# Certified AI Security Professional Training

Duration: **4 Day**

BILGINC IT Academy
info@bilginc.com
bilginc.com

What Can You Achieve with This Training?

- Learn how to secure **artificial intelligence systems throughout their development, deployment, and operational lifecycle**.
- Understand the architectures and security requirements of **machine learning, deep learning, large language models (LLMs), and AI agents**.
- Recognise AI-specific attack techniques such as **Prompt Injection, Jailbreaking, Data Poisoning, and Model Extraction**.
- Explore how **MITRE ATLAS** and the **OWASP LLM Top 10** support AI threat analysis and attack-path assessment.
- Develop knowledge of **AI red teaming, security testing, monitoring, and incident response**.
- Understand security approaches for **LLM applications, RAG architectures, AI agent permissions, and AI supply chains**.
- Prepare for the **PECB Certified Artificial Intelligence Security Professional** certification examination.

The **Certified AI Security Professional (CAISP) Training** is a **four-day programme** designed for cybersecurity and artificial intelligence professionals who need to understand and manage security risks throughout the AI system lifecycle.

AI solutions introduce security challenges that extend beyond traditional software environments. **Model training, data pipelines, inference services, large language models, external tool integrations, and autonomous agents** create additional assets, dependencies, and attack surfaces that require protection.

This course examines how **machine learning, deep learning, large language models (LLMs), and AI agents** are designed, developed, deployed, attacked, and defended.

Participants explore **AI threat actors, adversarial objectives, trust boundaries, and security controls**. Frameworks such as **MITRE ATLAS** and the **OWASP LLM Top 10** are introduced to support structured analysis of AI threats and attack techniques.

Technical modules address attacks including **Prompt Injection, Jailbreaking, data poisoning, model extraction, sensitive information disclosure, insecure output handling, and tool abuse**.

The defensive modules focus on **secure AI architecture, defence-in-depth, access controls, guardrails, secure RAG, agent permissions, and human oversight**.

The programme also covers **AI supply chain security, vulnerability assessment, AI penetration testing, red teaming, security monitoring, SIEM integration, and incident response**.

Additional topics examine the use of **AI as a defensive cybersecurity capability**, alongside AI security governance, risk assessment, and the development of organisational AI security programmes.

# Certified AI Security Professional Training

Duration: **4 Day**

BILGINC IT Academy
info@bilginc.com
bilginc.com

**Practical exercises, labs, knowledge checks, and quizzes** reinforce the technical content and support preparation for the **PECB Certified Artificial Intelligence Security Professional** examination.

### Exams and Assessments

Following the course, participants take the formal **PECB Certified Artificial Intelligence Security Professional certification examination**.

## Prerequisites

Participants should have:

- A **solid understanding of core cybersecurity and information security concepts**
- Knowledge or experience in **threat modelling, identity and access control, incident response, security monitoring, and general security architecture** is recommended
- A **basic understanding of AI and machine learning**, including model training, data pipelines, and API-based AI services

## Who Should Attend

This course is intended for professionals responsible for securing AI systems or developing specialist knowledge in **artificial intelligence security**.

- **Cybersecurity and information security professionals**
- **AI and machine learning engineers**
- **AI security engineers**
- **Security architects and consultants**
- **SOC analysts and incident responders**
- **Threat intelligence professionals**
- **Data scientists and AI developers**
- **AI risk and compliance professionals**
- **Technology managers and technical leads** responsible for AI security
- Professionals securing **AI-enabled applications and services**

# Certified AI Security Professional Training

Duration: **4 Day**

BILGINC IT Academy
info@bilginc.com
bilginc.com

## What You Will Learn

By the end of this course, participants will be able to:

- Explain the architectures and operation of **machine learning, deep learning, LLMs, and AI agents**.
- Identify security risks throughout the **AI development, training, deployment, and operational lifecycle**.
- Analyse **AI attack surfaces, threat actors, and trust boundaries**.
- Assess AI threats using **MITRE ATLAS** and the **OWASP LLM Top 10**.
- Recognise attack techniques such as **Prompt Injection, Jailbreaking, Data Poisoning, Model Extraction, and Model Denial of Service**.
- Explain the purpose and application of **AI red teaming and security testing**.
- Identify appropriate security controls for **AI models, data, applications, and infrastructure**.
- Evaluate defensive approaches including **secure RAG, guardrails, context isolation, and vector database protection**.
- Understand how to manage **AI agent permissions and tool access** according to least-privilege principles.
- Recognise supply chain risks involving **third-party models, open-source components, and training datasets**.
- Explain approaches to **AI security monitoring, telemetry, logging, SIEM integration, and detection engineering**.
- Understand how to **detect, investigate, contain, and recover from AI-related security incidents**.
- Evaluate the opportunities and limitations of AI in **threat intelligence, threat hunting, malware analysis, and reverse engineering**.
- Describe the role of **AI security governance, risk assessment, and security policies** in organisational programmes.
- Identify the main components of an **AI security programme**, including continuous assurance and supply chain governance.

## Training Outline

### 1. AI Security Foundations

This module establishes the technical foundations needed to understand AI systems and their security

BILGINC IT Academy

info@bilginc.com

bilginc.com


requirements.

- **Artificial intelligence concepts and terminology**
- **Machine learning fundamentals**
- **Supervised and unsupervised learning**
- **Deep learning and neural networks**
- **AI model development and training**
- **Training, validation, and testing data**
- **Data pipelines and the model lifecycle**
- **AI inference and deployment**
- Understanding the **AI technology stack**

## 2. Security for AI

This module examines how AI security differs from traditional application security and identifies the assets and trust boundaries that require protection.

