

# Towards Building Next-Generation AI Assistants

Date: 23 May 2024 (Thursday)

Time: 2:45pm - 3:45pm



Seminar Link: <https://cityu.zoom.us/j/91716775507>

## ABSTRACT

As technology continues to advance, the incorporation of artificial intelligence (AI) into our daily lives has become increasingly prevalent. One notable manifestation of this trend is the emergence of AI assistants, which serve a myriad of roles - from virtual personal assistants managing routine tasks, to social companion chatbots for entertainment, to digital consultants for recommendations. Yet, ensuring the trustworthiness of these AI assistants remains a critical challenge. They may produce irrelevant, non-factual, hallucinated, impolite, and even harmful responses, posing significant risks to their deployment in real-world applications. In this talk, I will share my research on building safe, user-centric AI assistants, with a particular focus on accurately modeling user intents and states. The underlying principle is that only by accurately capturing user needs can these AI assistants take right actions and provide reliable responses. Specifically, my research is centered around three key aspects: 1) user state tracking to maintain a structured and accurate representation of user needs; 2) user question rewriting to address ambiguities and ellipses in user inputs; and 3) user satisfaction estimation to evaluate user engagement and prevent unsatisfactory responses.



## Mr Fanghua YE

### GUEST SPEAKER'S PROFILE

Mr Fanghua YE is currently pursuing a fully-funded PhD in Computer Science under the supervision of Professor Emine Yilmaz and Professor Jun Wang at University College London (UCL). His research interests span a diverse range of fields, including conversational AI, large language models, trustworthiness, information retrieval, and social network analysis. His research outcomes have been published in top-tier conferences such as ACL, EMNLP, WWW, NeurIPS, ICDM, and SIGMOD. He has extensive academic and industry experience, having served as a research assistant at the Chinese University of Hong Kong and the National University of Singapore, as an area chair of EMNLP 2023, and as a research intern at Amazon Science and Tencent AI Lab.