Hospital GPT: Proactive Large Medical AI Models to Combat the Future Pandemics

Fenglin Liu, David A. Clifton
[1] Institute of Biomedical Engineering, University of Oxford, Oxford

Introduction: Why are we interested in this?

Background:

- **In the face of a pandemic, time is of the essence**. A slow response to pandemics can lead to **multifaceted and severe consequences**:
  - **Human Health**: Causes a rapid increase in infections, leading to avoidable morbidity and mortality;
  - **Hospitals**: Overwhelms healthcare systems, putting unsustainable pressure on resources and staff;
  - **Economy**: Costs billions in healthcare expenditures, lost productivity, and destabilizes markets;
  - **Social Life**: Disrupts daily life, education, and contributes to mental health challenges, among other effects.
- Since the COVID-19 pandemic, we have encountered **monkeypox (or m-pox)** and a new variant-of-concern of **bird influenza.**
  - The ability to **promptly respond to future pandemics**, especially in **resource-constrained regions**, is crucial.
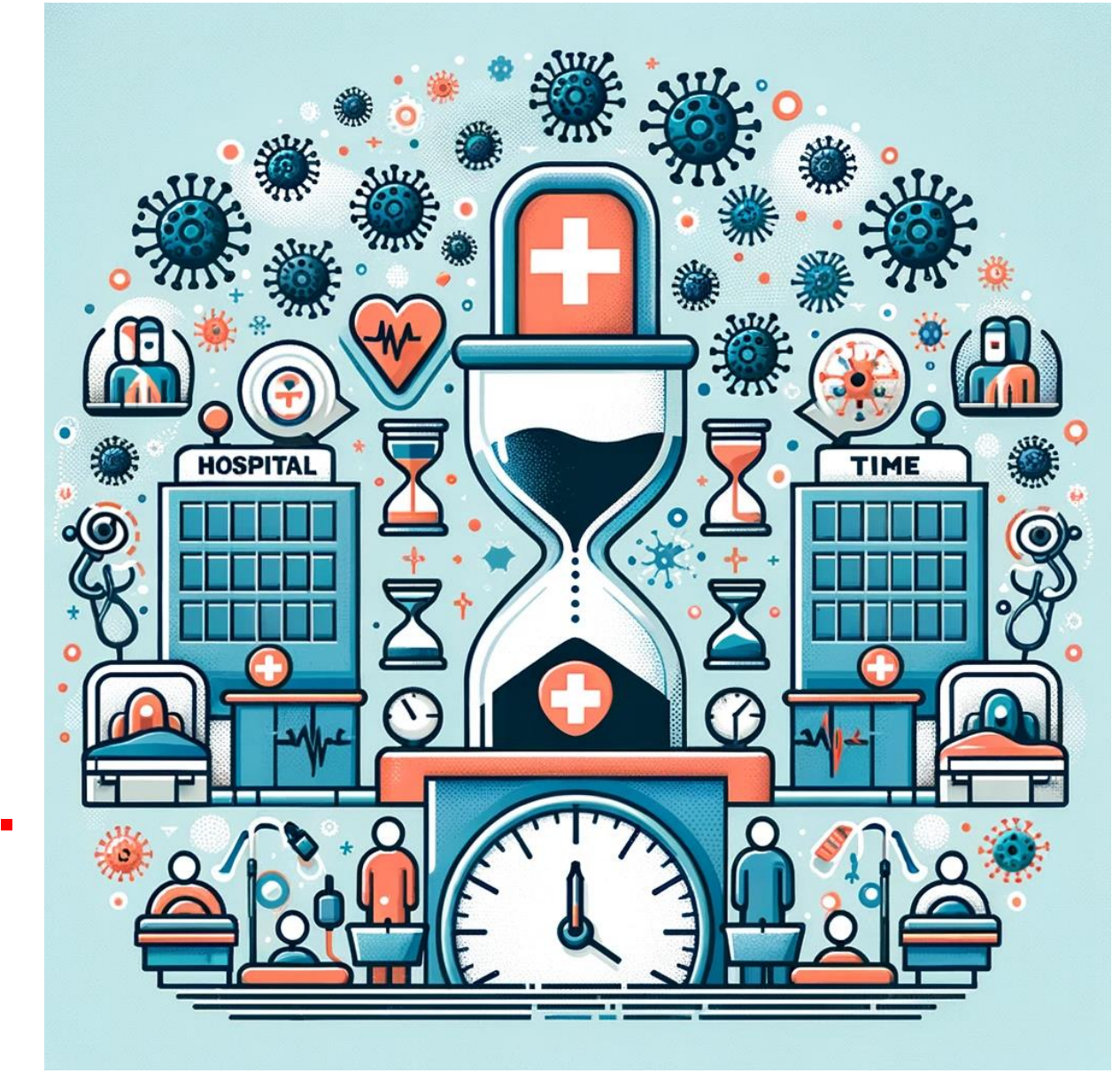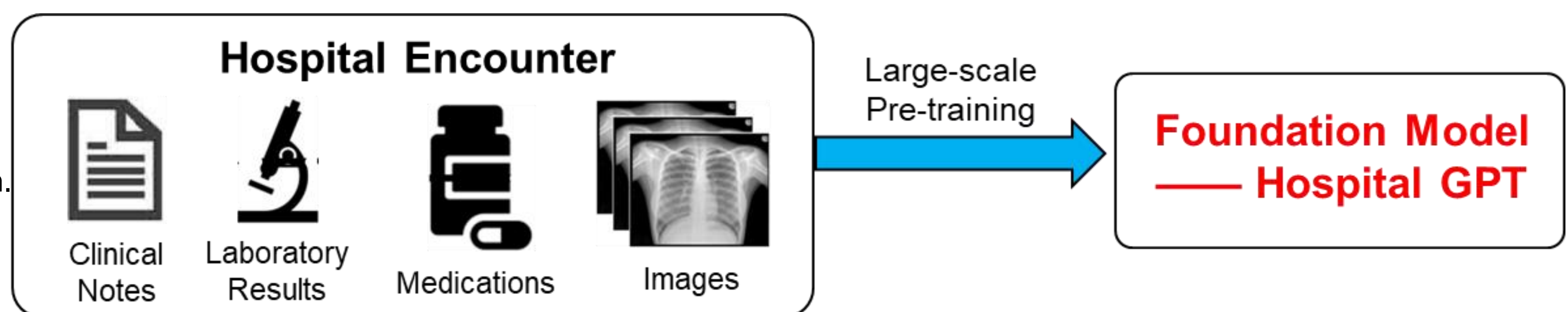
Limitation & Challenge:

