

APPLIED GEN AI SECURITY ENGINEERING

From Trust Boundary to Enterprise Assurance

Enterprise AI • LLM • RAG • Agentic AI • MCP • Red Teaming • Defense • Governance



4

MONTHS



16

WEEKS



80

HOURS



7

MODULES



8

HANDS-ON LABS

TABLE OF CONTENTS

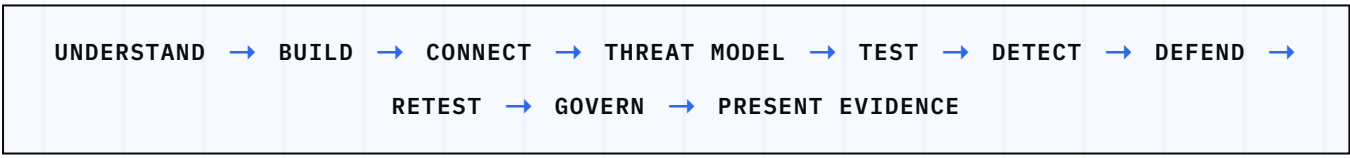
01	Program Overview	03
02	Who This Program Is For	04
03	Learning Progression	05
04	The Four-Month Journey	06
05	Seven-Module Architecture	07
06	Detailed 16-Week Curriculum	08
07	Hands-On Lab Journey	25
08	Enterprise Capstone Experience	29
09	Professional Capabilities Developed	30
10	Program Outcomes	31
11	Career Alignment	31
12	Prerequisites	32
13	Delivery Approach	33
14	Curriculum Governance Note	33

01 PROGRAM OVERVIEW

Applied Gen AI Security Engineering Program is a four-month, sixteen-week enterprise AI security journey built around seven progressive capability modules and a hands-on lab-based learning path. It carries learners from an accurate understanding of how AI and LLM systems actually behave through to threat-modeling, authorised testing, defending, and governing production-grade enterprise AI systems.

Rather than a tour of trending attack techniques, the program follows one deliberately sequenced difficulty curve: every week is strictly built on the one before it, so beginners are never dropped into offensive or governance material before its prerequisite has been established, while experienced security professionals still reach genuine enterprise assessment and governance work within the first two months. Eight instructor-led hands-on labs (one per module, plus a culminating enterprise capstone lab) give every learner a demonstrable portfolio of evidence-backed security work by the time they finish.

HOW THE PROGRAM PROGRESSES



THREE PILLARS OF THE PROGRAM

 SECURITY OF AI Assessing and improving the security of LLM, RAG, agentic and enterprise AI systems.	 AI FOR CYBERSECURITY Using AI as a controlled, human-validated accelerator for security operations.	 AI GOVERNANCE, RISK & ASSURANCE Converting technical evidence into defensible enterprise risk decisions.
--	--	---

02 WHO THIS PROGRAM IS FOR

- Cybersecurity, application-security and cloud-security professionals formalising AI-specific security skill
- SOC and security-operations professionals extending detection and response into AI-enabled workflows
- DevSecOps practitioners and developers moving toward secure AI engineering
- AI/ML practitioners who need genuine security depth beyond model performance
- Governance, risk and assurance professionals bringing AI systems into existing risk frameworks
- Experienced security professionals and architects transitioning into enterprise AI security leadership

03 LEARNING PROGRESSION

Every learner moves through the same three stages, in the same order. The only thing that varies is how much of it feels new.



FOUNDATION

WEEKS 1-4

No prior AI security experience is assumed at the start of the program. Learners first build an accurate understanding of how AI and LLM systems behave, then establish the cybersecurity and enterprise-architecture foundation needed to secure them: trust boundaries, RAG and retrieval security, conversation memory, data handling and API security. This foundation supports every assessment and governance stage that follows.



APPLIED ASSESSMENT & DEFENSE

WEEKS 5-12

The core assessment build-out: agentic AI and MCP security, structured threat modeling, authorised adversarial testing, and the detection and defensive engineering needed to close what testing reveals.



ENTERPRISE GOVERNANCE & CAPSTONE

WEEKS 13-16

Enterprise integration: using AI responsibly to support security operations, translating technical evidence into governance and assurance decisions, and a two-week capstone that ties every capability into one defensible enterprise AI security assessment.

