

Hacking AI: Angriffe und Abwehr für KI-Systeme und Infrastruktur

Lernen von den Profis – Ihre Trainer sind Ahmad Abolhadid und Matthias Ortmann.

Kursbeschreibung

Dieses Training ist eine praxisorientierte Einführung in das Angreifen und Verteidigen von agentischen KI-Systemen und Anwendungen auf Basis von Large Language Models (LLMs). Es werden verschiedene Techniken von Prompt-Injection-Angriffen vorgestellt, um das Verhalten von KI-Systemen zu manipulieren. Im Laufe des Trainings werden die Teilnehmer mit neuen Konzepten und Komponenten wie Guardrails und MCPs vertraut gemacht. Die Teilnehmer lernen, wie man diese angreift und wie man sie sicher implementiert. Der Kurs enthält Elemente der Gamifizierung, die das Lernen interaktiv, interessant und unterhaltsam gestalten.

Neben den offensiven und defensiven Techniken von LLM-Anwendungen behandelt das Training auch die Infrastrukturseite von LLM-Systemen. Anhand ausgewählter Komponenten der Plattform, die diesen Workshop betreibt – darunter Model Serving, Chat-Oberflächen, RAG-Pipelines und API-Gateways – lernen die Teilnehmer, wie häufige Standardkonfigurationen ausgenutzt werden können und wie man sie absichert.

Die Lab-Umgebung ist einfach über den Browser zugänglich. Daher ist keine Vorinstallation nötig und keine Vorerfahrung im Hacking erforderlich.

03. - 04. Dezember 2026, LIVE ONLINE
17. - 18. März 2027, LIVE ONLINE

Diese Veranstaltung wird als Weiterbildung bei Rezertifizierungsmaßnahmen von verschiedenen Instituten anerkannt.

Hacking AI: Angriffe und Abwehr für KI-Systeme und Infrastruktur

M 71

Lernen von den Profis

/// In diesem Workshop lernt ihr

Angreifen und Verteidigen von LLM-Anwendungen:

- Wie LLMs funktionieren und warum ihre Architektur von Grund auf unsicher ist
- OWASP Top 10 Schwachstellen von LLM-Anwendungen und KI-Agenten
- Direkte und indirekte Prompt-Injection-Techniken
- Abgreifen von vertraulichen Daten aus KI-Anwendungen
- Verstehen, Erfassen und Absichern von System-Prompts
- Implementierung von Guardrails zum Schutz von KI-Systemen sowie Techniken zu deren Umgehung
- Das Model Context Protocol (MCP) und seine Annahmen zum Vertrauensmodell
- Klassische Schwachstellen wie SQLi und RCE in KI-Systemen
- Defensive Muster für LLM-Pipelines und MCP-Server

Aufbau und Absicherung selbstgehosteter LLM-Infrastruktur:

- Selbstgehostete LLM-Komponenten im Überblick
 - Model Serving
 - Chat-Oberflächen
 - RAG-Pipelines
 - API-Gateways
- Wann und warum selbst hosten?
 - Datensouveränität
 - Kosten
 - Anpassbarkeit
 - Offline-Betrieb
 - Air-Gapped-Szenarien
- Warum reproduzierbare, deklarative Infrastruktur wichtig ist, wenn man auf von Natur aus nicht-deterministischen KI-Systemen aufbaut
- Praktische Abwägungen beim Selbsthosting: vom Single-GPU-Entwicklungssetup bis zur Multi-Node-Plattform
- Absicherung von LLM-Infrastrukturkomponenten:
 - Inferenz-Endpunkte
 - Vektordatenbanken
 - Agenten-Tools
 - Code-Interpreter

HM TRAINING SOLUTIONS ON-SITE SERVICE

Alle HM Training Solutions Seminare stehen auch firmenintern zur Verfügung. Sie können auf den Bedarf Ihrer Organisation zugeschnitten werden. Der Kurs kann auf Wunsch firmenintern in englischer Sprache gegen Aufpreis durchgeführt werden. Weitere Details erhalten Sie unter der Telefonnummer +49 (0) 6022 508 200.

Hacking AI: Angriffe und Abwehr für KI-Systeme und Infrastruktur

M 71

Lernen von den Profis

/// Was wir für euch vorbereitet haben

- Eine vollständige KI-Plattform, die die im Workshop besprochenen Infrastrukturmuster veranschaulicht.
- Eine persönliche Lab-Umgebung pro Teilnehmer mit Open WebUI, einem live laufenden Open-WebUI-Pipelines-Server und mehr – alles im Browser.
- Viele praktische Übungen zum Einüben der im Training erklärten Angriffs- und Abwehrtechniken.
- Deploy-Break-Harden-Infrastruktur-Szenarien: echte Fehlkonfigurationen aus selbstgehosteten LLM-Komponenten.