- **Existing data-driven AI systems rely heavily on large number of resources**, e.g., data and computing, for deployment.
- For **new pandemics** or **resource-constrained regions**, the sufficient resources are typically unavailable.
  - Taking COVID-19 as an example, the time taken to collect sufficient data means that **1-3 waves (6-12 months) pass** before an AI system can be deployed. By then, arguably, **the worst of the pandemic had passed**, and **hospitals had already come under huge pressure**.
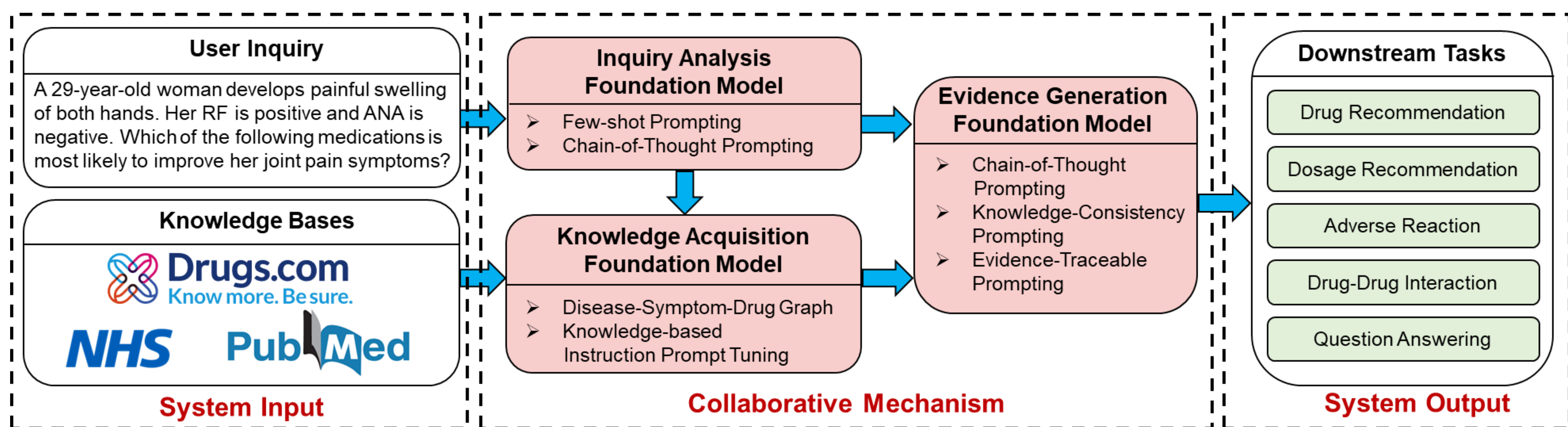
Approach: How did we do this?

System Development:

- This project adopts the **advanced machine learning algorithms**, to implement the Hospital GPT, a **proactive** and **interactive** knowledge-grounded medical large AI system.
- **Medical Foundation Model**: The system is designed to learn broad medical knowledge from diverse publicly-available medical data.

Hospital Encounter — Clinical Notes, Laboratory Results, Medications, Images → Large-scale Pre-training → **Foundation Model —— Hospital GPT**

System Deployment:

User Inquiry
A 29-year-old woman develops painful swelling of both hands. Her RF is positive and ANA is negative. Which of the following medications is most likely to improve her joint pain symptoms?