04 THE FOUR-MONTH JOURNEY

MONTH 1

WEEKS 1-4

Foundations & Enterprise AI Application Security

The program opens with no prerequisites beyond basic technical literacy assumed. Learners build an accurate mental model of AI and LLM behaviour, connect it to core cybersecurity principles, and establish a fully evidenced security baseline for a representative enterprise AI application: the shared starting line for every learner in the cohort, regardless of background.

MONTH 2

WEEKS 5-8

Agentic AI Security & Threat-Driven Test Planning

With the foundations in place, learners move into autonomous systems, assessing agent identity, delegated authority and MCP trust boundaries, then convert architectural understanding into a structured threat model and an authorised, evidence-driven security-test plan.

MONTH 3

WEEKS 9-12

Offensive Assessment, Detection & Defense

Month Three shifts from planning to execution. Learners conduct authorised adversarial testing across prompt, retrieval, agent, tool, MCP and API surfaces in a controlled cyber range, then design detection, implement layered controls, and validate risk reduction through retesting.

MONTH 4

WEEKS 13-16

AI-Enabled Security Operations, Governance & Capstone Mastery

The final month connects technical practice to enterprise decision-making: using AI responsibly to support security operations, translating findings into governance, risk and assurance language, and then spending two full weeks integrating every prior month's work into one coherent, defensible enterprise AI security assessment that learners design, review, and present.

05 SEVEN-MODULE ARCHITECTURE

The seven modules below are the conceptual backbone of the program. Some are compact; others are staged across several weeks because the underlying capability is too large to teach responsibly in a single session.

1	AI, LLM & Cybersecurity Foundations Building accurate mental models of AI and LLM behaviour and core cybersecurity principles as the shared foundation for enterprise AI security.	WEEKS 1-2
2	Enterprise AI Application Security Understanding how RAG, memory, data handling and APIs reshape the enterprise AI attack surface, and establishing a validated security baseline.	WEEKS 3-4
3	Agentic AI & MCP Security Securing autonomous, tool-using agents: identity, delegated authority, least privilege and MCP trust boundaries.	WEEKS 5-6
4	AI Threat Modeling & Security Test Planning Converting architecture understanding into structured threat models and authorised, evidence-driven test plans.	WEEKS 7-8
5	Offensive AI Security & Red Teaming Controlled, authorised adversarial assessment of prompt, retrieval, agent, tool and API surfaces with defensible evidence.	WEEKS 9-10
6	AI Security Detection & Defense Designing detection and monitoring, implementing layered controls, and validating risk reduction through retesting.	WEEKS 11-12
7	AI for Cybersecurity & Governance, Risk & Assurance Using AI responsibly to support security operations, and translating technical evidence into enterprise governance and assurance decisions.	WEEKS 13-14
C	Capstone Integration Experience Integrating the complete lifecycle into one defensible enterprise AI security assessment and executive presentation.	WEEKS 15-16

06 DETAILED 16-WEEK CURRICULUM

Every week below produces one concrete, demonstrable outcome. Applied-experience descriptions are intentionally kept high-level; the exact assessment methodology and step-by-step exercise design are delivered live in-session.

THE SIXTEEN WEEKS AT A GLANCE

WEEK 01	AI, LLM & Generative AI Foundations for Security	MODULE 1
WEEK 02	Cybersecurity Foundations & Enterprise AI Architecture	MODULE 1
WEEK 03	RAG, Retrieval & Knowledge-Base Security	MODULE 2
WEEK 04	Conversation Memory, AI Data & API Security	MODULE 2
WEEK 05	Agentic AI Architecture & Autonomous Capability Security	MODULE 3
WEEK 06	Agent Identity, Delegated Authority & MCP Security	MODULE 3
WEEK 07	AI Threat Modeling & Trust-Boundary Analysis	MODULE 4
WEEK 08	Security Test Planning, Hypotheses & Authorised Scope	MODULE 4
WEEK 09	AI Attack-Surface Assessment & Prompt/Retrieval Risk	MODULE 5
WEEK 10	Agent, Tool, MCP & API Security Testing	MODULE 5
WEEK 11	AI Security Detection & Monitoring Engineering	MODULE 6
WEEK 12	Defensive Controls, Runtime Protection & Retesting	MODULE 6
WEEK 13	AI-Assisted Cybersecurity Operations	MODULE 7
WEEK 14	AI Governance, Risk & Assurance	MODULE 7
WEEK 15	Capstone Architecture, Threat Model & Assessment Review	CAPSTONE INTEGRATION EXPERIENCE
WEEK 16	Capstone Evidence, Remediation & Executive Presentation	CAPSTONE INTEGRATION EXPERIENCE

