+33 (0)1 48 10 09 46
formation@ambient-it.net
140 Avenue Jean Lolive 93500 Pantin - France

Updated on 03/25/2026

Sign up

OSAI+™ Certification Training (AI-300)

ALL-IN-ONE: EXAM INCLUDED IN THE PRICE WITH THE AI-300 COURSE

OFFSEC CERTIFICATIONS - PURCHASE YOUR CERTIFICATIONS

5 days (35 hours)

Overview

OSAI+ - OffSec AI Red Teamer - is a leading certification dedicated to the offensive security of artificial intelligence systems. Positioned at an advanced level, it covers LLMs, RAG architectures, AI agents, multi-agent interactions, the Model Context Protocol, the AI supply chain, and modern deployment infrastructures.

Our OSAI+ Certification training will enable you to master red teaming techniques applied to AI as part of operational preparation for the OffSec certification: reconnaissance, mapping of attack surfaces, agent exploitation, compromise of RAG pipelines, abuse of MCP layers, supply chain attacks, and exploitation of cloud and containerized infrastructures.

You will learn to identify, analyze, exploit, and chain vulnerabilities specific to modern AI systems, while developing a rigorous offensive methodology based on an understanding of architectures, data flows, trust relationships, and stealth mechanisms.

Upon completion of the training, you will be able to conduct a realistic offensive assessment of a complex AI environment, exploit vulnerabilities in LLM, RAG, agent, and MCP components, model AI-specific threats, and produce a professional report that meets the expectations of an advanced certification track.

Like all our training courses, this one is based on a resolutely hands-on approach, centered around labs.

Course Content

- 120 days of self-paced lab access
- 10 expert coaching sessions: 10 half-day sessions (9:00 AM to 12:30 PM) for a total of 35 hours
- 1 certification attempt
- Access to the learner forum

Objectives

- Understand the fundamental concepts of AI red teaming and the offensive capabilities of artificial intelligence systems.
- Identify and exploit vulnerabilities in LLMs, RAG architectures, AI agents, and MCP environments.
- Analyze the attack surfaces of a modern AI infrastructure, from the application layer down to cloud and containerized components.
- Implement advanced compromise scenarios including prompt injection, memory poisoning, data poisoning, and toolchain abuse.
- Develop an offensive testing strategy and produce a technical report suitable for certification purposes.
- Effectively prepare for OSAI+ certification in an intensive, hands-on setting.

Target Audience

- Penetration testers
- Cybersecurity engineers
- AI/ML engineers
- Cloud architects

Prerequisites

- TCP/IP Networks – Intermediate to Advanced
- Linux/Windows System Administration – Intermediate level
- Bash/Python scripting – Intermediate level
- Knowledge of offensive cybersecurity (penetration testing, exploitation)
- Basic understanding of artificial intelligence (LLM, APIs, RAG) recommended
- Proficiency in technical English

Software Prerequisites

- **Kali Linux** – Downloadable here

Note: Ambient IT does not own OSAI+™; this certification belongs to OffSec® Services LLC.

OSAI+ Certification Training

[First Half-Day]

Understanding AI System Security

- Introduction to AI system security and the positioning of the OSAI+ certification
- Differences between traditional red teaming and AI red teaming
- Understanding the challenges of data theft, model manipulation, and workflow hijacking
- Overview of the MITRE ATLAS and OWASP Top 10 for LLMs frameworks
- Reading the offensive cycle: Reconnaissance, Access, Influence, Persistence, Exfiltration

[2nd half-day]

Mapping AI attack surfaces

- Identifying LLM attack surfaces: public APIs, fine-tuning, and on-premises deployments
- Understanding the components of a RAG architecture: retrievers, vector databases, and ingestion pipelines
- Examining orchestration layers: MCP, tools, and agents
- Identify critical dependencies between AI services and application components
- Detect vulnerabilities and potential attack paths in a modern AI environment

[3rd half-day]

Reconnaissance and information gathering

- Conducting passive reconnaissance: public repositories, exposed documentation, leaks, and AI-focused OSINT
- Conduct active reconnaissance: API enumeration, model fingerprinting, and endpoints
- Analyzing dependencies and trust cascades between ingestion, retrieval, and generation
- Understand the traces left by AI systems: structured prompts, embeddings, API calls
- Adopt stealth tactics: query variation, rate limiting, and honeypot evasion

[4th half-day] Attacking AI

agents

- Understand the architecture of an AI agent: memory, tool registry, planning loop, and system messages
- Exploiting direct and indirect prompt injection vulnerabilities
- Manipulate persistent memory to influence future decisions
- Exploiting tool invocation vulnerabilities and insufficient permission controls
- Implementing post-compromise stealth techniques in an agent-based environment

[5th half-day]

