

Roadmap to trustworthy medical LLMs

2024年6月16日

Benyou Wang



香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

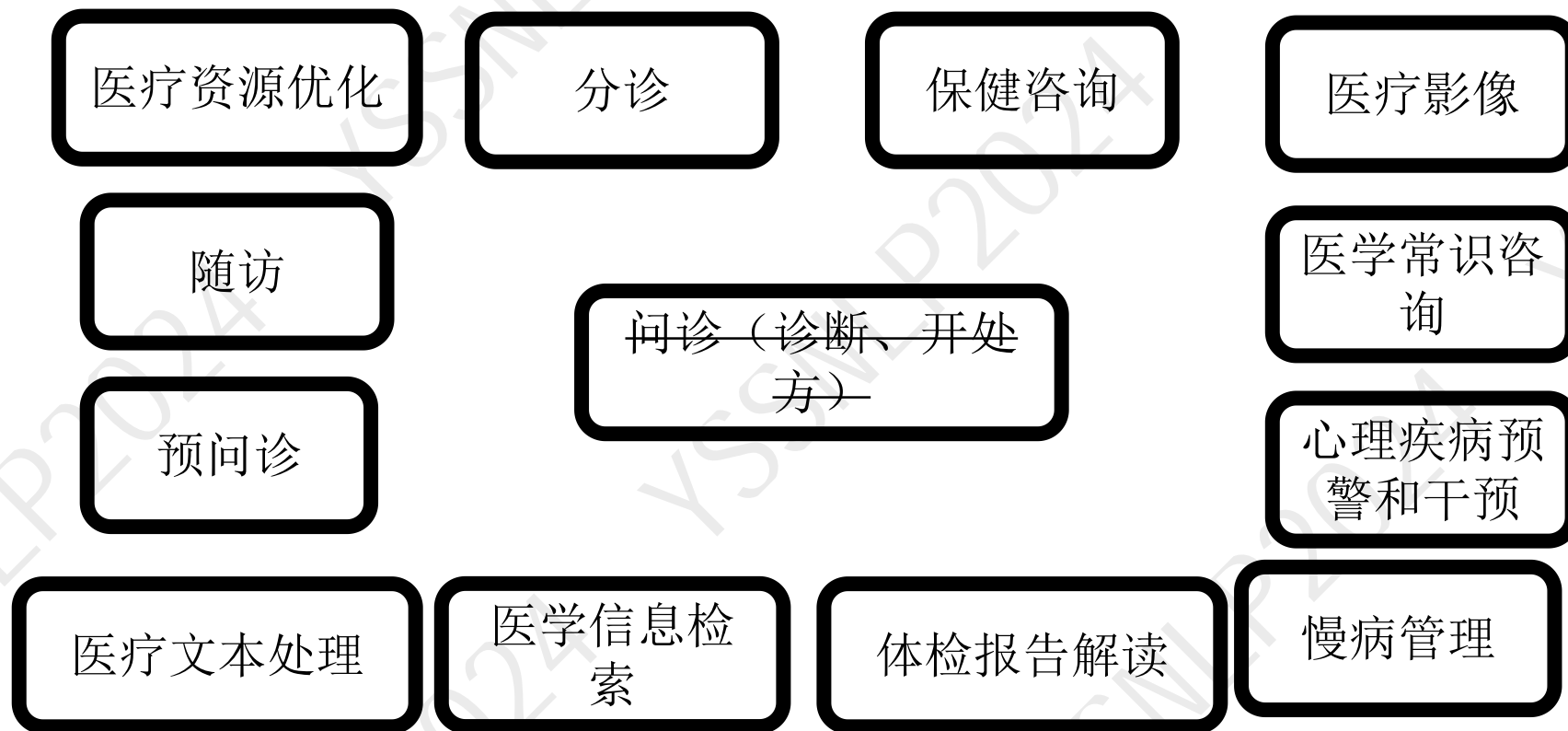


深圳市大数据研究院
Shenzhen Research Institute of Big Data



国家健康医疗
大数据研究院(深圳)
NATIONAL HEALTH DATA INSTITUTE, SHENZHEN

落地场景



Challenge: 医疗大模型的评估困难



Guiming Hardy Chen, Shunian Chen, Ruifei Zhang, Junying Chen, Xiangbo Wu, Zhiyi Zhang, Zhihong Chen, Jianquan Li, Xiang Wan, Benyou Wang. ALLaVA: Harnessing GPT4V-synthesized Data for A Lite Vision-Language Model. <https://arxiv.org/abs/2402.11684>