AI, LLM & GENERATIVE AI FOUNDATIONS FOR SECURITY



Establish an accurate technical understanding of how AI, machine learning and large language models actually behave: the shared vocabulary the rest of the program builds on for security work.

KEY LEARNING AREAS

- 01** Artificial intelligence, machine learning and generative AI concepts relevant to security practice
- 02** Foundation models and the current landscape of open versus closed, frontier versus efficient systems
- 03** LLM behaviour: tokens, context windows, inference and embeddings
- 04** How prompts and instructions shape model output
- 05** Non-determinism, hallucination and their security implications
- 06** Where generative AI genuinely changes an organisation's risk profile



APPLIED EXPERIENCE

A guided, controlled-environment exploration observing baseline model behaviour and comparing how system and user instructions are processed.



PROFESSIONAL OUTCOME

Learners can explain how an LLM actually behaves, in both business and technical language, as the foundation for evaluating its security.

CYBERSECURITY FOUNDATIONS & ENTERPRISE AI ARCHITECTURE



Connect core cybersecurity principles to the architecture of an enterprise AI system, establishing the trust-boundary vocabulary used for the rest of the program.

KEY LEARNING AREAS

- 01 Confidentiality, integrity and availability applied to AI-enabled systems
- 02 Identity, authentication and authorisation fundamentals for AI workflows
- 03 Enterprise AI architecture: components, data flows and integration points
- 04 Trust boundaries and initial attack-surface awareness
- 05 Mapping familiar web and API security concepts onto AI-enabled applications
- 06 Security responsibilities across the AI system lifecycle



APPLIED EXPERIENCE

A guided architecture walkthrough identifying assets, trust boundaries and security responsibilities within a representative enterprise AI system.



PROFESSIONAL OUTCOME

Learners can identify the components, trust boundaries and security ownership of an enterprise AI system before any application-layer detail is introduced.

RAG, RETRIEVAL & KNOWLEDGE-BASE SECURITY



Understand how retrieval-augmented generation changes what an AI application can access, and where that introduces new security surface.

KEY LEARNING AREAS

- 01 RAG architecture: retrieval, grounding and generation
- 02 Vector stores, embeddings and knowledge-base design
- 03 Retrieval boundaries and unauthorised-access risk
- 04 Document-ingestion pipelines as a security entry point
- 05 Distinguishing a model risk from an application or data risk
- 06 Establishing a clean, evidenced application baseline before testing



APPLIED EXPERIENCE

A guided build of an enterprise-representative RAG application using controlled documents, followed by baseline validation of expected retrieval behaviour.



PROFESSIONAL OUTCOME

Learners can explain how RAG reshapes an AI application's attack surface and can validate and evidence a clean security baseline.

CONVERSATION MEMORY, AI DATA & API SECURITY



Extend the enterprise AI security baseline to conversational memory, sensitive-data handling and the APIs that connect AI systems to the wider enterprise.

KEY LEARNING AREAS

- 01 Conversation state and persistent memory design
- 02 AI data classification and handling of sensitive information
- 03 API security, authentication and authorisation for AI-enabled services
- 04 Approval workflows and audit-event design
- 05 Data and authorisation boundaries across AI application components
- 06 Evidence capture for later security comparison



APPLIED EXPERIENCE

A guided exercise extending the baseline application with memory, data-handling and API interactions, then documenting the resulting security posture.



PROFESSIONAL OUTCOME

Learners can recognise data and authorisation boundaries across a full enterprise AI application and preserve a defensible baseline for later assessment.

