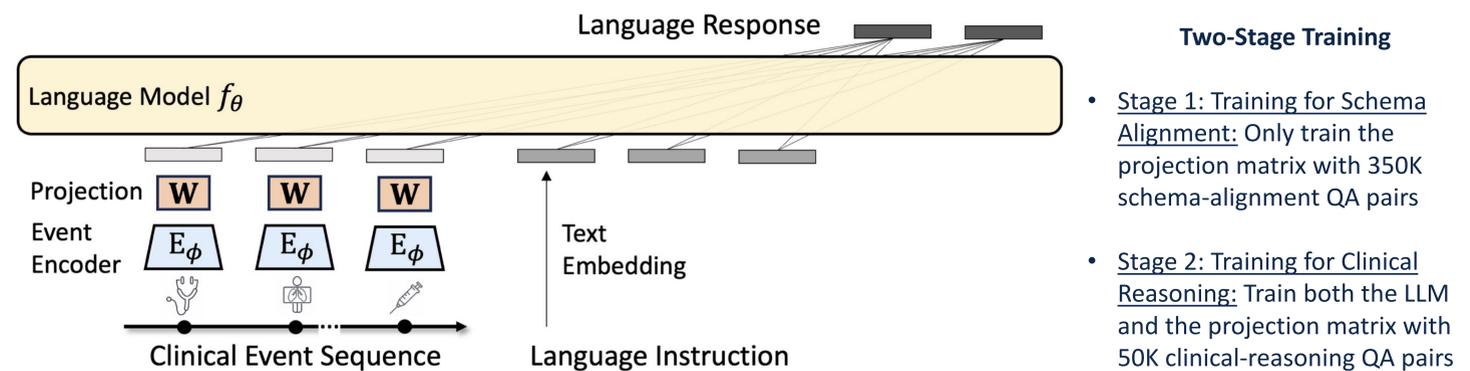


(a) Generation Pipeline for Schema Alignment Data



(b) Generation Pipeline for Instruction Following Data

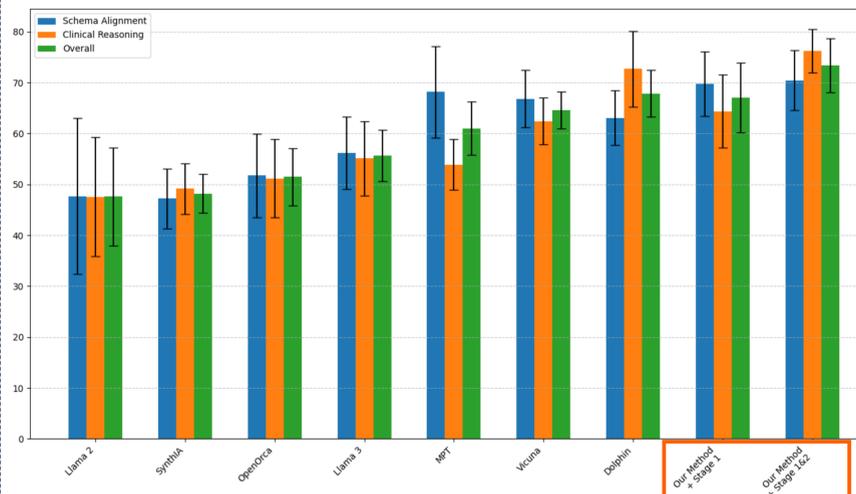
## Llemr: A Foundation Model for EHR Data



- Stage 1: Training for Schema Alignment: Only train the projection matrix with 350K schema-alignment QA pairs
- Stage 2: Training for Clinical Reasoning: Train both the LLM and the projection matrix with 50K clinical-reasoning QA pairs

## Experiment Results

- As Conversation AI Assistant



- On Clinical Predictive Benchmarks

Method	Mortality	Readmission	Length-of-Stay	Diagnosis
RNN	0.8002 (0.02)	0.6643 (0.01)	0.6833 (0.03)	0.7735 (0.01)
Transformer	0.8241 (0.03)	0.7006 (0.01)	0.6990 (0.01)	0.8025 (0.02)
RETAIN	0.8302 (0.02)	0.6994 (0.01)	0.7015 (0.01)	0.8073 (0.02)
GRASP	0.8362 (0.01)	0.7155 (0.01)	0.7100 (0.03)	0.8005 (0.02)
GenHPF	0.8258 (0.02)	0.7102 (0.01)	0.6993 (0.02)	0.8103 (0.03)
REMed	0.8346 (0.02)	0.7193 (0.02)	0.7018 (0.01)	<b>0.8128 (0.01)</b>
Llemr	<b>0.8388 (0.01)</b>	<b>0.7251 (0.03)</b>	<b>0.7132 (0.01)</b>	0.8086 (0.01)

- Example Responses

Schema Alignment	
User	Where was the patient admitted?
Gold	The patient was admitted to the emergency room.
Vicuna	The patient was admitted to the hospital with an emergency room complaint of agitation. The patient was transferred to the Medical Intensive Care Unit (MICU). The patient's diagnosis is not provided in the given data.
Llemr	The patient was initially admitted to the Emergency Room.
Clinical Reasoning	
User	What antibiotic was used for perioperative coverage during the aortic valve replacement surgery?
Gold	Cefazolin was used for perioperative antibiotic coverage during the aortic valve replacement surgery.
Vicuna	The antibiotic used for perioperative coverage during the aortic valve replacement surgery was likely Omeprazole.
Llemr	The data only includes various medications such as analgesics, sedatives, and other treatments, but no antibiotics were prescribed specifically for the surgery.