COURSE SYLLABUS

Applied AI Security Certificate Program

A three-course engineering path through AI security: securing AI/ML systems, monitoring and incident response, then red teaming and assurance.

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



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

Course Outline

This package includes these courses

- Securing AI/ML Systems: Protecting the Intelligent Attack Surface Course (12 Hours)
- AI Red Teaming and Assurance Course (12 Hours)
- AI Security Monitoring and Incident Response Course (6 Hours)

Securing AI/ML Systems: Protecting the Intelligent Attack Surface Course

An intermediate engineering course on securing AI and machine learning systems across data, model, pipeline and application, centred on adversarial machine learning and generative AI security.

- AI/ML system architecture, and models, datasets, pipelines, applications and infrastructure distinguished
- Traditional cybersecurity risk separated from AI-specific risk
- Security across the full AI lifecycle, from design through retirement
- Confidentiality, integrity, availability and provenance of training and operational data
- Data poisoning and manipulation across collection, labeling and preprocessing
- Poisoning, evasion, privacy and model-extraction attacks, with defenses that raise robustness
- Models, weights and checkpoints as protected assets, with provenance and integrity verification
- Third-party models, datasets, libraries and frameworks as supply-chain dependencies
- Direct and indirect prompt injection, insecure output handling, and RAG risk
- Securing ML pipelines, orchestration platforms, APIs and deployment
- Detecting abnormal inputs, outputs and behavior, and telling degradation from compromise
- AI-specific incident response, including containment, model rollback and post-incident validation

AI Red Teaming and Assurance Course

Advanced adversarial testing and assurance for AI systems — the capstone that connects breaking a system to proving it is trustworthy enough to run.

- Red teaming separated from adjacent assurance activities and placed in the assurance lifecycle
- The engagement lifecycle from planning through reporting
- Scoping to a system's risk classification and setting usable objectives
- Rules of engagement, team composition and deconfliction
- Matching adversarial testing techniques to the attack surface
- Manual expert-driven testing against automated and tool-assisted approaches
- Generative AI probing techniques
- Separating a genuine red-team finding from an ordinary bug
- TEVV records that support a risk decision
- Assurance cases built as claim, argument and evidence
- Model cards, system cards, and independent evaluation results
- Findings reports prioritized by impact over technical complexity
- Remediation, retesting, and human oversight as a mitigation category
- When a finding becomes an incident
- Extending testing and assurance to foundation models and agentic systems

AI Security Monitoring and Incident Response Course

AI security monitoring and incident response for practitioners: what to watch for, distinguishing incidents from anomalies, and the first-hour response.

- AI security risks that differ from or compound traditional IT risk
- Working vocabulary: data poisoning, adversarial examples, prompt injection, drift, extraction
- What continuous monitoring means for a high-impact AI system
- The four AI monitoring categories and matching techniques
- Risk-based security event logging for AI systems
- Reading monitoring signals and identifying likely causes
- Distinguishing anomalies from AI security incidents
- Preserving model version, prompt and response logs, and training data lineage
- Severity classified 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
- Post-incident root cause analysis and revalidation