Knowledge Bases
Drugs.com — Know more. Be sure.
NHS PubMed

System Input

Inquiry Analysis Foundation Model
- Few-shot Prompting
- Chain-of-Thought Prompting

Knowledge Acquisition Foundation Model
- Disease-Symptom-Drug Graph
- Knowledge-based Instruction Prompt Tuning

Evidence Generation Foundation Model
- Chain-of-Thought Prompting
- Knowledge-Consistency Prompting
- Evidence-Traceable Prompting

Collaborative Mechanism

Downstream Tasks
- Drug Recommendation
- Dosage Recommendation
- Adverse Reaction
- Drug-Drug Interaction
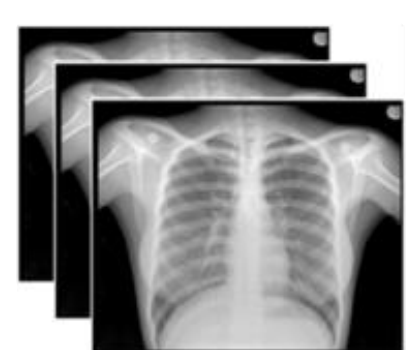- Question Answering

System Output

- We incorporate the **clinical-standard knowledge bases** to make **evidence-based medical recommendations**, improving its transparency and trustworthiness, which are critical in **achieving trust from clinical users** - and **in regulatory approvals that are essential for deployment in clinical practice**.
- **Our system can be easily developed in hospitals to support clinical tasks that involve any types of medical data.**

Result: What did we find?

- Our Hospital GPT can be deployed with **limited resources** and thus can be ready to deploy promptly **within 1-2 weeks** when encountering a new pandemic, rather than within months, for reducing the adverse impacts of future pandemics:
  - **Saving lives**, **protecting economies**, and **ensuring that healthcare systems run more intelligently, effectively, and smoothly**.

- **Clinical Note Generation**: *e.g., Medical Report Generation*
  ✓ Lungs are clear. Heart and mediastinum of normal size and contour. No pleural effusions or pneumothoraces. scoliosis.

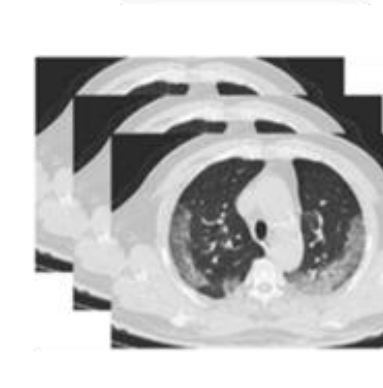- **Disease Diagnosis** *e.g., Question Answering*
  Question: Do granulomas on skin go away?
  Answer: Granulomas on the skin are a chronic condition that can last anywhere from a few months to a few years to go away on its own.

- **Health Management** *e.g., ICD-Coding*
  Clinical Note: Acute upper respiratory infection, unspecified. **ICD-10-CM Code**: J06.9
  Procedure: Computerized Tomography (CT Scan) **ICD-10-CM Code**: B22

Achievements: Why is this important?

- 1. **Academic Publications**: An exceptional number of high-quality research outputs for first-author publication in the top-tier literature for our field [1,2,3,4,5];
- 2. **Industry Commercialisation**:
  - **GlaxoSmithKline** is evaluating the research outputs of my project with their own datasets,
  - **EMIS**, the UK's largest provider of Primary Care computer systems, is keen to develop it for use in supporting GPs when they see patients.
  - **Royal Veterinary College** is likely to lead to commercialisation my research for societal impact, supported by OUI and Oxford Science Enterprises.
- 3. **Patents**:
  - Three research outputs [1,2,4] are **being filed for patenting** by Oxford University Innovation (OUI)
  - One of which [4] was **fast-tracked** (at considerable expense to OUI) due to its perceived importance.
  - My project is an important component of **attracting large quantities of institutional funding** for a new initiative at oxford university, named "Oxford X".

References

[1] A Medical Multimodal Large Language Model for Future Pandemics. **npj Digital Medicine** (2023).
[2] Retrieve, Reason, and Refine: Generating Accurate and Faithful Patient Instructions. **Advances in Neural Information Processing Systems** (2022).
[3] ZeroNLG: Aligning and Autoencoding Domains for Zero-Shot Multimodal and Multilingual Natural Language Generation. **IEEE Transactions on Pattern Analysis and Machine Intelligence** (2024)
[4] DrugGPT: A Knowledge-Grounded Collaborative Large Language Model for Evidence-based Drug Analysis. Submitted to **Nature Biomedical Engineering** (2023).
[5] A Prompt-based Deep Learning Framework for Accurate Novel Thorax Disease Reporting with Extremely Limited Labels. Submitted to **Nature Machine Intelligence** (2023).