Creating a ▓▓▓▓▓ Medicine, Instruction tuning dataset

Colleagues,

Thank you for your interest in creating a unique dataset of instructions to align a language model's output to clinician expectations (see ask below). ChatGPT summarizes our challenge via the poem below ...

In modeling, a challenge we faced,
To align a language model's taste,
With clinicians' demands,
Unique dataset in hand,
Instructions refined with great haste!

Best regards,
▓▓▓▓▓▓▓▓▓

As you know, we're building a [Large Language Model](#) (LLM) for use in the clinic, one that would be aware of the patient's EHR data and published medical literature (such as from PubMed). LLMs power applications that interact with humans – such as ChatGPT.

To ensure that the language model's output is aligned with its intended use, we need to train it to respond to instructions (also known as *prompts*) that clinicians like you might send to such a model. **We need your help in generating examples of instructions that might be posed to such a model.**

What:

- Instructions (e.g., "*Write a discharge summary for this patient*") that a clinician might want to submit to a LLM-powered application connected to other clinical technology applications within the hospital. We will be training a Large Language Model to respond to these instructions.
- This video ▓▓▓▓▓▓▓▓▓▓▓ explains the task and how to provide the examples.

When:

- By Tuesday, May 23rd 2023

Why should you care:

- This is an opportunity to influence how Large Language Models are built and used for healthcare
- Co-authorship opportunity (for individuals who contribute 200+ instructions)
- Prizes! We'll do a raffle with five $100 gift cards. Your probability of being selected will be proportional to the number of instructions/questions you submit

Where to provide examples: ████████████████████