論文研討Seminar

Hacking the Brain of AI

Time: 4/15(Wed) 15:30-17:20
Location: EDB27
Speaker: 陳仲寬

奧義智慧資安研究處長
台灣駭客協會理事

【講題大綱】

當各種語言模型的應用逐漸興盛，許多日常任務也逐漸被 Agentic AI 所取代，但箇中風險卻時常被大家所忽略。在這場議程中，我們深入探討過往 Agentic AI 系統，如：OpenClaw、CrewAI、LangGraph 的漏洞與資安事件的實際案例，並更廣泛地建構 Agentic AI 生態系中的各種攻擊面，並以紅隊的思維，了解如何對其進行有效的測試。我們將探討從 AI 模型、系統到最終產品的完整測試流程，建立信任的基礎。另一方面，攻擊者如何濫用 AI 工具，也是大家時常忽略的一個方向，攻擊者也可以使用 LLM 工具來最佳化攻擊的流程，本議程也將從實際觀測到的案例開始，深入說明攻擊者可以如何濫用 LLM 來進行攻擊。

【講者簡介】

陳仲寬 (CK) 現為奧義智慧資安研究處長，也是台灣駭客協會理事。畢業於國立交通大學網路安全實驗室博士班，專注於研究網路攻擊與防禦、機器學習、軟體漏洞、惡意程式分析等領域。曾發表多篇學術期刊與研討會技術文章。另也曾於多個國內外技術研討會上發表演講，如 BlackHat、HITCON、HITB、SANS DFIR Summit、CodeBlue、FIRST 以及 VXCON。目前亦擔任東吳大學兼任助理教授、HITCON 審稿委員會的主席和資安社群 CHROOT 的成員。

國立陽明交通大學
NATIONAL YANG MING CHIAO TUNG UNIVERSITY