Mahmoud E. Shabana

Associate AI Security Researcher - Carnegie Mellon University Software Engineering Institute, Pittsburgh, PA
mshabana@andrew.cmu.edu — +1 (646) 427-8840 — LinkedIn — SEI Author Page

RESEARCH INTERESTS

Adversarial Machine Learning, AI-driven System Security, Mechanistic Interpretability for Model Safety, Automated Vulnerability Discovery

EDUCATION

New York University Tandon School of Engineering, Brooklyn, NY — September 2022 — May 2024
Master of Science in Cybersecurity — Cumulative GPA: 3.90/4.00

Monmouth University, West Long Branch, NJ — September 2020 — January 2022
Master of Science in Software Engineering — Cumulative GPA: 3.68/4.00

SKILLS

- **Active Clearance:** TS/SCI
- **Programming:** Python, C/C++, JavaScript, SQL
- **Software:** PyTorch, Tensorflow, Keras, SciKit-Learn, Pandas, Pydantic-AI, Langchain, vLLM, Docker, AWS
- **Security Tools:** PyRit, Garak, CleverHans, Metasploit, Pwntools, Ghidra, IDA Pro, WinDbg, OllyDbg, AFL++, Symbolic Execution (Angr), Wireshark, Binary Intermediate Representations (LLVM, MLIR, PTX)
- **Certifications:** Certified AI Penetration Tester - Red Team (CAIPT-RT), OSCP (In-Progress), HackTheBox CPTS (In-Progress)

RESEARCH EXPERIENCE

Carnegie Mellon University Software Engineering Institute — Pittsburgh, PA
Principal Investigator & Associate AI Security Researcher — July 2024 — Present

- Lead a team of six researchers and engineers as Principal Investigator (PI) to reverse engineer deep learning executables extracted from AI-enabled edge systems and reconstruct low-level AI artifacts to High-level serialized AI model formats (ONNX and PyTorch)
- Directed a team of four researchers in collaboration with CMU to deliver an AI Red team auto-evaluation framework that utilizes RAG-based LLM evaluation to better identify and grade security risks in AI systems
- Prototyped and deployed a defensive cyber framework for automated experimentation of various Deep Learning and ML-based solutions for diverse malware detection to a DoD agency
- Designed a hierarchal agentic reverse engineering platform that utilizes static and dynamic tools like GhidraMCP, Windbg MCP, and Vagrant VMs for autonomous binary summarization and sandboxing
- Conducted vulnerability research of AI artifacts in modern anti-virus solutions to develop bypasses to behavioral and signature-based detection utilizing UEFI parser and custom model binary analysis tools
- Selected as technical committee member for Malware Technical Exchange Meeting (MTEM) 2026 to plan, organize, and direct conference events and presentations
- Nominated for Newcomer of the Year award in 2025, a Software Engineering Institute (SEI) wide award highlighting key contributors among new-hires in their first year at the SEI

New York University Tandon - Center for Cybersecurity — Brooklyn, NY
Graduate Student Research Assistant - Dr. Brendan Dolan-Gavitt — September 2023 — May 2024

- Fine-tuning open-source Large Language Models (LLMs) to conduct code-evaluation on decompiled output from Ghidra and IDA pro
- Established a novel benchmark dataset of vulnerable software programs to rigorously evaluate the security assessment capabilities of fine-tuned code LLMs
- Cross-examined performance to existing code-based LLMs, such as StarCoder, CodeLlama, WizardCoder, and Replit's custom LLM
- Presented implementation and findings to Carnegie Mellon Software Engineering Institute - CERT Division

United States Cyber Command - Cyber Recon Research Program — Brooklyn, NY
Graduate Researcher - Dr. Travis Trammell — November 2022 — April 2023

- Gathered intelligence on foreign Advanced Persistent Threat (APT) organizations to analyze potential attack vectors against US Election Infrastructure through influence campaigns

- Monitored Iranian APT social media activity with sock puppet accounts, Tails OS, and ProtoVPN for anonymity
- Documented tactics, techniques, and procedures (TTP) of Iranian APTs that were used on social media platforms to bolster influence and recruit individuals to their political cause
- Analyzed the emerging threat of generative AI (e.g., Stable Diffusion) in amplifying Iranian state-sponsored influence operations, focusing on the stable creation of synthetic media for sock puppet personas and propaganda
- Presented our findings to cyber professionals in industry and military intelligence agencies at US Cyber Command

Monmouth University - Summer Research Program West Long Branch, NJ
Research Assistant May 2018 — September 2018

- Analyzed the efficacy of photogrammetric point clouds for infrastructure damage assessment by generating 3D models from drone imagery using Python and OpenCV
- Conducted a comparative analysis of feature detection algorithms, identifying SIFT as the most accurate method over ORB and SURF for matching keypoints in aerial photographs
- Pioneered a novel visualization method using a 3D interactive scatter plot of drone GPS coordinates to preserve the high resolution of original photographs, after determining that point cloud reconstructions lacked sufficient detail for damage assessment
- Co-authored and published findings in the peer-reviewed proceedings of the 2019 International Conference on Computer Vision and Graphics (AISC 943)

WORK EXPERIENCES

Johns Hopkins Applied Physics Laboratory (APL) - Asymmetric Operations Howard, MD
Reverse Engineering Intern May 2023 — August 2023

- Generalized image classification of printed circuit board components by diversifying training and testing datasets with data augmentations refactoring Keras ImageDataGenerator
- Streamlined testing of newly trained ML models by implementing a backend CLI application using Python Typer library
- Researched and replicated various CVEs to exploit backdoors on vendor-provided embedded systems to force unintended behavior and allow for privileged commands to disable certain features on the embedded hardware
- Captured and reversed CAN Bus messages to map physical components to CAN IDs and retransmit messages for remote execution utilizing SavvyCAN and a CAN BUS sniffer

Broadridge Financial Solutions - FXL Product Dev Team Newark, NJ
Software Developer January 2022 — July 2022

- Exposed and documented company-owned API functions for external users to connect custom frontends to company backend services using XML and Swagger YAML
- Developed and tested the design and implementation of a Full-Stack desktop trading application to meet business and technical requirements for multi-million dollar clients
- Configured interface adapters to store financial records from third-party applications using C#, .NET HTTP Handlers, and SQL queries

AWARDS

SEI Internal Project Proposal Grant 2025
Research project proposal accepted and appointed by SEI Chief Technical Officer and technical committee as Principal Investigator (PI) for my own 2-year long research project. Awarded a Congressional Department of Defense (DoD) grant of $1.8 million in funding

CyberCorp Scholarship for Service (SFS) Recipient 2022
One of 12 graduate students at NYU Tandon School of Engineering selected for a full academic scholarship

Bill Boylan Student-Athlete Award Recipient 2021
Selected among all senior student-athletes for demonstrating excellence in leadership, sportsmanship, and scholarship

Chi Alpha Sigma Honor Society Recipient 2020
Two-time award recipient for outstanding academic and athletic achievement