COASP – Certified Offensive AI Security Professional (e-Learning)

EC-Council

Com certificação

- **Nível:**
- **Duração:** 26h

Sobre o curso

Spring Deals: Cursos EC-Council com 5% de desconto até 30 de abril

Master the Tactical Methodology to Hack LLMs and Secure Agentic AI, the Global Command for Offensive Teams

LLMs are vulnerable. Prompt injection bypasses guardrails. Data poisoning corrupts models. This credential enables you to red-team AI systems, exploit vulnerabilities in LLMs and agents, and build defenses that survive real-world attacks.

- Red-team LLMs: prompt injection, jailbreaking, and prompt chaining
- Exploit AI agents: memory corruption, tool misdirection, and checkpoint manipulation
- Master OWASP LLM Top 10 & MITRE ATLAS attack frameworks

What is the Certified Offensive AI Security Professional (COASP) certification?

COASP is EC-Council's offensive AI security program designed for cybersecurity professionals who must think like attackers and defend AI like engineers. It trains you to red-team LLMs, exploit AI systems, and defend enterprise AI before attackers do.

Attack Vectors

Prompt injection. Data poisoning. Model theft. Attackers are exploiting AI faster than security teams can learn to defend.

Your Credential

Certified Offensive AI Security Professional. Validate you can simulate attacks, find vulnerabilities, and

harden AI systems.

AI Red-Teaming Is a New Discipline

Traditional pentesting doesn't cover LLM vulnerabilities. Prompt injection, data poisoning, and model manipulation require specialized offensive skills. COASP is the first credential built specifically for AI red teamers.

The Problem:

Traditional pentesting doesn't cover LLM vulnerabilities. Security teams lack the specialized skills to exploit and defend AI systems at scale.

- Prompt injection impacts 73%+ of production AI deployment[#]
- No standardized AI red-teaming methodology
- Security teams lack LLM exploitation skills

COASP Credential Validates:

The COASP certification equips you to red-team AI systems end-to-end, from prompt injection to model exploitation. Master offensive techniques that break AI before attackers do.

- Master prompt injection, jailbreaking, and prompt chaining
- Learn OWASP LLM Top 10 & MITRE ATLAS attack chains
- Build AI defenses that survive adversarial testing

*Source: * IBM X-Force Threat Intelligence, 2024; ** Palo Alto Networks; # Resecurity, citing OWASP Top 10 for LLM Applications (2025)*

Destinatários

Who should take the COASP certification?

COASP is ideal for red-team and blue-team professionals, SOC analysts, penetration testers, AI/ML engineers, DevSecOps specialists, and compliance managers responsible for AI safety in regulated industries like finance, healthcare, and defense.

Who is COASP Ideal For

COASP is designed for security professionals who want to master offensive and defensive AI security techniques.

Offensive Security

Penetration Tester/Ethical Hacker
Red Team Operator/Red Team Lead
Offensive Security Engineer
Adversary Emulation/Purple Team Specialist

Defensive Security

SOC Analyst (Tier 2/3)/Detection Engineer
Blue Team Engineer/Threat Detection Engineer
Incident Responder (IR)/DFIR Analyst
Security Operations Manager (SOC Lead)

Threat Intelligence

Malware Analyst/Threat Researcher
Cyber Threat Intelligence (CTI) Analyst – AI Focus
Fraud/Abuse Detection Analyst (AI-enabled threats)

AI/ML Engineering

ML Engineer/Applied AI Engineer
GenAI Engineer (RAG/Agents)
AI/LLM Application Developer
MLOps/AI Platform Engineer

Security Engineering

DevSecOps/Secure DevOps Specialist
Application Security Engineer (LLM Apps/APIs)
Product Security Engineer/AI Product Security

AI Security Architecture

Secure AI Engineer/AI Security Architect
LLM Systems Engineer

Objetivos

What This Credential Validates

Verify the skills that make you the AI leader organizations need:

- Your ability to red-team AI systems
- Skills that differentiate you in prompt injection, jailbreaking, and model exploitation
- Confidence in testing enterprise AI defenses
- Industry-recognized proof of offensive AI security expertise

- Your ability to find vulnerabilities before attackers do

What Your Organization Gets

Solve the AI security crisis:

- Identify AI vulnerabilities before production deployment
- Build robust defenses against adversarial AI attacks
- Protect LLMs and AI agents from exploitation
- Clear security validation for AI investments

Pré-requisitos

Do I need prior cybersecurity experience?

