COURSE SYLLABUS

Applied AI Security Certificate Program (Self-Paced)

A three-course self-paced engineering path: securing AI/ML systems, monitoring and responding to AI incidents, then red teaming and assurance.

> Group classes in Live Online and onsite training is available for this course. For more information, email onsite@graduateschool.edu or visit: https://www.graduateschool.edu/certificates/applied-ai-security-certificate-program-self-paced

support@graduateschool.edu •
(888) 744-4723

Course Outline

This package includes these courses

- Securing AI/ML Systems: Protecting the Intelligent Attack Surface Course (Self-Paced) (12 Hours)
- AI Red Teaming and Assurance Course (Self-Paced) (12 Hours)
- AI Security Monitoring and Incident Response Course (Self-Paced) (8 Hours)

Securing AI/ML Systems: Protecting the Intelligent Attack Surface Course (Self-Paced)

Engineering-tier AI and machine learning security spanning data, model, pipeline and application, built around adversarial machine learning and generative AI risk.

- AI/ML system architecture, with models, datasets, pipelines, applications and infrastructure distinguished
- AI-specific security risk separated from traditional cybersecurity risk
- Security mapped across the AI lifecycle from design to retirement
- Protecting confidentiality, integrity, availability and provenance of training and operational data
- Data poisoning and manipulation risks in collection, labeling and preprocessing
- Poisoning, evasion, privacy and model-extraction attacks, and defenses that increase robustness
- Models, weights and checkpoints as protected assets, with provenance verified before deployment
- Third-party models, datasets, libraries and frameworks assessed as supply-chain dependencies
- Direct and indirect prompt injection, insecure output handling, and retrieval-augmented generation risk
- Security for ML pipelines, orchestration platforms, APIs and deployment
- Detecting abnormal inputs, outputs and model behavior versus performance degradation
- AI-specific incident response: containment, model rollback and post-incident validation

AI Red Teaming and Assurance Course (Self-Paced)

An advanced course and the capstone of the applied AI security path, covering adversarial testing of AI systems and the assurance evidence that justifies running them.

- Red teaming distinguished from adjacent assurance activities and placed in the assurance lifecycle
- The red-team engagement lifecycle from planning through reporting
- Scoping to risk classification and setting usable objectives
- Rules of engagement, team composition and deconfliction
- Matching adversarial testing techniques to attack surfaces
- Manual expert-driven testing weighed against automated approaches
- Probing techniques for generative AI systems
- Telling a red-team finding apart from an ordinary software bug
- TEVV records that support a risk decision
- The assurance case as claim, argument and evidence
- Model cards, system cards and independent evaluation
- Findings reports prioritized by impact rather than technical complexity
- Remediation, retesting and human oversight as a mitigation
- Recognizing when a finding becomes an incident
- Testing and assurance for foundation models and agentic, tool-using systems

AI Security Monitoring and Incident Response Course (Self-Paced)

Practitioner-level AI security monitoring and incident response, covering what to watch, how to tell an incident from a blip, and what to do in the first hour.

- How AI security risks differ from and compound traditional IT risks
- The vocabulary of AI incidents: data poisoning, adversarial examples, prompt injection, drift, extraction
- Continuous monitoring for a high-impact AI system, beyond uptime and throughput
- The four monitoring categories and the technique that fits each
- Risk-based security event logging for AI systems
- Reading monitoring signals and identifying likely causes
- Telling anomalies apart from genuine AI security incidents
- Preserving AI-specific evidence: model version, prompt and response logs, data lineage
- Classifying severity by impact on safety, rights, mission and data
- The four-phase incident response lifecycle applied to AI
- AI-specific containment, including model rollback and disabling automated decisioning
- Root cause analysis and revalidation before return to automated operation