Challenge: 鲁棒的医疗诊断和医疗决策任重道远

Medical diagnosis and clinical decision making

HuatuoGPT-II
Pass some exams

Accuracy of
professional knowledge

HuatuoGPT

Chat like a doctor

We are picking the **low-hanging lychee**

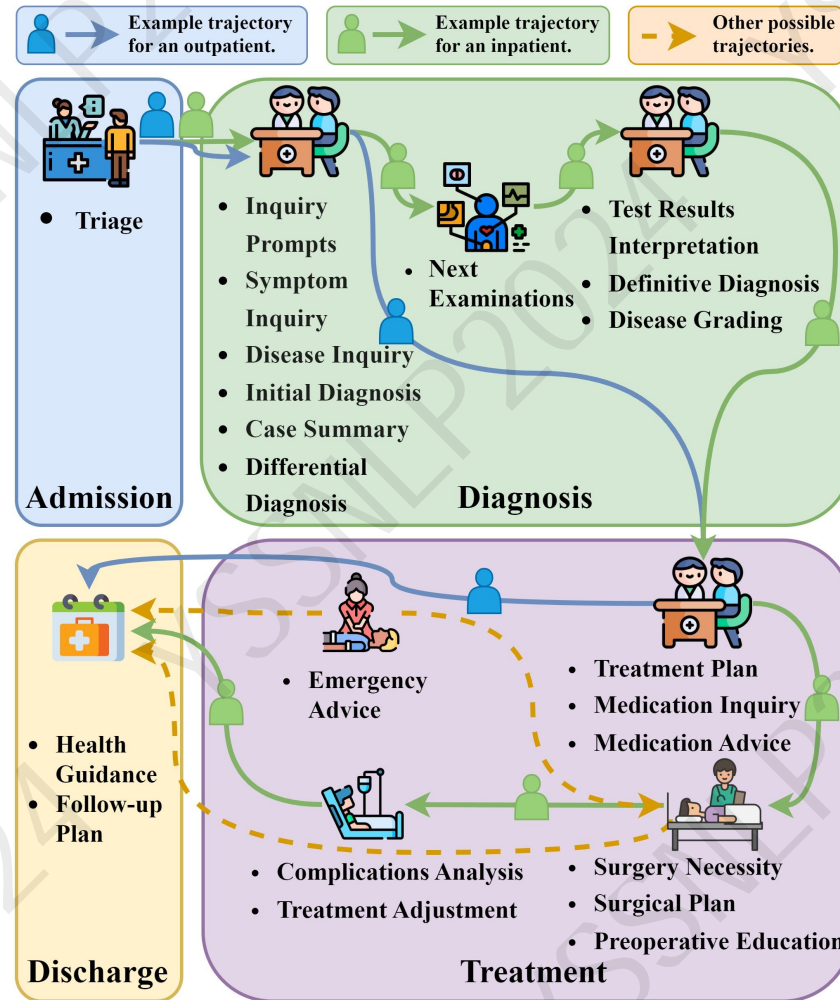


Select a well-defined problem

Desterata:

- 不与医生竞争，服务医生
- **评估**客观简单可复现
- 更大的受众，可以服务更广泛的社会
- achievable in a few years

Standpoint: LLM for Doctors not as Doctors



Wenya Xie, et.al. Benyou Wang. LLMs as for Doctors: Leveraging Medical LLMs to Assist Doctors, Not Replace Them. in Progress

Problem Definition

Definition:

