

CHIRP AI Launch Event

Sean Mooney, PhD



Providers and AI

- Eric Topol recently posted on LinkedIn that there are now 5 reports of AI outperforming physicians, even when they use that same AI.



Jonathan Chen, Stanford

Original Investigation | Health Informatics

October 28, 2024

Large Language Model Influence on Diagnostic Reasoning A Randomized Clinical Trial

Ethan Goh, MBBS, MS^{1,2}; Robert Gallo, MD³; Jason Hom, MD⁴; Eric Strong, MD⁴; Yingjie Weng, MHS⁵; Hannah Kerman, MD^{6,7}; Joséphine A. Cool, MD^{6,7}; Zahir Kanjee, MD, MPH^{6,7}; Andrew S. Parsons, MD, MPH⁸; Neera Ahuja, MD⁴; Eric Horvitz, MD, PhD^{9,10}; Daniel Yang, MD¹¹; Arnold Milstein, MD²; Andrew P. J. Olson, MD¹²; Adam Rodman, MD, MPH^{6,7}; Jonathan H. Chen, MD, PhD^{1,2,13}

» [Author Affiliations](#) | [Article Information](#)

JAMA Netw Open. 2024;7(10):e2440969. doi:10.1001/jamanetworkopen.2024.40969

A.I. Chatbots Defeated Doctors at Diagnosing Illness

A small study found ChatGPT outdid human physicians when assessing medical case histories, even when those doctors were using a chatbot.

 Listen to this article · 9:25 min [Learn more](#)



AI is Now Operational

- AI is being implemented in the workplace as an advanced IT system, **without a data scientist necessarily in the loop**
- Examples include **Office 365 message management/drafting (Co-pilot), minute taking from meetings**, secure ChatGPT access like CHIRP
- Responsible AI takes a new meaning, we must ensure that **implementations are effective, safe, legal, ethical, and equitable**

Timely Access to Advanced AI is critical for NIH Research

- LLMs are causing disruption, both opportunities and challenges, throughout the world
- In biomedicine, construction of foundation models from vast biotechnology data sets is a considerable activity and LLMs are being deployed in healthcare settings
- IRP access to these methods is critical for our role

Considerations

- Operational AI tools are advanced IT systems and must be treated as such, including going through IT/cybersecurity governance.
- They must be secure and approved for the type of data being shared with them (e.g. confidential federal information).
- Operational Use Cases are reported to NIH through a data call in response to M-24-10.
- Rights and/or safety impacting AI tools should be flagged.
- In my opinion, any use of an AI that will result in a different work product of an individual than if the individual did their work and did not use an AI, it's use should be considered and evaluated carefully.

CHIRP is an example of an advanced IT system providing operational AI in support of research

- A great collaboration between the IRP, ODSS, enterprise computing, and others. A model for others.
- I expect this will continue the deployment of systems like this to support researchers
- I want to acknowledge the many individuals who contributed to make CHIRP successful including, but not limited to, Yang Fann, Nick Weber, Susan Gregurick, and Maureen Falvella

NIH Operational AI Systems ('Use Cases' from Data Call)

77 of the 195 use cases are non-retired, operational AI not qualified as R&D or productivity COTS. Of the 77:

27

Relates to Grants Management

16

Include a Chatbot

68

Include natural language processing (inclusive of chatbots)

97

Of the 195 are R&D

9

of the 195 use cases are non-retired productivity COTS

12

Are retired use cases



Center for
Information
Technology

Today's Information Technology,
Tomorrow's Cyberinfrastructure

Thank You!