Department of Computer Science
香港城市大學
City University of Hong Kong

COMPUTER SCIENCE COLLOQUIUM

Building Multimodal Intelligent Agents as Human Copilots

SPEAKER **Dr. Kevin Qinghong LIN**

Postdoctoral Researcher
University of Oxford

DATE 27 Jan, 2026 (Tue)
TIME 2:00 PM - 3:00 PM
VENUE CS Seminar Room, Y6405, 6th Floor, Yellow Zone, Yeung Kin Man Academic Building, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon Tong

ABSTRACT

Building intelligent assistants that can solve real-world human tasks has long been a central goal of artificial intelligence. Such assistants are inherently multimodal, seamlessly integrating visual and linguistic information, and should be able to observe, reason, and act under human instruction and in alignment with human goals. This raises a fundamental challenge: how can we develop agents that truly combine these interconnected capabilities? A promising path forward is to learn directly from human experience, by observing how humans demonstrate and solve diverse tasks. I decompose this agent learning problem into three core dimensions: (i) Perception, enabling agents to interpret rich multimodal observations; (ii) Reasoning, allowing them to model deeper context and causality; and (iii) Action, supporting meaningful, goal-directed interaction with dynamic environments. Beyond building general-purpose agents, it is equally important to ask how such systems can be deployed as specialist human copilots—acting as co-scientists that accelerate research workflows (e.g., creating posters and presentation videos), or as tutors that enhance learning and education. These efforts point toward a new generation of AI assistants that become effective collaborators in everyday life.

BIOGRAPHY

Dr. Kevin Qinghong LIN is a Postdoctoral Researcher at the University of Oxford, working with Prof. Philip Torr. Kevin received a Ph.D. from the National University of Singapore, working with Prof. Mike Shou. His research focuses on vision-language, video understanding, and intelligent agents. His work has been published in leading venues such as CVPR, ICCV, ECCV, and NeurIPS, and was selected for CVPR 2025 Doctoral Consortium. Kevin received awards including the PREMIA 2023 Best Student Paper Gold Award, CVPR 2024 EgoVis Distinguished Paper Award, and NeurIPS 2024 Open-World Agents workshop Outstanding Paper Award. He was an Outstanding Reviewer at CVPR 2024 and a Top Reviewer at NeurIPS 2024, and an Area Chair for NeurIPS 2025. He won first place in the EPIC-Kitchens and Ego4D challenges at CVPR 2022. Kevin is an active open-source contributor with over 10K GitHub stars. He leads ShowUI, an agent that learns to use computers, which has been downloaded more than 240,000 times. Kevin brings industry experience from Tencent, Meta AI, Meta Reality Labs, and Microsoft Research.

All are welcome!

In case of questions, please contact Prof. MA Jiawei at phoenix.jma@cityu.edu.hk, or visit the CS Departmental Seminar Web at https://www.cs.cityu.edu.hk/events/cs-seminars/recent-cs-colloquiums.