- Given a user with implicit/explicit information
- Goal: predict a label (分类任务: 科室/开药/诊断)
- Key
 - The system could actively ask followup questions to reduce uncertainty
 - The user could lose patience if with too many interactions

科室分类是一个最容易来验证的场景！作为一个GMAI的sandbox

Overall

- **More Knowledge**

- from **input** side: RAG native training
- from **output** side: Agent with external knowledge sources
- internalize it in model: Inject medical rules with synthesized data

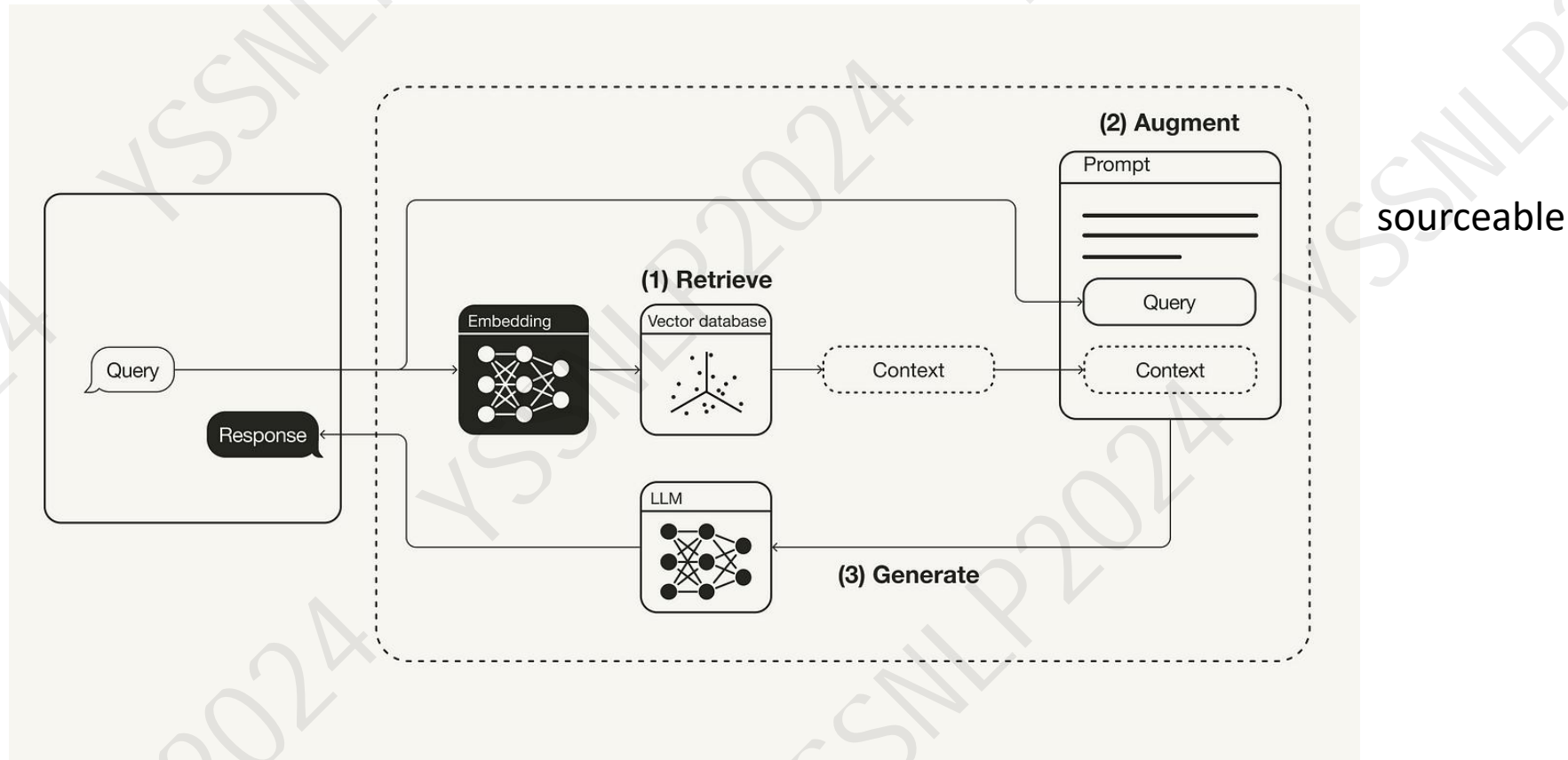
- **More transparent reasoning path**

- Interpretability with Chain of diagnosis

- **With more modality**

RAG native training

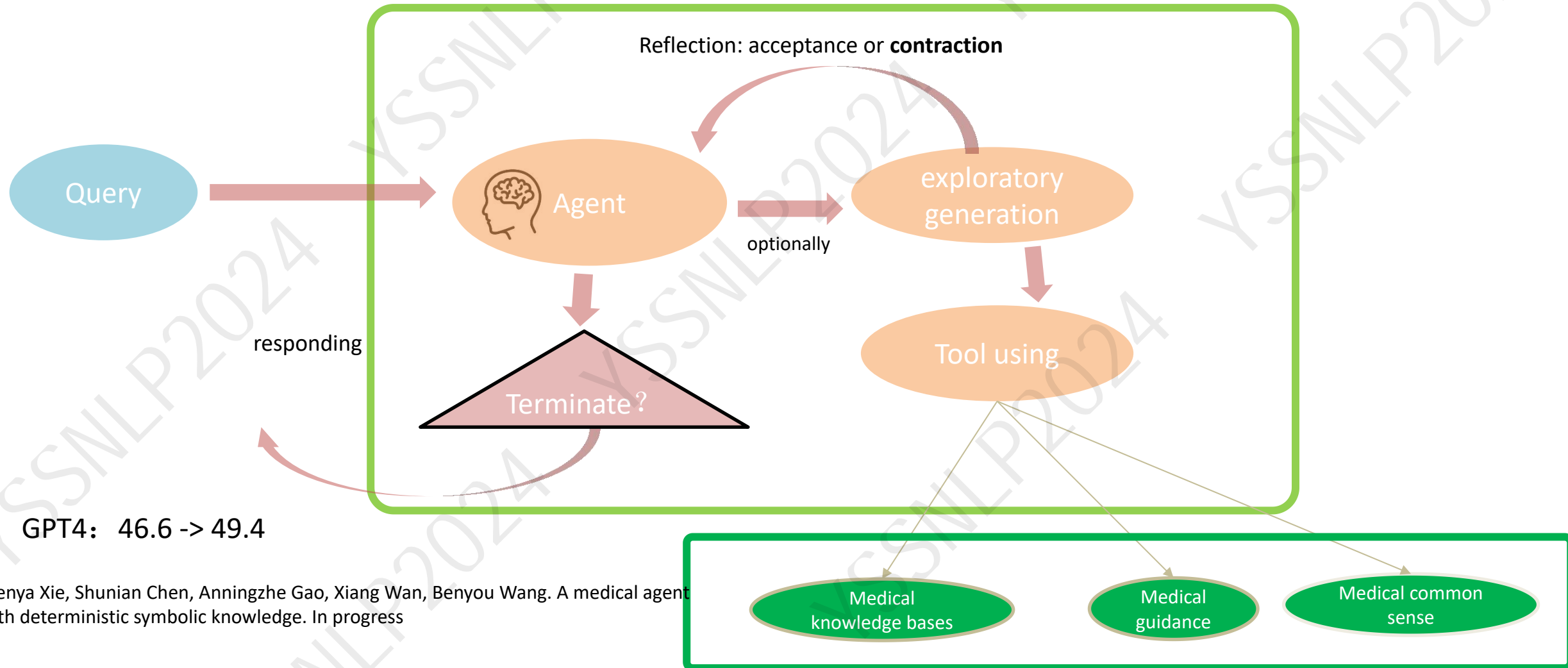
知识应该内化还是外挂？ 边界在哪儿？



RAG Native training 把大模型从记忆中解放出来，把学习的重心放到推理等高级能力
encourage grounded generation (based on limited sourceable data) instead of creative generation