AGENTIC AI ARCHITECTURE & AUTONOMOUS CAPABILITY SECURITY



Understand what changes, from a security perspective, when an AI system can select capabilities and take action rather than only generate text.

KEY LEARNING AREAS

- 01 Agentic AI architecture and autonomous capability selection
- 02 Tool-enabled workflows and action-taking AI systems
- 03 Read versus write capability and the risk difference between them
- 04 Autonomy levels and appropriate human-in-the-loop control points
- 05 Early identification of excessive-agency risk
- 06 Governance considerations for agent actions



APPLIED EXPERIENCE

A guided assessment of a controlled agentic AI environment, evaluating how capability selection and action execution actually behave.



PROFESSIONAL OUTCOME

Learners can explain agentic AI architecture and identify where autonomous action introduces new security responsibility.

AGENT IDENTITY, DELEGATED AUTHORITY & MCP SECURITY



Examine how identity, delegated authority and the Model Context Protocol define trust between users, agents and connected tools.

KEY LEARNING AREAS

- 01 User identity versus agent identity, and delegated authority
- 02 Least-privilege design for autonomous agents
- 03 MCP clients, servers, tools and resources
- 04 Trust and secure context exchange between agents and tools
- 05 Approvals, monitoring and governance of agent actions
- 06 Recommending controls for connected-tool ecosystems



APPLIED EXPERIENCE

A guided evaluation of a controlled Agentic AI and MCP environment, assessing trust, authority and access-boundary design.



PROFESSIONAL OUTCOME

Learners can assess an Agent → MCP → Tool architecture and recommend proportionate authority and monitoring controls.

AI THREAT MODELING & TRUST-BOUNDARY ANALYSIS



Convert architectural understanding of an AI system into a structured threat model that identifies credible, prioritised risk.

KEY LEARNING AREAS

- 01 Assets, data flows, entry points and trust boundaries in AI systems
- 02 AI-specific threat modeling using structured security methods
- 03 RAG, memory, agent, MCP, identity and API risk within a single model
- 04 Third-party and supply-chain risk considerations
- 05 Prioritising risk by architecture and business impact
- 06 Aligning threat models to recognised industry frameworks



APPLIED EXPERIENCE

A guided threat-modeling exercise producing a structured, evidence-ready threat model for a representative enterprise AI system.



PROFESSIONAL OUTCOME

Learners can produce an AI-focused threat model and prioritise findings by business impact rather than technical novelty.

SECURITY TEST PLANNING, HYPOTHESES & AUTHORISED SCOPE



Translate a threat model into a defensible, authorised security-test plan before any testing activity begins.

KEY LEARNING AREAS

- 01 Translating credible threats into testable hypotheses
- 02 Defining authorised scope, rules of engagement and evidence requirements
- 03 Security objectives and success criteria for AI assessments
- 04 Ethical and legal considerations in AI security testing
- 05 Evidence-handling and chain-of-custody principles
- 06 Planning repeatable versus one-time evaluation activities



APPLIED EXPERIENCE

A guided planning exercise producing a scoped, authorised test plan derived directly from the prior week's threat model.



PROFESSIONAL OUTCOME

Learners can define defensible security-test cases and authorised scope before testing begins: the professional standard for AI security assessment.

AI ATTACK-SURFACE ASSESSMENT & PROMPT/RETRIEVAL RISK



Begin authorised, controlled adversarial assessment of an AI application's attack surface, starting with prompt and retrieval risk.

KEY LEARNING AREAS

- 01 AI attack-surface reconnaissance and assessment methodology
- 02 Prompt and indirect-instruction risk categories
- 03 RAG, retrieval and vector-store manipulation risk
- 04 Memory-manipulation risk categories
- 05 Maintaining scope, evidence and ethical controls throughout testing
- 06 Distinguishing demonstrable weakness from anecdotal claims



APPLIED EXPERIENCE

Authorised adversarial testing exercises in a controlled AI cyber range, focused on prompt and retrieval-layer risk categories.



PROFESSIONAL OUTCOME

Learners can demonstrate AI security weaknesses at the prompt and retrieval layer with evidence suitable for root-cause analysis.

