Human-AI Interaction in Healthcare: Three Case Studies About How Patient(s) And Doctors Interact with AI in a Multi-Tiers Healthcare Network

Yunzhi Li[1]
liyunzhi15@mails.ucas.ac.cn

Liuping Wang[1,2]
wangliuping17@mails.ucas.ac.cn

Shuai Ma[1,2]
mashuai171@mails.ucas.ac.cn

Xiangmin Fan[1,2]
xiangmin@iscas.ac.cn

Zijun Wang[3]
Zijun.Wang@ibm.com

Junfeng Jiao[4]
pgjjf@163.com

Dakuo Wang[3]
Dakuo.Wang@ibm.com

[1]University of Chinese Academy of Sciences, Beijing, China

[2]Institute of Software, Chinese Academy of Sciences, Beijing, China

[3]IBM Thomas J. Watson Research Center, Yorktown Heights, NY USA

[4]Beijing Pinggu District Health commission, Beijing, China

Abstract
This position paper presents three ongoing research projects that aim to study how to design, develop, and evaluate the systems supporting human-AI interaction in the healthcare domain. Collaborating with the local government administrators, hospitals, clinics and doctors, we get a valuable opportunity to study and improve how AI-empowered technologies are changing people's life in providing or receiving healthcare services in a suburb district in Beijing, China. We hope this work will ground the discussion with other participants in the workshop and build further collaborations with the health informatics community.

Author Keywords
Health informatics; Human-AI collaboration; AI-empowered healthcare; multi-tiers healthcare.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Introduction
The advancement of AI technologies in recent years have empowered healthcare systems and is changing the landscape of how people provide or receive

healthcare services. For example, researchers have built neural network models to use a user's text input to recommend the diagnosis at a fairly high accuracy [4]. IBM Watson Oncology, with some controversy, has proved that it can predict the symptom based on scan image data as good as human doctors[2]. However, we have rarely seen a massive user adoption of these systems, neither at the healthcare service provider side, nor at the patient side.

In this work, we have a rare opportunity to collaborate with various stakeholders in the healthcare system in a suburb district in Beijing, China. The local government agencies are collaborating with AI start-ups, local hospitals and clinics to construct a new multi-tiers healthcare infrastructure. From the patient's perspective (shown in Figure 1), there are four different tiers:

1. Self-diagnosis and early symptom awareness with AI agents
2. Telehealth (or in-person visit) with PCPs in nearby clinics in the community or in the rural village
3. Examinations in a lab at examination centers or in a general hospital
4. and In-person visit with a specialist.

In addition to these four steps of the user journey, a user often needs to browse an online community regarding her own symptom or medical condition (e.g., cancer [3]or pregnancy [5]). And after the diagnosis and treatment, the patient needs to interact (explicitly or implicitly) with pharmacies or with medical issuance companies' information systems. Thus, we include these aspects in our picture too.

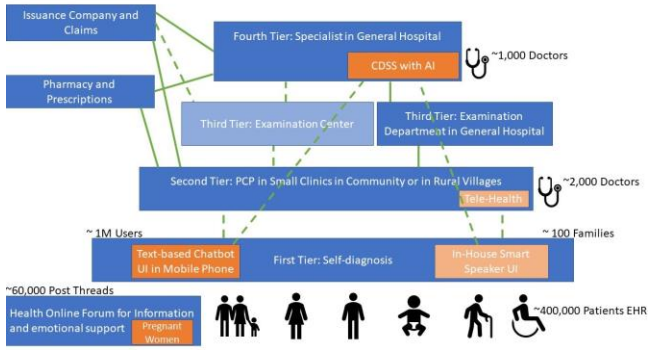

Figure 1: Multi-tiers Healthcare Infrastructure in Pinggu, a suburb district in Beijing China.

The whole infrastructure is still evolving and some of the components are in development (e.g., an in-house smart speaker system for self-diagnosis; and 3rd-party examination centers that can share data within the network). But the present picture is already very complex that deserves an exploration. In this workshop paper, we would like to focus on three components of the system, and we have ongoing research project in each component:

1. A NN-based AI moderator for online forum. It can generate new responses to users in an online community for pregnant women;
2. How users (millions) are using a chatbot in mobile phones for self-diagnosis, and the research questions around explainability and trusted AI;
3. How PCPs and specialists (thousands) are using an AI-empowered Clinical-Decision Support System in their daily work practice and what

Figure 2: AI Interact with User in a Pregnant Woman Online Community

Figure 3: Four Categories of Posts and Examples of Corresponding AI-Generated Responses in the Pregnant Women Online Community

the challenges are in this Human-AI collaboration.