Towards a medical Agent

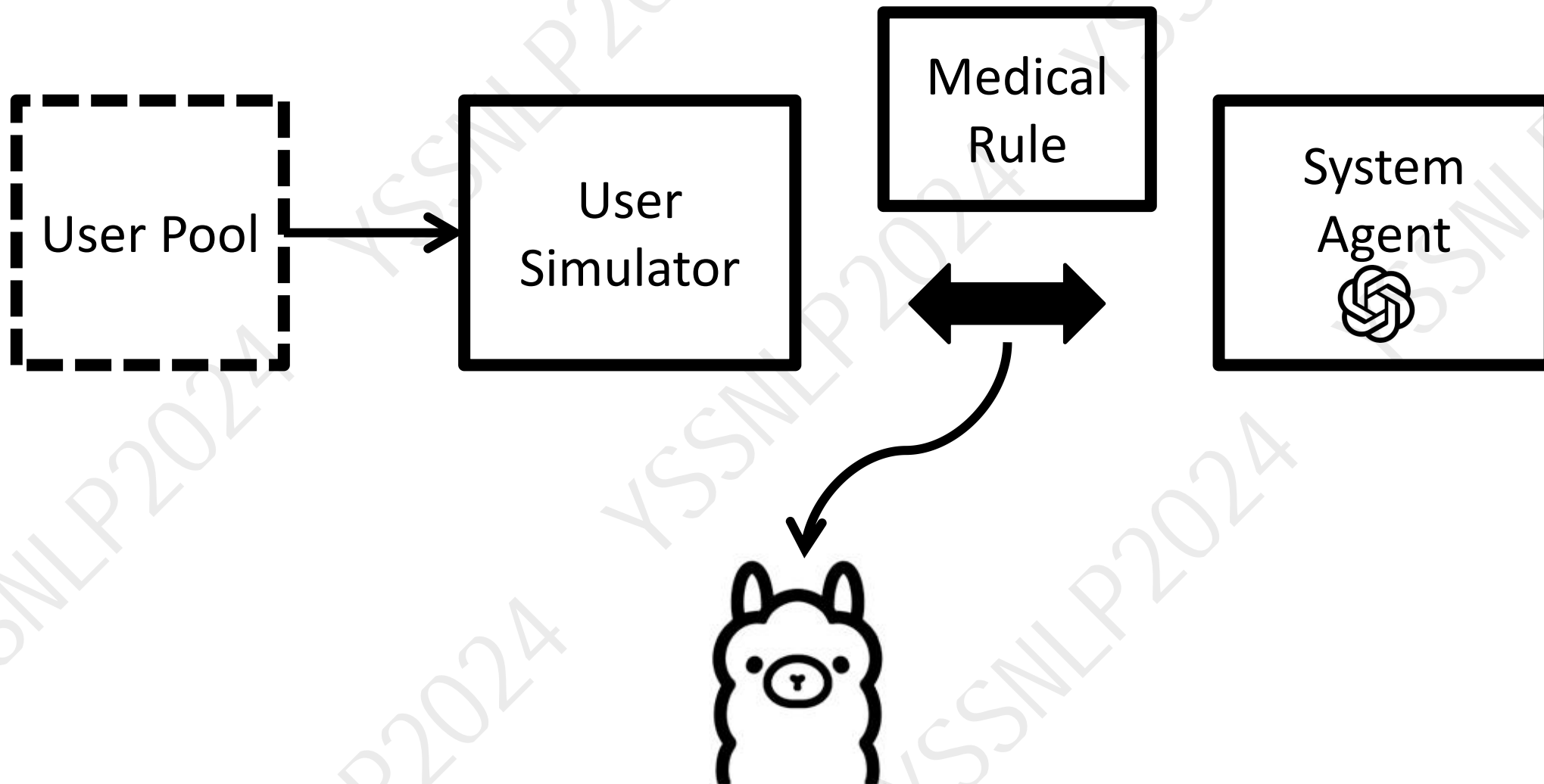
Medical LLM Agent in a **neural-symbolic** fashion