AGENT, TOOL, MCP & API SECURITY TESTING



Extend authorised adversarial assessment to agent, tool, MCP and API surfaces, and introduce repeatable evaluation concepts.

KEY LEARNING AREAS

- 01 Agent, tool and MCP abuse scenarios
- 02 API abuse and authorisation-testing concepts for AI-enabled services
- 03 Repeatable and automated AI security evaluation concepts
- 04 Differentiating manual offensive testing from structured AI red teaming
- 05 Evidence capture supporting root-cause and business-impact analysis
- 06 Professional reporting standards for AI security findings



APPLIED EXPERIENCE

Authorised adversarial testing exercises in a controlled AI cyber range, focused on agent, tool, MCP and API risk categories.



PROFESSIONAL OUTCOME

Learners can differentiate manual testing from structured AI red teaming and capture evidence that supports defensible findings.

AI SECURITY DETECTION & MONITORING ENGINEERING



Use evidence from prior assessment activity to design meaningful detection and monitoring for AI-enabled workflows.

KEY LEARNING AREAS

- 01 Telemetry sources across prompt, retrieval, memory, agent, MCP and tool layers
- 02 Designing detections that reflect real attack behaviour rather than noise
- 03 Alert quality, triage considerations and false-positive management
- 04 Monitoring requirements across the AI application stack
- 05 Connecting detection design back to threat-model findings
- 06 Detection-engineering documentation standards



APPLIED EXPERIENCE

A guided exercise designing detection and monitoring requirements using evidence captured during the prior month's assessment activity.



PROFESSIONAL OUTCOME

Learners can design useful AI security telemetry and detections grounded in demonstrated attack behaviour.

DEFENSIVE CONTROLS, RUNTIME PROTECTION & RETESTING



Implement layered defensive controls at the correct trust boundary, then validate their effectiveness through retesting.

KEY LEARNING AREAS

- 01 Secure RAG, memory protection, agent identity and tool-authorisation controls
- 02 Runtime policy enforcement, execution limits and approval controls
- 03 Secure AI delivery and supply-chain awareness
- 04 Security regression concepts for AI-enabled systems
- 05 Original retest, modified bypass attempt and legitimate-workflow validation
- 06 Residual-risk analysis after control implementation



APPLIED EXPERIENCE

A guided exercise implementing layered controls in response to prior findings, then retesting to confirm risk reduction and preserved business functionality.



PROFESSIONAL OUTCOME

Learners can implement layered AI security controls and prove their effectiveness through retesting and residual-risk analysis.

AI-ASSISTED CYBERSECURITY OPERATIONS



Use AI responsibly as an accelerator for established cybersecurity workflows, with human validation built in throughout.

KEY LEARNING AREAS

- 01 AI-assisted log analysis and alert triage
- 02 AI-assisted threat hunting and incident analysis
- 03 Human-validation requirements and hallucination risk in security use cases
- 04 Appropriate versus inappropriate use of AI in security operations
- 05 Documenting AI-assisted analysis for audit and review
- 06 Limits of AI reliability in high-stakes security decisions



APPLIED EXPERIENCE

A guided exercise using AI to support a controlled cybersecurity analysis workflow, with structured human validation of every output.



PROFESSIONAL OUTCOME

Learners can use AI responsibly to support cybersecurity analysis while maintaining accountable human judgement.

AI GOVERNANCE, RISK & ASSURANCE



Translate technical AI security evidence into the governance, risk and assurance language enterprise decision-makers require.

KEY LEARNING AREAS

- 01 AI system inventory and risk-register practices
- 02 Provider risk, privacy considerations and control mapping
- 03 Alignment with recognised AI risk and assurance frameworks
- 04 Production-readiness and assurance-evidence standards
- 05 Residual-risk communication for non-technical stakeholders
- 06 Building an evidence-backed AI risk and assurance package



APPLIED EXPERIENCE

A guided exercise converting a technical AI security assessment into a structured risk, assurance and production-readiness package.



PROFESSIONAL OUTCOME

Learners can translate technical AI security findings into governance and business-risk decisions defensible to an executive audience.

CAPSTONE ARCHITECTURE, THREAT MODEL & ASSESSMENT REVIEW



Review the end-to-end architecture, threat model and assessment plan for the capstone scenario before finalising testing and remediation.