Yes, this program requires foundational cybersecurity knowledge. This is not a beginner course, it’s hands-on offensive security training for professionals who already understand security fundamentals.

Metodologia

Master the offensive techniques that break AI systems before attackers do. From prompt injection to model extraction, learn to think like an adversary and defend like an engineer.

- 20+ AI Security Testing Tools
- 20+ Hands-On Offensive AI Techniques
- 30 Lab Exercises
- 15+ MITRE ATLAS Techniques

Programa

- Offensive AI and AI System Hacking Methodology
- AI Reconnaissance and Attack Surface Mapping
- AI Vulnerability Scanning and Fuzzing
- Prompt Injection and LLM Application Attacks
- Adversarial Machine Learning and Model Privacy Attacks
- Data and Training Pipeline Attacks
- Agentic AI and Model-to-Model Attacks
- AI Infrastructure and Supply Chain Attacks
- AI Security Testing, Evaluation, and Hardening
- AI Incident Response and Forensics

Offensive AI and AI System Hacking Methodology

- AI and machine learning fundamentals from an offensive security perspective
- AI attack surfaces, threat landscapes, and adversary techniques (MITRE ATLAS–aligned)
- AI system hacking methodologies, frameworks, and risk implications
- AI attack taxonomies and classification models
- Offensive AI scoping fundamentals and foundations for securing AI systems
- Overview and mapping of OWASP LLM & ML Top 10 (2025) to AI threat and governance considerations

AI Reconnaissance and Attack Surface Mapping

- Apply OSINT tools and techniques to identify and profile AI assets
- Gather intelligence from AI data sources and training pipelines
- Discover and map AI attack surfaces using publicly available intelligence
- Enumerate AI endpoints, services, APIs, and exposed parameters
- Identify and analyze AI models and vector stores from an attacker’s perspective
- Evaluate OSINT exposure and apply hardening controls to reduce risk
- Use AI threat intelligence to support continuous monitoring and defensive readiness

AI Vulnerability Scanning and Fuzzing

- Core principles of AI vulnerability assessment and threat discovery
- Tools and techniques for scanning vulnerabilities in AI models, pipelines, and deployments
- Practical fuzzing methods tailored for AI systems and model interfaces
- How to integrate scanning and fuzzing into AI security workflows for proactive risk mitigation

Prompt Injection and LLM Application Attacks

- Prompt injection and jailbreaking techniques in real-world LLM applications
- Sensitive information disclosure and system prompt leakage risks
- Improper output handling vulnerabilities and misinformation threats
- Advanced prompt-based attack techniques and exploitation strategies
- Secure LLM application design principles and defensive controls

Adversarial Machine Learning and Model Privacy Attacks

- Core adversarial machine learning attack classes
- Practical adversarial input attacks across data modalities
- Privacy, inference, and model extraction attack techniques
- Robustness, trustworthiness, and risk evaluation methods
- Defensive strategies for model privacy and resilience

Data and Training Pipeline Attacks

- AI data and training pipeline architecture and threat surfaces
- Practical data poisoning techniques and attack scenarios
- Backdoor and trojan insertion during model training
- Security measures to safeguard data and training pipelines

Agentic AI and Model-to-Model Attacks

- Agentic AI architecture and attack surface
- Excessive agency and autonomy exploitation techniques
- Cross-LLM and model-to-model attack vectors
- Denial-of-wallet risks and unbounded resource consumption
- Attacks targeting AI workflows and orchestration layers
- Defensive strategies for securing agentic AI applications

AI Infrastructure and Supply Chain Attacks

- AI infrastructure components and system integration architectures
- Vulnerabilities in AI systems, frameworks, and deployment pipelines
- Abuse of tools, plugins, and APIs in AI-enabled applications
- AI supply chain threats and dependency risks (deep dive)
- Hardening strategies for AI infrastructure and supply chains

AI Security Testing, Evaluation, and Hardening

- AI security testing methodologies and evaluation techniques
- Red team frameworks for offensive AI assessment
- AI vulnerability identification, validation, and risk reporting
- Security hardening and mitigation best practices for AI systems

AI Incident Response and Forensics

- Detect and respond to AI-specific security incidents
- Collect and analyze AI logs, telemetry, and digital evidence
- Analyze root causes in post-incident analysis