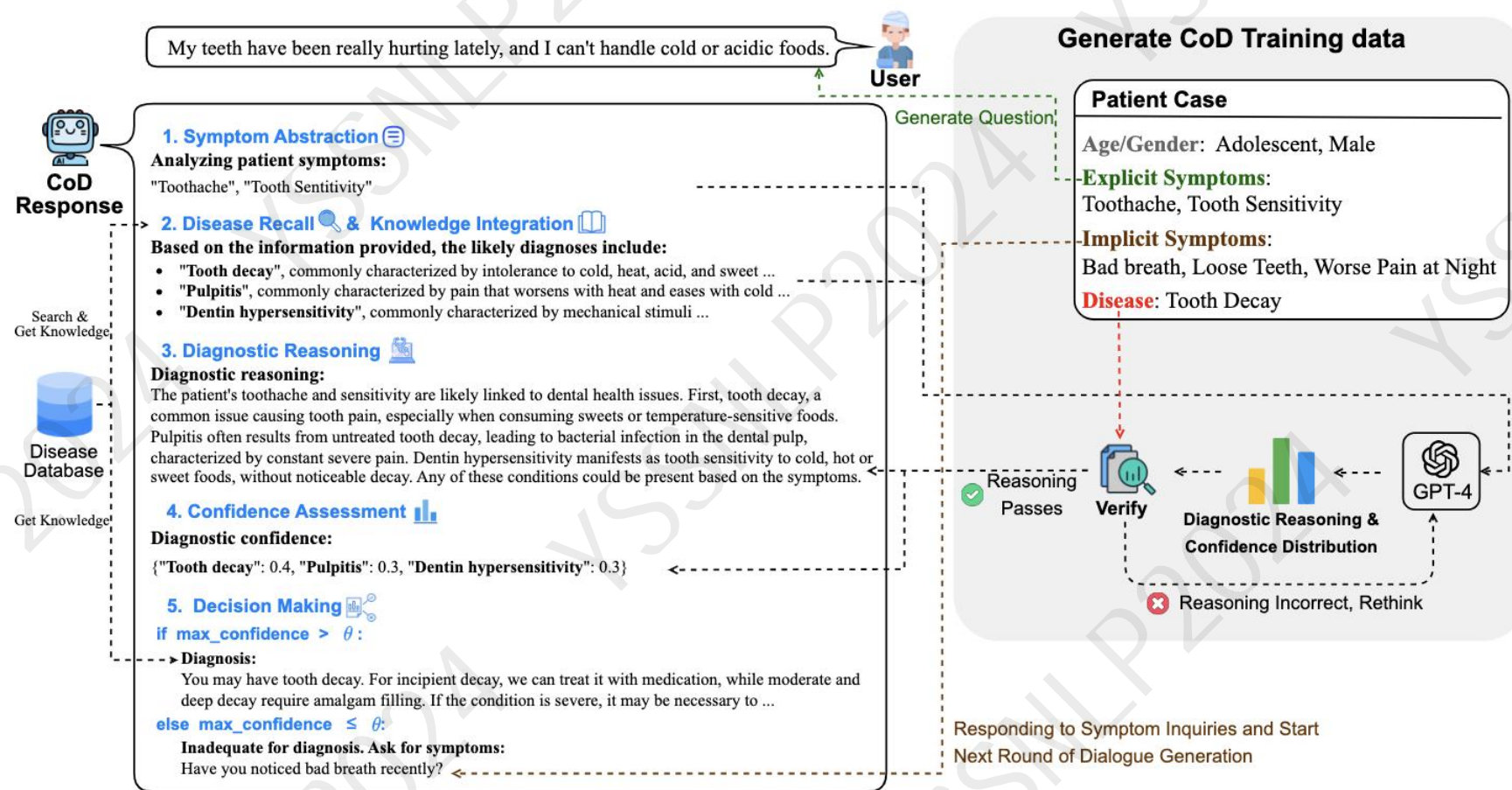
GPT4: 46.6 -> 49.4

Wenya Xie, Shunian Chen, Anningzhe Gao, Xiang Wan, Benyou Wang. A medical agent with deterministic symbolic knowledge. In progress