Attacking multi-agent systems and A2A communications

- Study multi-agent architectures and distributed coordination mechanisms
- Understand the fundamentals of A2A protocols: message structure, transport, and authentication models
- Performing impersonation, relay, and replay attacks on inter-agent traffic
- Exploit validation, schema, and serialization flaws
- Alter a workflow without breaking it while maintaining state continuity

[6th half-day] Exploiting

RAG pipelines

- Understand the RAG chain: ingestion, chunking, vectorization, storage, retrieval, and LLM synthesis
- Identify weaknesses in the knowledge base, retrieval, and embedding layers
- Implementing knowledge base leakage and retrieval hijacking attacks
- Executing ingestion poisoning and embedding attack scenarios
- Hiding malicious payloads within legitimate documents to increase stealth

[7th half-day]

Exploit the Model Context Protocol

- Understand the role of the Model Context Protocol in orchestrating and mediating tools
- Identify tool description poisoning and tool shadowing attacks
- Exploiting endpoint redirection and constraint bypass scenarios
- Chain tools to achieve privilege escalation or unintended actions
- Adopt stealth techniques by mimicking legitimate usage and expected outputs

[8th Half-Day] Attacking the

AI Supply Chain

- Map the AI supply chain: datasets, model weights, adapters, scripts, and registries
- Understand dataset poisoning and trigger injection mechanisms
- Examining model compromises: backdoored weights, LoRA poisoning, and adversarial layers
- Exploit distribution, update, and trust mechanisms
- Disguise malicious modifications to appear compliant with standard controls

[9th half-day]

Exploiting AI infrastructures and deployments

- Understanding AI deployment architectures: AWS SageMaker, Azure OpenAI, Vertex AI, self-hosted servers, and Kubernetes clusters
- Identify cloud configuration errors: overly permissive roles, public endpoints, and secrets Presentations
- Exploiting containers, orchestration, and Kubernetes RBAC abuses
- Attacking model servers via hot-reload, control APIs, or custom handlers
- Maintaining discretion in MLOps and CI/CD environments

[10th half-day]

Threat modeling and exam preparation

- Identify trust boundaries, dependencies, and escalation paths in an AI system
- Mapping critical assets: proprietary weights, sensitive embeddings, MCP endpoints, and RAG sources
- Build a comprehensive and realistic AI Red Team strategy
- Structure a professional report that meets OffSec's expectations
- Effectively prepare for the practical exam and reporting phase of the OSAI+ certification

FAQ – QUESTIONS & ANSWERS

HOW IS THE OSAI+™ CERTIFICATION EXAM CONDUCTED?

You must read the official guide before taking your exam.

The practical phase of the exam lasts 24 hours and involves attacking a complete AI environment by exploiting as many vulnerabilities as possible. After this phase, you will have 24 hours to complete and submit the exploitation report, in which you will explain your approach.

IN WHAT LANGUAGE IS THE OSAI+™ TRAINING TAUGHT?

Exam preparation will be in French. However, the supplementary materials provided by OffSec are in English.

IS THE OSAI+™ CERTIFICATION EXAM INCLUDED IN THE TRAINING PRICE?

Yes, you will be able to take the exam after completing the training.

HOW LONG IS THE LAB ACCESSIBLE?

You have access to the lab for 120 days.

IN WHAT LANGUAGE IS THE EXAM CONDUCTED?

The exam is conducted in English.

DO I NEED A WEBCAM?

Yes, your webcam must be active throughout the entire exam, and it must be able to capture your entire room.

DO I NEED A GOOD INTERNET CONNECTION?

Yes, because your computer must be able to maintain a TeamViewer connection for 24 hours while continuously running simulations.

HOW MUCH DOES THE CERTIFICATION COST?

The certification alone costs 1,749 euros.

Target Audience

This training is intended for both individuals and companies, large or small, wishing to train their teams in new advanced IT technology or to acquire specific professional knowledge or modern methods.

Entry-level assessment

The pre-training assessment complies with Qualiopi quality standards. Upon final registration, the learner receives a self-assessment questionnaire that allows us to evaluate their estimated proficiency in various types of technologies, as well as their expectations and personal goals for the upcoming training, within the limits imposed by the selected format. This questionnaire also allows us to anticipate certain connection or internal security issues within the company (intra-company or virtual classroom) that could pose challenges for monitoring and ensuring the smooth running of the training session.

Teaching Methods

Practical Course: 60% Practical, 40% Theory. Training materials distributed in digital format to all participants.

Organization

The course alternates between theoretical input from the trainer, supported by examples and reflection sessions, and group work.

Certification

At the end of the session, a multiple-choice quiz is used to verify that the skills have been properly acquired.

Certification

A certificate will be issued to each trainee who has completed the entire training program.