EC-Council
Building A Culture Of Security

C|OASP
Certified Offensive AI Security Professional

La credencial Certified Offensive AI Security Professional (C|OASP) está diseñada para transformar a los profesionales en especialistas en seguridad ofensiva de IA listos para el entorno empresarial. Este programa se enfoca en técnicas reales de ataque a sistemas de IA y en metodologías de adversarios, preparando a los profesionales para identificar, explotar y remediar vulnerabilidades en LLMs, sistemas de IA agéntica y pipelines de IA, utilizando marcos de la industria como OWASP LLM Top 10 y MITRE ATLAS.

La realidad de la seguridad en IA

- 87% de las organizaciones reportan ataques impulsados por IA (IBM).
- Incremento del 890% en tráfico de GenAI (Palo Alto Networks).
- La inyección de prompts impacta a más del 73% de los despliegues de IA en producción (Resecurity).

El problema del mercado: Por qué las organizaciones son vulnerables a ataques de IA

- Los pentesters no saben cómo explotar LLMs o agentes de IA.
- No existe una metodología estandarizada para red-teaming de IA.
- Los escáneres tradicionales de vulnerabilidades no detectan fallas específicas de IA.
- Los equipos SOC no pueden detectar ataques impulsados por IA.
- Los arquitectos de seguridad no entienden los modelos de amenazas de IA.

Nuestra solución: Certified Offensive AI Security Professional (C|OASP)

La certificación de EC-Council para profesionales responsables de asegurar sistemas de IA, no solo de probarlos. Proporciona un camino claro para identificar riesgos, mitigar vulnerabilidades y fortalecer las defensas de la IA.

Capacitación:

- **Título del curso:** Certified Offensive AI Security Professional (C|OASP).
- **Versión:** 1.
- **Duración:** 5 días.
- **Prerequisito:** 2 años de experiencia en ciberseguridad.
- **Modalidad:** Instructor-led Training (ILT), iWeek (en línea sincrónico), iLearn (en línea asincrónico).

Examen:

- **Título del examen:** Certified Offensive AI Security Professional (C|OASP).
- **Código del examen:** 312-52.
- **Número de preguntas:** 70 (65 de opción múltiple + 5 prácticas basadas en desempeño).
- **Duración:** 6 horas.
- **Disponibilidad:** ECC Exam Portal.
- **Puntaje de aprobación:** 70–80%.
- **Formato:** Preguntas de opción múltiple y preguntas basadas en desempeño.

Módulos del programa C|OASP

- Metodología de hacking ofensivo de IA y sistemas de IA.
- Reconocimiento de IA y mapeo de superficie de ataque.
- Escaneo de vulnerabilidades de IA y fuzzing.
- Inyección de prompts y ataques a aplicaciones LLM.
- Aprendizaje automático adversarial y ataques a la privacidad de modelos.
- Ataques a datos y pipelines de entrenamiento.
- IA agentiva y ataques de modelo a modelo.
- Infraestructura de IA y ataques a la cadena de suministro.
- Pruebas de seguridad en IA, evaluación y hardening.
- Respuesta a incidentes y forense en IA.

¿Para quién es ideal?

C|OASP está diseñado para profesionales de seguridad que desean dominar técnicas ofensivas y defensivas de seguridad en IA.

1. SEGURIDAD OFENSIVA

- Pentester / Hacker Ético
- Operador Red Team / Líder Red Team
- Ingeniero de Seguridad Ofensiva
- Especialista en Emulación de Adversarios / Purple Team

2. SEGURIDAD DEFENSIVA

- Analista SOC (Nivel 2/3) / Ingeniero de Detección
- Ingeniero Blue Team / Detección de Amenazas
- Respondedor de Incidentes (IR) / Analista DFIR
- Gerente de Operaciones de Seguridad (Líder SOC)

3. INTELIGENCIA DE AMENAZAS

- Analista de Malware / Investigador de Amenazas
- Analista de Inteligencia de Amenazas (CTI) – Enfoque en IA
- Analista de Detección de Fraude/Abuso (amenazas habilitadas por IA)

4. INGENIERÍA DE IA / ML

- Ingeniero de Machine Learning / Ingeniero de IA Aplicada
- Ingeniero de IA Generativa (RAG / Agentes)
- Desarrollador de Aplicaciones de IA / LLM
- Ingeniero MLOps / Ingeniero de Plataforma de IA

5. INGENIERÍA DE SEGURIDAD

- Especialista DevSecOps / Secure DevOps
- Ingeniero de Seguridad de Aplicaciones (LLM / Apps de IA)
- Ingeniero de Seguridad de Producto / Seguridad de Productos de IA

6. ARQUITECTURA DE SEGURIDAD EN IA

- Ingeniero de IA Segura / Arquitecto de Seguridad en IA
- Ingeniero de Sistemas LLM

EC-Council
Building A Culture Of Security

C|OASP
Certified Offensive AI Security Professional

Técnicas prácticas de seguridad ofensiva en IA cubiertas

Reconocimiento multiprotocolo
Enumerar endpoints relacionados con IA en servicios REST y gRPC.