KEY LEARNING AREAS

- 01 A structured review checklist spanning the full program lifecycle
- 02 Common integration gaps across architecture, threat model and test scope
- 03 Running a staged review process with checkpoints
- 04 Presenting technical security reasoning to a structured review panel
- 05 Refining the assessment and remediation plan based on feedback



APPLIED EXPERIENCE

A staged review process: present the capstone architecture, threat model and test plan for structured feedback, then revise before proceeding to final testing and remediation.



PROFESSIONAL OUTCOME

Learners can identify and close assessment and remediation gaps before they become costly to fix.

CAPSTONE EVIDENCE, REMEDIATION & EXECUTIVE PRESENTATION



Complete the capstone assessment lifecycle and deliver a professional technical report and executive security presentation.

KEY LEARNING AREAS

- 01 Finalising evidence, retesting and residual-risk assessment
- 02 Technical report structure and evidence packaging
- 03 Presenting security decisions to an executive-level audience
- 04 Incorporating structured feedback under a deadline
- 05 Mapping completed capability to professional AI security roles



APPLIED EXPERIENCE

A final assessment, retest and reporting cycle, followed by a timed oral security presentation and executive Q&A.



PROFESSIONAL OUTCOME

Learners leave with a defensible, evidence-backed AI security assessment and the ability to present it to both technical and executive audiences.

07 HANDS-ON LAB JOURNEY

Each module culminates in one instructor-led, hands-on lab in a controlled AI environment: seven module labs across the program, building a portfolio of evidenced, practical security work one capability at a time.

LAB 1 AI, LLM & CYBERSECURITY FOUNDATIONS LAB

WEEKS 1-2

MODULE 1 · AI, LLM & CYBERSECURITY FOUNDATIONS

LAB OBJECTIVE · Work hands-on in a controlled AI environment to observe real model behaviour and map it onto core cybersecurity principles and enterprise AI architecture.

CAPABILITIES APPLIED

- Baseline LLM behaviour observation and instruction analysis
- Trust-boundary and asset identification in a representative AI system
- Mapping cybersecurity fundamentals onto AI-enabled architecture

PROFESSIONAL OUTCOME · A documented architecture and trust-boundary walkthrough of a representative enterprise AI system.

LAB 2 ENTERPRISE AI APPLICATION SECURITY BASELINE LAB

WEEKS 3-4

MODULE 2 · ENTERPRISE AI APPLICATION SECURITY

LAB OBJECTIVE · Build and validate a clean, evidenced security baseline for an enterprise AI application spanning RAG, memory, data handling and APIs.

CAPABILITIES APPLIED

- RAG and knowledge-base security configuration
- Conversation memory and AI data-handling review
- API authentication, authorisation and audit-event validation

PROFESSIONAL OUTCOME · A fully evidenced enterprise AI application baseline, ready for later security comparison.

LAB 3 AGENTIC AI & MCP SECURITY LAB

WEEKS 5-6

MODULE 3 • AGENTIC AI & MCP SECURITY

LAB OBJECTIVE • Assess a connected, tool-using AI agent environment to evaluate identity, delegated authority and MCP trust boundaries.

CAPABILITIES APPLIED

- Agent capability-selection and action-execution assessment
- Identity, delegated-authority and least-privilege evaluation
- MCP trust and connected-tool ecosystem review

PROFESSIONAL OUTCOME • A documented trust-and-authority assessment of an Agent → MCP → Tool architecture.

LAB 4 AI THREAT MODELING & TEST PLANNING LAB

WEEKS 7-8

MODULE 4 • AI THREAT MODELING & SECURITY TEST PLANNING

LAB OBJECTIVE • Produce a structured AI threat model and translate it into an authorised, evidence-driven security-test plan.

CAPABILITIES APPLIED

- Structured AI-specific threat modeling
- Risk prioritisation by architecture and business impact
- Authorised scope and test-hypothesis definition

PROFESSIONAL OUTCOME • A defensible threat model and authorised test plan, ready for controlled testing.

