

Abstract

We've moved way beyond the old days of building discovery, recommendation, decision support, and other AI tools using traditional ML and pattern recognition techniques. The future of universal personal assistance for discovery and learning is upon us. How will multimodality image, video, and audio understanding, and reasoning abilities of large foundation models change how we build these systems? I will shed some initial light on this topic by discussing 3 trends: First, the move to a single multimodal large model with reasoning abilities; Second, the fundamental research on personalization and user alignment; Third, the combination of System 1 and System 2 cognitive abilities into a single universal assistant.

Biography

Ed H. Chi is VP of Research at Google DeepMind, leading machine learning research teams working on large language models (from LaMDA leading to launching Bard/Gemini), and universal assistant agents. With 39 patents and ~200 research articles, he is also known for research on user behavior in web and social media. As the Research Platform Lead, he helped launched Bard/Gemini, a conversational chatbot experiment. His research also delivered significant improvements for YouTube, News, Ads, Google Play Store at Google with >950 product landings and ~\$10.4B in annual revenue since 2013. Prior to Google, he was Area Manager and Principal Scientist at Xerox Palo Alto Research Center's Augmented Social Cognition Group in researching how social computing systems help groups of people to remember, think and reason. Ed earned his 3 degrees (B.S., M.S., and Ph.D.) in 6.5 years from University of Minnesota. Inducted as an ACM Fellow and into the CHI Academy, he also received a 20-year Test of Time award for research in information visualization. He has been featured and quoted in the press, including the Economist, Time Magazine, LA Times, and the Associated Press. An avid golfer, swimmer, photographer and snowboarder in his spare time, he also has a blackbelt in Taekwondo.