Análisis de telemetría para mapear límites de decisión de la IA
Analizar salidas del modelo para hacer ingeniería inversa de la lógica de decisión y los umbrales.

Reconocimiento de APIs
Descubrir y mapear endpoints de APIs de IA, parámetros y mecanismos de autenticación.

Reconocimiento de IA mediante "fingerprinting" del modelo
Identificar modelos de IA, versiones y configuraciones a través de análisis de comportamiento.

Ataques de transferencia, frontera (boundary) y ruido
Ejecutar ataques adversariales de caja negra en distintas arquitecturas de modelos de IA.

Ataques PGD en modelos de audio
Desplegar ataques por gradiente en modelos de clasificación y transcripción de audio.

Ataques entre LLMs (Cross-LLM)
Evaluar y explotar vectores de ataque en sistemas que operan entre múltiples LLMs.

Reconocimiento de APIs y extracción de modelos
Descubrir endpoints de APIs de IA y extraer pesos del modelo desde infraestructura expuesta.

Ataques de envenenamiento de RAG (RAG Poisoning)
Inyectar contenido malicioso en pipelines de generación aumentada por recuperación (RAG).

Roles profesionales habilitados por C|OASP

- Especialista en AI Red Team / Ingeniero de IA adversarial.
- Ingeniero de seguridad ofensiva (enfoque IA/LLM).
- Analista de seguridad en IA adversarial.
- Investigador en aprendizaje automático adversarial.
- AI Threat Hunter / Analista de seguridad en IA.
- Analista de malware y exploits de IA.
- Ingeniero de respuesta a incidentes en IA.
- Especialista en pruebas y evaluación de IA.
- verificada en IA.
- Ingeniero de IA segura / Arquitecto de seguridad en IA.
- Especialista en seguridad de MLOps / AIOps.
- Especialista en riesgos de modelos de IA.
- Ingeniero de sistemas LLM.
- Especialista en riesgos y aseguramiento en IA.
- Analista de inteligencia de amenazas (CTI) con enfoque en IA.
- Gerente de programa de seguridad en IA.
- Gerente de seguridad de producto en IA.
- Asesor / consultor en riesgos de IA.

Habilidades que C|OASP valida

- Ejecutar ataques de inyección de prompts, jailbreaking y prompt chaining.
- Hacer red-team a agentes de IA: corrupción de memoria, desvío de herramientas y manipulación de checkpoints.
- Aplicar los frameworks OWASP LLM Top 10 y MITRE ATLAS.
- Realizar ataques adversariales de ML: data poisoning y extracción de modelos.
- Construir reglas de detección y estrategias de hardening para sistemas de IA.

Competidores

Área de Capacidad / Resultados	COASP (EC-Council)	SANS - SEC535	OffSec - LLM Red Teaming	Hack The Box - AI Red Teamer	TONEX - COAIS
Metodología de IA ofensiva y hacking de sistemas de IA	✓	◐	◐	✓	◐
Reconocimiento en IA y mapeo de la superficie de ataque	✓	✓	◐	◐	◐
Escaneo de vulnerabilidades en IA y técnicas de fuzzing	✓	◐	✗	✗	✗
Inyección de prompts y ataques a aplicaciones basadas en LLM	✓	✗	✓	✓	✓
Machine Learning adversarial y ataques a la privacidad de	✓	✗	✗	✓	◐
Ataques a datos y a la cadena de entrenamiento	✓	✗	◐	✓	◐
IA agentica y ataques modelo-a-modelo	✓	◐	◐	◐	◐
Ataques a la infraestructura de IA y a la cadena de suministro	✓	✗	✓	◐	◐
Pruebas de seguridad en IA, evaluación y fortalecimiento (hardening)	✓	✗	◐	◐	✗
Respuesta a incidentes en IA y análisis forense	✓	✗	✗	✗	✗

Nota: Esta es la perspectiva de EC-Council sobre nuestro panorama competitivo.