LAB 5 OFFENSIVE AI SECURITY & RED-TEAMING LAB**WEEKS 9-10****MODULE 5 · OFFENSIVE AI SECURITY & RED TEAMING**

LAB OBJECTIVE · Conduct authorised adversarial testing in a controlled AI cyber range across prompt, retrieval, agent, tool, MCP and API surfaces.

CAPABILITIES APPLIED

- Authorised adversarial testing methodology
- Evidence capture for root-cause and impact analysis
- Differentiating manual testing from structured AI red teaming

PROFESSIONAL OUTCOME · An evidence-backed record of demonstrated AI security weaknesses across the assessed surfaces.

LAB 6 AI DETECTION & DEFENSIVE ENGINEERING LAB**WEEKS 11-12****MODULE 6 · AI SECURITY DETECTION & DEFENSE**

LAB OBJECTIVE · Design detection and monitoring from prior evidence, implement layered controls, then retest to confirm risk reduction.

CAPABILITIES APPLIED

- AI security telemetry and detection design
- Layered defensive-control implementation
- Retesting and residual-risk validation

PROFESSIONAL OUTCOME · A validated set of detective and defensive controls with confirmed risk reduction through retesting.

LAB 7 AI-ENABLED SECURITY OPERATIONS & GOVERNANCE LAB**WEEKS 13-14****MODULE 7 · AI FOR CYBERSECURITY & GOVERNANCE, RISK & ASSURANCE**

LAB OBJECTIVE · Use AI responsibly to support a cybersecurity workflow, then convert the resulting technical evidence into a governance and assurance package.

CAPABILITIES APPLIED

- AI-assisted log analysis, triage and threat hunting with human validation
- AI risk-register and control-mapping construction
- Assurance-evidence and executive-risk communication

PROFESSIONAL OUTCOME · An evidence-backed AI risk and assurance package, ready for executive review.

The eighth and culminating lab is the Enterprise AI Security Capstone Lab, run across the final two weeks of the program and detailed separately in Section 8 as the program's integrated, end-to-end experience.

08 ENTERPRISE CAPSTONE EXPERIENCE

LAB 8 • CAPSTONE INTEGRATION EXPERIENCE • WEEKS 15-16



ENTERPRISE AI SECURITY ASSURANCE CAPSTONE

CHALLENGE STATEMENT • Step into the role of enterprise AI security assessor for a fictional organisation operating several interconnected AI systems, and deliver one coherent, end-to-end security assessment, from architecture review through governance sign-off, that a real security or executive committee could act on.

CAPABILITIES APPLIED

- End-to-end AI architecture and threat-model review
- Authorised assessment, detection design and defensive-control validation
- AI-assisted security-operations judgement
- Governance, risk and assurance packaging
- Executive-level technical communication

HOW THE TWO WEEKS RUN

The capstone runs across the final two weeks of the program: one week to review architecture, threat model and test plan under structured feedback, and one week to complete assessment, remediation validation, and deliver a stakeholder-ready oral security presentation with live executive Q&A.

PROFESSIONAL OUTCOME

A reviewed, presentation-ready enterprise AI security assessment package, and the ability to defend both technical evidence and business-risk conclusions to an executive audience.

09 PROFESSIONAL CAPABILITIES DEVELOPED

AI & LLM Security Foundations

- Explaining LLM behaviour, tokens, context and inference in security terms
- Mapping cybersecurity fundamentals onto AI-enabled architecture
- Identifying trust boundaries and security ownership across an AI system

Enterprise AI Application Security

- Assessing RAG, retrieval and knowledge-base security
- Evaluating conversation memory, AI data and API security
- Establishing and evidencing a clean security baseline

Agentic AI & MCP Security

- Assessing agent identity, delegated authority and least privilege
- Evaluating MCP trust boundaries and connected-tool ecosystems
- Recommending proportionate authority and monitoring controls

AI Threat Modeling & Test Planning

- Producing structured, AI-specific threat models
- Prioritising risk by architecture and business impact
- Defining authorised, evidence-driven security-test plans

Offensive AI Security & Red Teaming

- Conducting authorised adversarial testing across prompt, retrieval, agent, tool, MCP and API surfaces
- Capturing evidence that supports root-cause and impact analysis
- Differentiating manual testing from structured AI red teaming

