AI Institute for Agent-based Cyber Threat Intelligence and Operation

https://action.ucsb.edu

ACTION Year 3 Annual Report: Summary

Reporting period: June 2025 to May 2026 (Published November 2026)

The NSF AI Institute for Agent-based Cyber Threat Intelligence and Operation (ACTION) develops AI-enabled security agents that work with human defenders across the full cybersecurity life cycle: assessing vulnerabilities, detecting attacks, attributing them, and responding and recovering. ACTION is led by UC Santa Barbara with 11 partner universities and is supported by the U.S. National Science Foundation, with additional support from the Department of Homeland Security and IBM.

Research. In its third year, ACTION advanced 20 cross-campus research projects across four foundational AI thrusts, four cybersecurity thrusts, and its shared testbed, GATE. Each project paired faculty from at least two campuses. Highlights include:

- Game-theoretic learning algorithms with provable guarantees that adapt to adversaries ranging from slow-moving to fully adaptive, extended to defending cyber-physical systems.
- An adaptive human-AI decision-making framework published at AAAI 2026, with related work receiving an honorable mention award at CHI 2026.
- New methods for evaluating the security of AI-controlled robots, including two benchmarks for embodied AI.
- Research using reinforcement learning to decide how to respond to network intrusions, evaluated in a real-world pilot deployment.

Open tools and datasets. ACTION researchers released TOUCAN, developed with IBM, a dataset of 1.5 million examples from nearly 500 real-world tool environments for training AI agents. They also released CVE-GENIE, a multi-agent system that reproduced working exploits for about 51% of recent real-world vulnerabilities at an average cost of $2.77 each, and six benchmarks spanning AI safety, reasoning, and agent security. All are publicly available.

Competitions. A University of Illinois Urbana-Champaign team including ACTION researchers won first place and a $250,000 prize in the Amazon Nova AI Challenge. Team Shellphish, founded by ACTION Director Dr. Giovanni Vigna, finished fifth in DARPA's AI Cyber Challenge, with the highest patching success rate among the finalists.

Education and workforce. ACTION received a $2 million NSF education supplement to build a modular AI and cybersecurity curriculum with the University of Washington, Georgia Tech, and Western Washington University. Other education highlights include:

- zero2hero, developed by UC Santa Barbara and Norfolk State University, which teaches cybersecurity and AI to students with no prior background through 40 new hands-on challenges.
- Summer research experiences at UC Santa Barbara and Rutgers, including student work presented at a NeurIPS 2025 workshop and AAMAS 2026.
- Capture the Flag training camps at Norfolk State University for 25 high school students and seven teachers, and for 22 college students, including five U.S. veterans.
- iCTF 2026 "Battle of the Agents," the first major international Capture the Flag competition with agentic challenges, with high school and undergraduate divisions and participants from four countries.

Impact and engagement. ACTION's publications exceed 300 manuscripts, with more than 8,000 citations as of May 2026. Dr. Vigna gave the opening keynote at ACM CCS 2025 and briefed the federal Cyber Security and Information Assurance Interagency Working Group. ACTION researchers share findings each month with the Department of Homeland Security. Through IBM, ACTION was introduced to the Coalition for Secure AI. Dr. Vigna was also interviewed by the Government Accountability Office for its forthcoming work for Congress on AI policy.

Looking ahead. In Year 4, ACTION's research centers on a shared scenario: AI agents defending water and chemical plants against coordinated cyberattacks. Priorities include defining the Institute's AI stack, opening GATE to outside researchers, and expanding education partnerships.

The full report and ACTION's Strategic and Implementation Plan are available at https://action.ucsb.edu/knowledge-transfer. Please direct any questions to action@cs.ucsb.edu.

This material is based upon work supported by the U. S. National Science Foundation under grant no. 2229876 and is supported in part by funds provided by the U. S. National Science Foundation, by the Department of Homeland Security, and by IBM.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U. S.National Science Foundation or its federal agency and industry partners.