/// Wer kann an diesem Workshop teilnehmen

Es sind keine Vorkenntnisse in LLMs oder KI-Sicherheit erforderlich. Der Workshop ist praxisorientiert und beginnt bei den Grundlagen. Er ist geeignet für:

- Penetrationstester und Red Teamer, die in Assessments auf LLM-integrierte Systeme treffen
- Application Security Engineers, die an KI- oder agentischen Produkten arbeiten
- Entwickler, die auf der Open-WebUI-API oder ähnlichen Backends aufbauen und die Risiken verstehen möchten
- Ingenieure und Architekten, die LLM-Infrastruktur für ihr Team oder ihre Organisation selbst hosten möchten
- Alle, die sich für die offensive Seite der KI-Sicherheit interessieren

/// Voraussetzungen

- Ein Laptop mit einem modernen Browser.
- Grundlegende Python-Lesekenntnisse.
- Die Bereitschaft, Dinge kaputtzumachen.

HM TRAINING SOLUTIONS ON-SITE SERVICE

Alle HM Training Solutions Seminare stehen auch firmenintern zur Verfügung. Sie können auf den Bedarf Ihrer Organisation zugeschnitten werden. Der Kurs kann auf Wunsch firmenintern in englischer Sprache gegen Aufpreis durchgeführt werden. Weitere Details erhalten Sie unter der Telefonnummer +49 (0) 6022 508 200.

Hacking AI: Angriffe und Abwehr für KI-Systeme und Infrastruktur

M 71

Lernen von den Profis

/// Über die Trainer

Ahmad Abolhadid ist Senior Security Analyst bei ERNW. Er verfügt über umfangreiche Erfahrung im Pentesting von Mobil- und Webanwendungen, Infrastrukturen und KI-Systemen. Er entwickelt spezialisierte Werkzeuge für Sicherheitstests und gestaltet mit Freude technische Trainings, um sein praktisches Wissen mit der Community zu teilen. Er hat einen Masterabschluss in Computer and Media Engineering der Hochschule Offenburg, Deutschland.

Matthias Ortmann ist Security Consultant bei ERNW mit Schwerpunkt auf Webanwendungssicherheit, Netzwerkschutz und der Sicherheit von LLM-integrierten Systemen. Er hat einen Masterabschluss in IT-Sicherheit der TU Darmstadt. Bei ERNW führt er Penetrationstests durch, entwickelt Sicherheitswerkzeuge und konzipiert reproduzierbare Infrastruktur für KI-Sicherheitsforschung und Trainingsumgebungen – einschließlich der selbstgehosteten LLM-Plattform, die diesen Workshop betreibt.

DETAILS ZUM ANMELDEFORMULAR

/// Drei Wege zur Anmeldung

Per Post: Bitte dieses Anmeldeformular ausfüllen und an HM Training Solutions senden.

Per E-Mail: Info@hm-ts.de

Per Webseite: <https://www.hm-ts.de>

/// Gebühren

2.690 € pro Person + 19% MwSt.

/// Bestätigungsbrief

Ihre Anmeldung bestätigen wir per Mail oder Brief. Er enthält Details über die Veranstaltung.

/// Änderungen

HM Training Solutions behält sich das Recht vor, bei Eintreten nicht vorhersehbarer Umstände das Seminar räumlich und/oder zeitlich zu verlegen, einen anderen Referenten ersatzweise einzusetzen oder die Veranstaltung zu stornieren. Weitergehende Ansprüche bestehen nicht.

/// Stornierung seitens des Teilnehmers

Bitte reichen Sie Stornierungen schriftlich per Post oder Email (info@hm-ts.de), ein. Bestätigte Anmeldungen können bis zu sechs Wochen vor Seminarbeginn kostenfrei storniert werden, danach berechnen wir die gesamte Seminargebühr. Eine Übertragung an einen Ersatzteilnehmer ist jederzeit möglich.

/// Firmeninterne Seminare

Alle Trainings von HM Solutions können auch firmenintern und zugeschnitten auf den Bedarf der jeweiligen Organisation durchgeführt werden. Weitere Informationen erhalten Sie unter der Telefon-Nr. +49 (0) 6022 508 200.

**Die Teilnehmerzahl ist begrenzt.
Wir berücksichtigen Ihre Anmeldung
in der Reihenfolge des Eingangs.**

ANMELDEFORMULAR

Hacking AI: Angriffe und Abwehr für KI-Systeme und Infrastruktur

☐ (M 71) 03. - 04. Dezember 2026, Live Online

☐ (M 71) 17. - 18. März 2027, Live Online

☐ Bitte reservieren Sie _____ Platz/Plätze zum Einzelpreis von **2.690 € + 19% MwSt.**

Herr/Frau Vorname Nachname

Funktion _____

Firma _____

Adresse _____

Postleitzahl _____ Ort _____

Land _____

Telefonnummer _____

Mobilfunknummer _____

E-Mail _____

Unterschrift _____

BUCHUNGSREFERENZ

HM71

/// Zahlung

☐ Bitte um Rechnungsstellung

Rechnungsadresse (falls nicht identisch mit obiger Anschrift).

PO-Nummer _____

/// Zusätzliche Teilnehmer

1. Herr/Frau Vorname Nachname

Funktion _____

E-Mail _____

2. Herr/Frau Vorname Nachname

Funktion _____

E-Mail _____

3. Herr/Frau Vorname Nachname

Funktion _____

E-Mail _____