AI Detection & Defensive Engineering

- Designing AI security telemetry and detections
- Implementing layered controls at the correct trust boundary
- Validating risk reduction through retesting

AI-Enabled Security Operations

- Using AI responsibly for log analysis, triage and threat hunting
- Applying human validation to manage hallucination risk
- Documenting AI-assisted analysis for audit and review

AI Governance, Risk & Assurance

- Building AI risk registers and control mappings
- Aligning assessments to recognised industry frameworks
- Producing assurance evidence and executive-ready risk communication

10 PROGRAM OUTCOMES

- Can assess, test, defend and govern enterprise AI systems across the full security lifecycle, not just discuss AI risk in the abstract
- Comfortable across the full stack: AI/LLM foundations, enterprise application security, agentic/MCP security, threat modeling, authorised testing, detection, defense and governance
- Leaves with a portfolio of evidenced assessments plus one enterprise-grade capstone security package
- Can communicate technical security evidence and business-risk conclusions to a non-technical executive audience
- Possesses a combined security-engineering and AI-security skillset sought by enterprises deploying AI at scale

11 CAREER ALIGNMENT

Learners completing this program are prepared for security and architecture roles across the enterprise AI stack, and for further specialisation into AI red-teaming or AI governance practice. The program's emphasis on evidence-based testing, defensive engineering, and assurance communication, alongside a demonstrable hands-on lab portfolio, is designed to support technical interview readiness at a practitioner and architecture level.

TARGET ROLES

AI Security Engineer

AI Red Team Specialist

AI Security Architect

Cloud & AppSec Engineer (AI)

AI Governance & Risk Specialist

12 PREREQUISITES

✓ REQUIRED KNOWLEDGE

- No prior AI security or AI red-teaming experience is required.
- Basic cybersecurity familiarity (networks, web applications, or general IT security) is helpful, but the program builds the necessary AI-specific foundation.
- Basic computer literacy and willingness to work with technical tools.
- Basic problem-solving ability and willingness to practise through guided, authorised security exercises.

☆ RECOMMENDED BACKGROUND

- Prior exposure to cybersecurity, IT, or software fundamentals is helpful, but not required.
- Basic familiarity with APIs, cloud services, or web applications is beneficial.
- Basic awareness of scripting or command-line environments is useful.

i HOW THE PROGRAM BUILDS TECHNICAL READINESS

The program begins with a controlled, no-assumptions introduction to AI and LLM behaviour, then progressively builds the cybersecurity and AI-architecture foundation required for enterprise AI security: from trust boundaries and application security through authorised testing, defense and governance.

⊗ NOT REQUIRED

- Deep machine learning or data-science background.
- Prior AI red-teaming or offensive-security experience.
- Existing security certification.
- Production AI-engineering experience.

13 DELIVERY APPROACH

- 80 instructor-led sessions delivered across 16 weeks (Monday to Friday), with one hour of structured learning per day. Total Instructor-Led Training: 80 Hours.
- Each session blends concept instruction, live demonstration, guided applied exercises in a controlled environment, and review
- Eight instructor-led hands-on labs (one per module, plus the capstone lab), each producing evidence for the learner's portfolio
- Individual and team-based applied work throughout
- Progressive, lab-based assessment culminating in a two-week capstone integration and executive-presentation phase

14 CURRICULUM GOVERNANCE NOTE

This curriculum is reviewed and refreshed on a regular cycle to keep pace with the fast-moving AI security threat and defense landscape, so that learners consistently train on current, industry-relevant security practice rather than a static snapshot of today's tools and techniques.

PROTECTED CURRICULUM IP

Detailed lab design, tool configuration, assessment payloads, and instructor delivery materials are maintained separately as protected curriculum intellectual property and are not part of this public curriculum overview.

LET'S BUILD YOUR **AI** CAREER.



WEBSITE

quantumacademics.ai

EMAIL

hello@quantumacademics.ai

PHONE / WHATSAPP

8098 809 809



CAMPUS

Jubilee Hills, Hyderabad, Telangana 500033



BOOK A FREE DEMO

quantumacademics.ai/book-demo