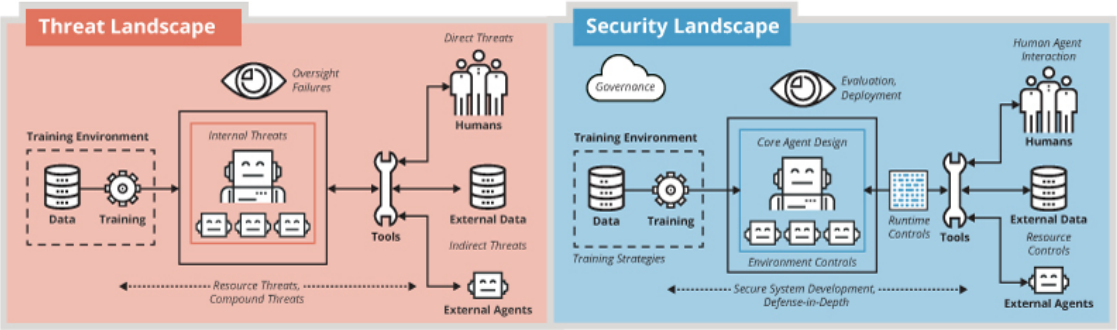
SEI Bulletin 40 years



Bridging Research and Practice in LLM Agent Security

December 3, 2025 — Large language model (LLM) agents perceive their environment and take actions in the real world, opening new attack surfaces and pathways to harm. But standard practices for securing LLM agents remain fragmented, write SEI researchers in a new paper. They describe their systematic review of agentic LLM security, which informed a proposed framework for developing risk-mitigation strategies.

“Large language model agents are rapidly transitioning from research prototypes to deployed systems, raising new and urgent security challenges,” write the authors of *Bridging Research and Practice in LLM Agent Security*.

They go on to say, “A consensus-driven synthesis of security best practices that spans both research and practice is essential to translate emerging insights into actionable guidance. To address this gap, we systematize knowledge across academia, industry, and real-world deployments to provide a unified framework for the secure design, development, and operation of LLM agents.”

The paper includes three recommendations for security research on LLM agents and demonstrates how to use the framework in a systematic risk

assessment.

Read the paper »

SEI News

Secure Software by Design 2025 Presentations Available
The August event featured presentations on data bills of materials, vulnerability detection, AI, and more.

SEI Study Analyzes Applicability of Security and Zero Trust Principles to Weapon Systems
The study explores the risks and tradeoffs when adapting enterprise-IT security and zero trust principles to weapon systems.

See more news »

Latest Blogs

AI-Powered Memory Safety with the Pointer Ownership Model *(David Svoboda, Lori Flynn)*
Bugs related to temporal memory safety, such as use-after-free and double-free vulnerabilities, are challenging issues in C and C++ code. The SEI is automating C code security with AI-powered memory safety.

How to Align Security Requirements and Controls to Express System Threats *(Elias Miller, Matthew Sisk)*
This blog post presents a method that combines information about security requirements, controls, and capabilities with analysis regarding cyber threats to enable more effective risk-guided system planning.

See more blogs »

Latest Podcasts

Orchestrating the Chaos: Protecting Wireless Networks from Cyber Attacks
Joseph McIlvenny and Michael Winter discuss common radio frequency (RF) attacks and investigate how software and cybersecurity play key roles in preventing and mitigating these exploitations.

From Data to Performance: Understanding and Improving Your AI Model
Drift in data and concept, evolving edge cases, and emerging phenomena can undermine the correlations that AI classifiers rely on. Linda Parker Gates, Nicholas Testa, and Crisanne Nolan discuss a new tool to help improve AI classifier performance.

See more podcasts »

Latest Publications

Bridging Research and Practice in LLM Agent Security *(Keltin Grimes, Julie Lawler, Robert C. Garrett, Emil Mathew, Marco Christiani, Sara Kingsley, Zhiwei Steven Wu, Nathan M. VanHoudnos)*
This systematic review discusses academic surveys, grey literature sources, and real-world case studies on securing LLM agents.

Minimally Viable Architecture: Architecture Early in Development *(Manuel Rosso-Llopart)*
This technical note explores MVAs and offers guidance on what teams should do with an MVA to get their certificate to field and authorization to operate.

See more publications »

Latest Videos

How to Address the Problem of Poorly-Defined Requirements in Software System Design

Lori Flynn and Lyndsi Hughes offer a solution to the problem of poorly defined requirements in system design that can lead to software flaws, cost and time overruns, and stakeholder dissatisfaction.

5 Essential Questions for Implementing the Software Acquisition Pathway and the Tools to Tackle Them
Eileen Wrubel and Brigid O'Hearn discuss the launch of the Software Acquisition Go Bag, the SEI's latest effort to help defense programs deliver software-enabled capability through data-driven insights, software engineering research, and acquisition science.

See more videos »

Upcoming Appearances

Hawaii International Conference on System Sciences (HICSS), January 6-9, 2026
SEI researchers are chairing the mini-track "AI-Driven Program Analysis and Software Synthesis" in the HICSS 59 Software Technology Track.

AIAA SciTech Forum 2026, January 12-16, 2026
Visit the SEI at booth 106.

AFCEA WEST 2026, February 10-12, 2026
Visit the SEI at booth 817.

See more opportunities to engage with us »

Upcoming Training

Software Architecture Design and Analysis
January 26-29, 2026 (SEI Live Online)

Insider Risk Management: Measures of Effectiveness
February 18-20, 2026 (SEI Live Online)

Insider Threat Program Manager: Implementation and Operation
February 24-26, 2026 (SEI Live Online)

E-learning - CERT Artificial Intelligence (AI) for Cybersecurity Professional Certificate

E-learning - Introduction to Artificial Intelligence (AI) Engineering

E-learning - Effective Communication of Technical Concepts

Online registration is now available for January-December 2026 public courses, both live online and in person.

See more courses »

Employment Opportunities

Business Development Manager

Reverse Engineer Researcher

Senior AI Security Researcher

All current opportunities »