One note before we move to introduce each project respectively, we would like to elaborate our findings generalizability. Even though the case studies included in this paper focus only on a city in China, we do argue that the experience and lessons learned from these projects can generate to other geographic areas. For example, the four-tiers healthcare network design is compatible to the U.S.'s telehealth initiative for the country side. We also believe that the explainability and trust issues that our users, both doctors and patients, are facing in interacting with AI systems are similar to the issues happening in other high-stake decision-making scenarios[1]. Thus, we sincerely hope that we can join this effort at the workshop to discuss our works with other participants.

AI Community Moderator for Online Forum for Pregnant Women

Women in pregnancy often look for a group to seek informational and emotional support [4]. Online forum is such a place. However, it is not possible to get a response in a timely manner. In this project, we propose to tackle this problem through latest Neural Network language generation models. Then, we design and build a chatbot user interface on top of the backend algorithm.

The research questions include:

1. How to integrate AI generated responses into the forum?
2. How much do the pregnant women trust in the AI partners in this community?
3. What impact will AI have on this community?

At this stage, we have completed the work of building a seq2seq-based language model to generate creative responses trained with the users' response data in this pregnancy forum (Figure2). Appropriate advice can be provided by AI for pregnant women, whether they want to seek emotional or informational support or simply share their life. At the same time, we carefully analyzed the data, including the title, text and picture content of the post, the reply content, and the basic information of the user, such as the expected date of childbirth. We manually coded 2,000 post-comment pairs and found there are four categories: seeking emotional support, seeking informational support, sharing positive emotions and share daily life. Then, we used Convolutional NN to train on these coded data and classify the rest of the unlabeled messages.

The next step of this project is that we are building a chatbot interface to communicate our AI generated responses back to the forum. Our chatbots will identify the posts that have not been answered for a long period of time, and reply to provide their emotional or information support. We will have a user study with follow up interview user sessions to solicit users' feedback to improve. We are particularly interested in and will research on the pregnant women's trust issue in AI.

Self-Diagnosis of Disease With A Chatbot In Your Hand (Left Hand Doctor)

With the development of the economy and the advancement of modern technologies, human’s life expectancy continues to increase and people have increasingly paid attention to their own health.

Figure 4: LHD Self-Diagnosis UI

Therefore, the expense of supporting the relevant medical ecosystems would become a serious social burden, especially with the aging of a country's population. In China, the world's largest developing and populous country, this phenomenon is particularly significant. Recently, with the aim of providing the general public with intelligent and professional self-diagnosis services, a Chinese medical smartphone application named Left Hand Doctor (LHD) came out. LHD has two major components: 1) intelligent self-diagnosis component 2) intelligent medication consultation component. The intelligent self-diagnosis component is a text-based and system-initiative dialogue system, as well as a new kind of Personal Informatics System. By asking numbers of questions about the user's symptoms, previous medical history and o ther relevant information, LHD will provide several highly possible diseases annotated by their possibilities, relevant hospital departments and other medical advices such as possible treatment, as shown in Figure 4. LHD will also automatically generate a standard medical report to provide documentary support for user's further professional counsel. The intelligent medication consultation component has a similar working principle and recommend proper medicine to the user, based on the self-reported information such as the allergy history, liver function, etc.

Under this scenario, we have found several valuable research topics: 1) Whether people are more willing to talk about their private issues by interacting with chatbots rather than interacting with real clinicians and what is the reason behind this situation? 2) What is the composition of the LHD's user community in different aspects such as age, health condition and education background? 3) How do the users view the advice or diagnosis provided by LHD? Are they really trust the advices provided by the system or they still need confirmations from real clinicians?

Therefore, we plan to conduct both interview study and laboratory study in the next one or two months to fully investigate these problems.

Human-AI Collaboration in Clinical-Decision Support System

Recently, numbers of research results within the intersection of Artificial Intelligence and Clinical Diagnostics have clearly demonstrated the powerful performance of AI-based automated diagnosis. Pinggu District is in the NorthEast of Beijing, with 4 general hospitals in the town center and more than 20 clinics in rural villages. More than 430,000 permanent residents receive medical services from this large medical infrastructure. However, unlike most hospitals in China, even all over the world, an integrative AI-based diagnosis support system has already been deployed in all the general hospitals and rural healthcare stations there, with the aim of helping nearly 3,000 clinical doctors there with their clinical diagnosis and medical decision-making.