- Differences between **AI security and traditional application security**
- **AI attack surfaces**
- **Threat actors and adversarial objectives**
- **AI assets requiring protection**
- **Confidentiality, integrity, and availability** in AI environments
- **Model and data integrity**
- **AI system trust boundaries**
- **Security across the AI lifecycle**
- **Threat modelling AI systems**

## 3. AI Architecture and Security

This module explores the security implications of AI infrastructure, services, and external dependencies.

- **AI infrastructure and platforms**
- **Model hosting and inference services**
- **APIs and application integration**
- **Cloud-based AI services**

BILGINC IT Academy

info@bilginc.com

bilginc.com


- **Third-party AI dependencies**
- **AI supply chain considerations**
- Security implications of **model provenance**

### 4. Large Language Model Architecture

This module introduces the architecture and operational components of large language models and LLM-based applications.

- **LLM architecture and operation**
- **Transformers and tokenisation**
- **Model training and fine-tuning**
- **Foundation models**
- **Retrieval-Augmented Generation (RAG)**
- **Embeddings and vector databases**
- **LLM application architectures**
- **System prompts and context**
- **AI agents and autonomous systems**

### 5. LLM and Generative AI Attacks

This module examines common attack techniques targeting large language models and generative AI applications.

- **Prompt Injection**
- **Direct and indirect Prompt Injection**
- **Jailbreaking**
- **Sensitive information disclosure**
- **Model manipulation**
- **Model extraction**
- **Training data attacks**
- **Data poisoning**
- **Adversarial inputs**
- **Model denial of service**

BILGINC IT Academy

info@bilginc.com

bilginc.com


- **Insecure output handling**
- **Excessive agency**
- **Tool and plugin abuse**

### 6. AI Threat Frameworks

This module introduces frameworks used to analyse AI threats and map adversarial techniques.

- **AI attack lifecycle**
- **MITRE ATLAS**
- **OWASP LLM Top 10**
- **Mapping AI attack techniques**
- **Threat scenarios and attack paths**
- Identifying weaknesses across **AI architectures**

### 7. AI Red Teaming

This module introduces security testing from an adversarial perspective.

- **Purpose of AI red teaming**
- **AI security testing methodologies**
- **Prompt-based security testing**
- Testing **model boundaries and safeguards**
- Identifying **exploitable behaviours**
- **Documenting and communicating findings**

### 8. Securing AI Systems

This module focuses on protecting AI models, data, applications, and supporting infrastructure through security controls.

- **Secure AI architecture**
- **Defence-in-depth for AI**
- **Input validation**
- **Output filtering**
- **Model access controls**
- **Authentication and authorisation**

BILGINC IT Academy

info@bilginc.com

bilginc.com


- **Privileged AI operations**
- **Secrets and credential protection**
- **Data protection**
- **Model protection**

### 9. Securing LLM Applications

This module explores defensive controls for LLM applications, RAG solutions, and AI agents.

- **Prompt Injection mitigations**
- **Guardrails**
- **System prompt protection**
- **Context isolation**
- **Secure RAG architectures**
- **Protecting vector databases**
- **Controlling tool access**
- **Agent permissions**
- **Least privilege for AI agents**
- **Human oversight and approval controls**

### 10. AI Supply Chain Security

This module examines security risks associated with external models, data sources, and software components used in AI solutions.

- **Third-party models**
- **Open-source models**
- **Model repositories**
- **AI libraries and dependencies**
- **Training datasets**
- **Model provenance**
- **Supplier assurance**
- **Supply chain compromise**
- **Dependency and component management**

BILGINC IT Academy

info@bilginc.com

bilginc.com


## 11. AI Security Testing

This module introduces approaches for identifying weaknesses and evaluating the effectiveness of AI security controls.

- **Vulnerability assessment**
- **Adversarial testing**
- **Model security testing**
- **AI penetration testing**
- **AI red teaming**
- **Testing guardrails and controls**
- **Security assurance before deployment**

## 12. AI Security Operations

This module examines how AI-specific security events can be monitored and investigated.

- **AI-specific telemetry**
- **Logging AI activity**
- **Monitoring model interactions**
- **Detecting abnormal AI behaviour**
- **Security monitoring integration**
- **SIEM integration**
- **Detection engineering for AI systems**
- **Identifying adversarial activity**

## 13. AI Incident Response

This module covers the fundamental activities involved in responding to security incidents affecting AI systems.

- **AI security incident scenarios**
- **Detecting AI compromise**
- **Incident triage**
- **Investigation**
- **Containment**

BILGINC IT Academy

info@bilginc.com

bilginc.com


- **Model isolation**
- **Recovery**
- **Evidence collection**
- **Post-incident analysis**

### 14. AI-Enhanced Cybersecurity

This module explores the use of AI as a defensive capability within cybersecurity operations.

- **AI-assisted security monitoring**
- **AI-assisted threat hunting**
- **Threat intelligence**
- **Security analysis**
- **Malware analysis**
- **AI-assisted reverse engineering**
- **Opportunities and limitations of defensive AI**

### 15. AI Security Management

This module examines governance, policies, risk management, and organisational responsibilities for AI security.

- **AI security governance**
- **Roles and responsibilities**
- **AI security policies**
- **AI risk assessment**
- **Security assurance**
- **Regulatory and compliance considerations**
- **Managing third-party AI risk**
- **Governance of AI models and agents**

### 16. Building an AI Security Programme

This module focuses on integrating AI security into established organisational cybersecurity practices.

- **AI asset identification**
- **AI security strategy**

# Certified AI Security Professional Training

Duration: **4 Day**

BILGINC IT Academy
info@bilginc.com
bilginc.com

- **Secure architecture principles**
- **AI security controls**
- **Security testing and red teaming**
- **Monitoring and incident response**
- **Supply chain governance**
- **Continuous assurance**
- **Integrating AI security into existing cybersecurity programmes**