Escuela de Verano de Ciberseguridad

25 al 29 de agosto del 2025 en el Palacio de Minería de la Ciudad de México

El presente documento describe el objetivo, estructura, cronograma y perfil de ingreso del curso *Sistemas de Inteligencia Artificial Seguros* a ser impartido por el Laboratorio de Ciberseguridad del Instituto Politécnico Nacional en la Escuela de Verano de Ciberseguridad.

El responsable académico del curso es el Dr. Eleazar Aguirre, Profesor Titular e Investigador Nacional del Centro de Investigación en Computación del IPN con apoyo del M. en C. Hugo Sebastian Pacheco Rodríguez estudiante de doctorado y especialista en seguridad para inteligencia artificial.

Título del curso

Sistemas de inteligencia artificial seguros

Aspectos generales del curso

Modalidad	Presencial
Duración total	8 horas
Cupo	60
Lengua de impartición	Español

Resumen general

El curso ofrece un panorama actual de los marcos de trabajo y mejores prácticas de seguridad para sistemas de inteligencia artificial. Se revisarán los modelos de amenaza desde una perspectiva del ciclo de vida de desarrollo de los sistemas de inteligencia artificial. Se identificarán contramedidas y mejores prácticas de mitigación a técnicas, tácticas y procedimientos conocidos.

Objetivo del curso

Que el estudiante conozca los marcos de trabajo de gestión de riesgos en inteligencia artificial y las mejores prácticas que protegen de técnicas, tácticas y procedimientos empleados por los modelos de amenaza a sistemas de inteligencia artificial.

Temario del curso

1. Sistemas de Inteligencia Artificial

1.1. Antecedentes y marco conceptual
1.2. Sistemas de IA
1.3. Arquitectura y componentes de los sistemas de IA
1.4. Ciclo de vida del desarrollo de los sistemas de IA
1.5. Desarrollo y Operaciones de seguridad para sistemas de IA

2. Modelos de amenaza de sistemas de IA

2.1. MITRE Adversarial Threat Landscape for AI Systems
2.2. OWASP Machine Learning Security Top 10
2.3. CSA Large Language Model (LLM) Threats Taxonomy

3. Marcos de trabajo de seguridad de la inteligencia artificial

3.1. ENISA Multilayer framework for good cybersecurity practices for AI
3.2. NIST AI 100-1, Artificial Intelligence Risk Management Framework (AI RMF 1.0)

4. Mejores prácticas de seguridad para IA
    4.1. Mejores prácticas de seguridad para LLM y datos de inteligencia artificial generativa
    4.2. UK NCSC, Guidelines for secure AI system development
    4.3. Contramedidas de seguridad para agentes inteligentes artificiales

Perfil de ingreso

Estudiantes de los últimos semestres de nivel licenciatura y de nivel posgrado de inteligencia artificial, ciencias de la computación, ciencias de datos o carreras afines. Es deseable que el estudiante haya cursado asignaturas de inteligencia artificial, sistemas de cómputo, ingeniería de software y ciberseguridad.

Metodología del curso

Los instructores presentarán los conceptos, marcos de trabajo, modelos y paradigmas de seguridad. Se revisarán tácticas y técnicas de los modelos de amenaza. Se revisarán conocimientos teóricos y ejemplos prácticos en casos de uso o contextos específicos. Durante las exposiciones se incentivará la reflexión, discusión y análisis de los estudiantes. El estudiante aplicará procedimientos y mejores prácticas de seguridad a un caso de uso. El reporte de estas actividades se entregará 10 días posteriores al término del curso.

Cronograma del curso

Fecha	**Temas por desarrollar**
25 de agosto	1.1 – 1.5 y 2.1
26 de agosto	2.2 y 2.3
27 de agosto	3.1 y 3.2
28 de agosto	4.1 – 4.3
5 de septiembre	Entrega del reporte final

Bibliografía del curso

1. ETSI TS 104 223 V1.1.1, (2025-04), Securing Artificial Intelligence (SAI); Baseline Cyber Security Requirements for AI Models and Systems.
2. Securing LLM Backed Systems: Essential Authorization Practices, AI Technology and Risk Working Group, 2024 Cloud Security Alliance.
3. OWASP Top 10 for LLMs - LLM and Gen AI Data Security Best Practices Guide 1.0, Febrary 2025.
4. Agentic AI – Threats and Mitigations, OWASP Top 10 for LLM Apps & Gen AI, Agentic Security Initiative.
5. Large Language Model (LLM) Threats Taxonomy, AI Controls Framework Working Group, CSA.
6. Guidelines for secure AI system development, UK NCSC.