This system has three major components: 1) intelligent diagnostic component 2) prescription recommendation component 3) medical knowledge base component.

Specifically, the intelligent diagnostic component, which is shown in Figure 5, will rank and display possible diseases (including probabilities) according to the confidences which are generated by analyzing the symptoms mentioned in the previous medical history,

the patient's self-statement and other relevant information. The prescription recommendation component will recommend medication and give corresponding reasons based on the patient's medical history, allergic history, basic information and other relevant medical records to help clinicians make better decisions. The medical knowledge database component contains massive and diverse electronic health records and medical knowledge, such as the definition and symptom of hyperthyroidism, which is anticipated to assist doctors in their enquiries and references.

However, clinical diagnosis is a high-stakes decision-making problem with high uncertainty and complexity. Doctors are fully responsible for the correctness of their diagnosis and any subtle diagnostic error could lead to a significant impact on the patient's health. Therefore, we are interested in understanding the usage patterns of the users using this AI-based decision support system. To be more specific, we are trying to answer three questions: 1) Is this system actually used by clinicians to support their diagnosis for real patients? 2) How did they use this system and how did they combine the advice provided by the system with their own judgment, especially under the circumstances that the two judgments conflict with each other? 3) How do they view the advice provided by the intelligent system? Are doctors really trust the advices provided by the system when they differ from or are similar to the doctors'? In addition, we are interested in exploring that are there any differences between the usage patterns of experienced doctors and newly qualified doctors.

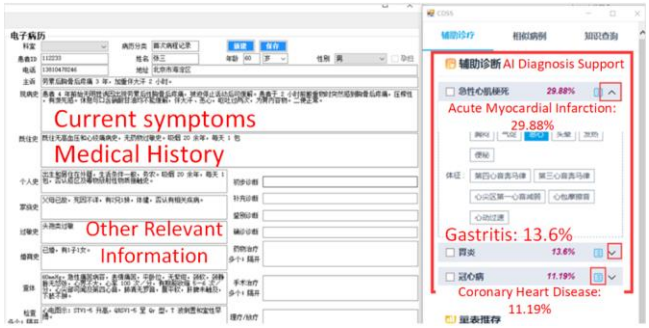

Figure 5: Human-AI Collaboration in CDSS

Therefore, in a month, we plan to conduct a one-week observational study by shadowing the doctors in general hospitals, as well as in the clinics in rural areas in Pinggu district. We will keep detailed records and carefully analyze the doctors' usage patterns. In addition, we plan to conduct interview studies with doctors and patients there through well-designed questionnaires to further discuss the trust issues between AI and human in the field of healthcare.

Conclusion

In this work, we present a research plan with 3 cases studies that aim to systematically investigate the human-AI interaction experiences in the various domains of healthcare. Our context is in Beijing China, and our space is still limited. There are more opportunities, such as AI use in ER room or in patient portal that we do not cover. We plan to leverage on this workshop to exchange knowledges and resources with other participants.

References

1. Matthew Arnold, Rachel K. E. Bellamy, Michael Hind, et al. 2018. FactSheets: Increasing Trust in AI Services through Supplier's Declarations of Conformity. *arXiv:1808.07261 [cs]*.

2. Alessandra Curioni-Fontecedro. 2017. A new era of oncology through artificial intelligence. *ESMO open* 2, 2: e000198–e000198.

3. Mies C van Eenbergen, Lonneke V van de Poll-Franse, Peter Heine, and Floortje Mols. 2017. The Impact of Participation in Online Cancer Communities on Patient Reported Outcomes: Systematic Review. *JMIR cancer* 3, 2: e15–e15.

4. Robert Lindsey, Aaron Daluiski, Sumit Chopra, et al. 2018. Deep neural network improves fracture detection by clinicians. *Proceedings of the National Academy of Sciences* 115, 45: 11591–11596.

5. Yi-Chia Wang, Robert Kraut, and John M. Levine. 2012. To stay or leave?: the relationship of emotional and informational support to commitment in online health support groups. *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work - CSCW '12*, ACM Press, 833.