CURRENT TOPICS IN PRIVACY SEMINAR

CarnegieBosch INSTITUTE

The Block Center FOR TECHNOLOGY AND SOCIETY

Carnegie Mellon University privacy ENGINEERING

Hank Lee - HCII

Title

Privacy in the age of AI: What has changed and what should we do about it?

Presentation abstract

How does AI change privacy? Are the designers, engineers, and technologists who create AI technologies equipped to recognize and mitigate the unique privacy risks entailed by the AI products and services they create? Addressing these questions is crucial to steer the development of AI products and services toward their promise and away from privacy invasions. In this presentation, I will detail our research on privacy in AI. I will begin by introducing a taxonomy of AI privacy risks we created, highlighting how AI changes the landscape of privacy by introducing risks not previously accounted for and amplifying the existing ones. Following this, I will discuss insights from an interview study with 35 AI practitioners, which reveal their practices of AI privacy work and the barriers therein — awareness, motivation, and ability. Lastly, I will introduce an ongoing effort to develop Privy, an interactive AI privacy risk identification tool. The tool assists AI practitioners in identifying and prioritizing privacy risks specific to the capabilities and requirements of their AI products, promoting more informed and responsible AI development.

Bio

Hank Lee is a PhD student at the Human-Computer Interaction Institute at Carnegie Mellon University, advised by Professors Sauvik Das and Jodi Forlizzi. His research lies at the interaction of usable privacy security, human-computer interaction (HCI), and human-centered AI. He studies and builds tools that enable practitioners to identify, reason about, and mitigate AI-entailed privacy risks during the development of consumer AI products. His work has been published in leading Privacy & Security and HCI venues, including IEEE S&P, USENIX Security, CHI, TOCHI, and the International Journal of Human-Computer Studies (IJHCS), and has received a Best Paper Award at CHI (2024) and a Distinguished Paper Award at USENIX (2024).

TUESDAY, FEBRUARY 11TH