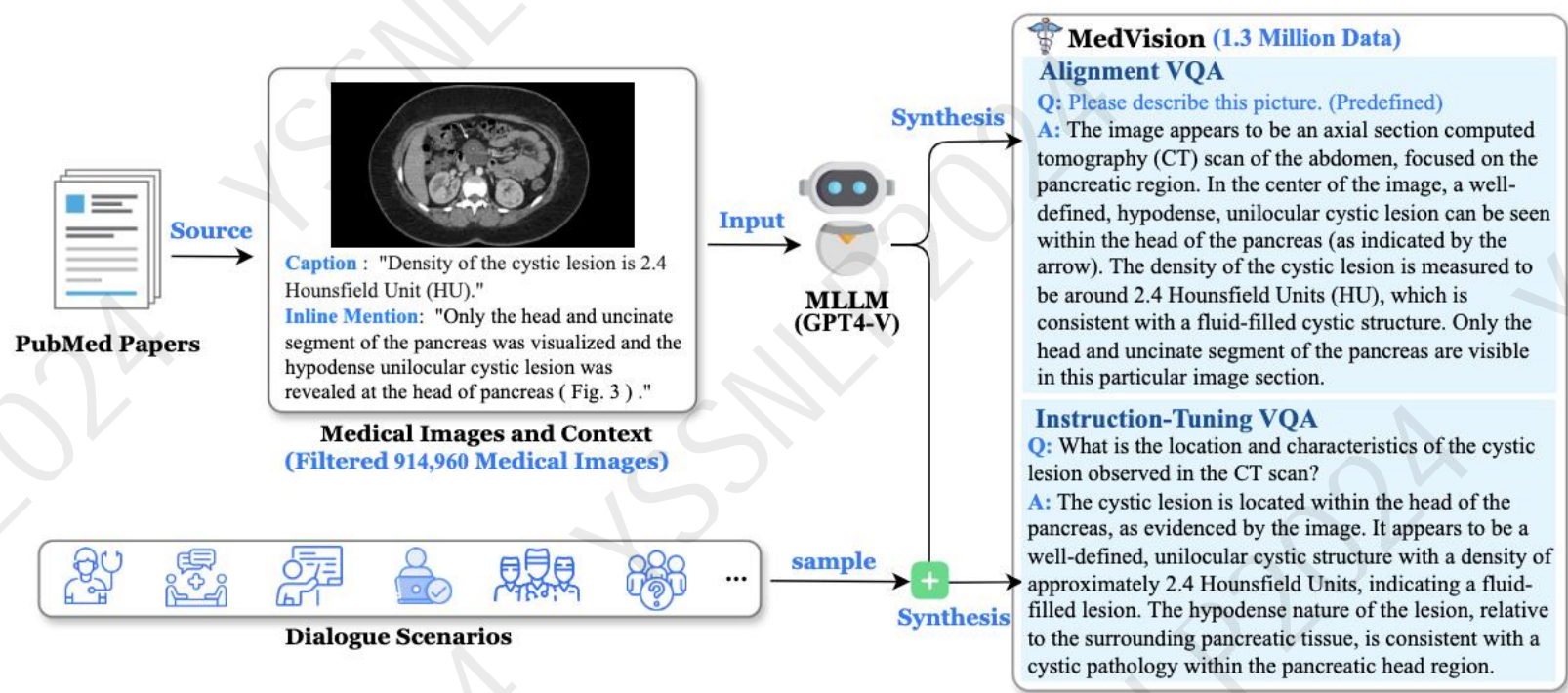
合成数据来注入知识



可解释的诊断链CoD



多模态信息



Thanks

wangbenyou@cuhk.edu.cn

We are hiring RAs & PhD (even for 2024 fall)

