GSDC.
Global Skill Development Council

6-FRAMEWORK FIELD GUIDE | CGAIC CERTIFICATION | PRINTABLE PDF

The full 6-framework field guide.

The matrix in print. Per-framework usage notes. The 6-layer stack diagram. Per-framework hands-on artifacts from the CGAIC toolkit. Plus the 9-module syllabus and sample exam.

6	30+	2.5L+
FRAMEWORKS	LBD LABS	CERTIFIED PROS

Inside the toolkit:

- **6 frameworks compared on 8 dimensions**
- **6-layer stack diagram (printable)**
- **4 framework-specific artifact build sheets**
- **9 module syllabi · verbatim**
- **30+ Learn-by-Doing labs catalog**

Program: Certified Generative AI in Cybersecurity (CGAIC)

Exam: 40 MCQ · 90 min · free retake | **Duration:** 90 days

Used by 2,50,000+ certified professionals worldwide.

The 6 Frameworks · At a Glance

Six frameworks dominate AI cybersecurity work in 2026. Each has a different author, scope, and use-case. You don't pick one — you stack them. This page is the one-line introduction; the deep notes follow on pages 4–5.

1 OWASP LLM Top 10

OWNER · OWASP FOUNDATION · TYPE · VULNERABILITY LIST · BEST FOR · DEVELOPERS, APPSEC

The ten most critical security risks for applications built on top of LLMs — **prompt injection, sensitive info disclosure, model denial of service, supply-chain weaknesses, excessive agency, and more**. Application-layer focus. Updated regularly; the de facto starting list for AI security reviews.

2 MITRE ATLAS

OWNER · MITRE · TYPE · ADVERSARIAL TACTICS & TECHNIQUES · BEST FOR · THREAT INTEL, RED TEAMS, SOC

Adversarial Threat Landscape for AI Systems — the **MITRE ATT&CK analog for AI**. Catalogues attacker tactics, techniques, and real-world case studies against ML and generative-AI systems. Use for threat modelling and detection-mapping.

3 NIST AI RMF

OWNER · NIST (USA) · TYPE · RISK MANAGEMENT FRAMEWORK · BEST FOR · GOVERNANCE, RISK OFFICERS

The NIST AI Risk Management Framework — **voluntary but widely adopted** — defines the Govern / Map / Measure / Manage functions for AI risk across the full lifecycle. The single most-referenced risk-management framework in U.S. enterprise programmes.

4 ISO/IEC 42001

OWNER · ISO/IEC · TYPE · MANAGEMENT-SYSTEM STANDARD · BEST FOR · CERTIFICATION, REGULATED FIRMS

The international management-system standard for AI — **certifiable**, the way ISO/IEC 27001 is for information security. Defines what an AI Management System (AIMS) must contain. Use when you need third-party certification or are operating in EU-regulated markets.

5 OWASP ASVS

OWNER · OWASP · TYPE · APPLICATION SECURITY VERIFICATION STANDARD · BEST FOR · APPSEC TESTING, SDL

The Application Security Verification Standard — **not AI-specific, but the verification spine** AI features bolt onto. Defines verification levels (L1, L2, L3) for security controls; AI-specific extensions plug into the same scheme.

6 Microsoft AI Red Team

OWNER · MICROSOFT · TYPE · OPERATIONAL RED-TEAM METHODOLOGY · BEST FOR · OFFENSIVE TESTING

Microsoft's **published methodology for red-teaming generative AI** — what to test, in what order, with what

The 8-Dimension Matrix (Printable)

All six frameworks compared on eight practitioner-relevant dimensions. Tear this page out — it's the single-page reference you'll keep on your desk.

Framework	1 · Scope	2 · Layer	3 · Audience	4 · Format	5 · Update Cadence	6 · Certifiable	7 · Regulator Weight	8 · Effort
OWASP LLM Top 10	LLM application risks	App / feature	Developers, AppSec	Ranked list, ~10 items	Annual+	No	Med	Low
MITRE ATLAS	Adversary tactics for AI	Threat / attacker	Red team, SOC, TI	Matrix + case studies	Quarterly	No	Med	Med
NIST AI RMF	AI risk management	Enterprise / lifecycle	Governance, CISO	Functions + categories	Major revisions	No	High	Med
ISO/IEC 42001	AI management system	Enterprise / org	Compliance, audit	Clauses + Annexes	5-year revisions	Yes	High	High
OWASP ASVS	App security verification	App / code	AppSec, devs, QA	Verification levels L1–L3	Annual+	No	Med	Med
MS AI Red Team	Red-team methodology	Operational testing	Red team, security eng	Methodology + playbook	Continuous	No	Low	Med

How to read each dimension

- **Scope** · what the framework actually covers, in one phrase.
- **Layer** · which layer of the stack (page 6) it primarily addresses.
- **Audience** · who should keep it on their desk day to day.
- **Format** · is it a list, a matrix, a management system, a playbook?
- **Update cadence** · how often it changes; affects how stale your version gets.
- **Certifiable** · can a third party certify your organisation against it?
- **Regulator weight** · how heavily auditors and regulators reference it in 2026.
- **Effort** · rough first-deployment effort for a mid-sized enterprise.

Per-Framework Usage Notes (1 of 2)

Frameworks 1–3. For each: when to use it, when not to, and the one mistake most teams make.

FRAMEWORK 1 · OWASP LLM TOP 10

Use it as the starter list for every AI security review

Use when: kicking off a security review of an LLM feature, training developers, or scoping a pen-test engagement. **Don't use when:** you need a complete enterprise risk picture — it's deliberately a top-N list, not a comprehensive catalogue. **Common mistake:** treating the top-10 as exhaustive. Many real-world findings sit just outside the list (rate-limit abuse, prompt-template injection variants). Combine with ATLAS to round out the threat picture.

FRAMEWORK 2 · MITRE ATLAS

Use it to map adversary behaviour, not control coverage

Use when: threat-modelling an AI system, designing detections, or briefing leadership on attacker capability. **Don't use when:** trying to prove control coverage to an auditor — ATLAS describes attacker behaviour, not defender controls. **Common mistake:** conflating ATT&CK and ATLAS coverage. Many SOCs claim "ATLAS coverage" while their detections only fire on classical MITRE techniques. Build per-tactic detection traceability before claiming coverage.

FRAMEWORK 3 · NIST AI RMF

Use it as the spine of your AI governance programme

Use when: building an enterprise AI governance programme from scratch, briefing the board, or aligning with a U.S. regulator's expectations. **Don't use when:** you need certifiable evidence — NIST is voluntary, not certifiable. Pair with ISO 42001 if you need an external badge. **Common mistake:** writing policy that quotes Govern/Map/Measure/Manage as labels without operationalising them. Each function needs concrete owners, evidence, and review cadence — or the framework adds zero protection.

Frameworks 1–3 cover the breadth — application risks, attacker behaviour, and the enterprise risk spine. None alone is enough; together they're 80% of the practical coverage most teams need.

Per-Framework Usage Notes (2 of 2)

Frameworks 4–6. Same structure: when, when not, and the common mistake.

FRAMEWORK 4 · ISO/IEC 42001

Use it when you need a certifiable AI management system

Use when: you operate in EU-regulated markets, supply enterprise customers who demand certification, or are aligning with a global compliance posture alongside ISO 27001. **Don't use when:** you're a small team without governance maturity — the management-system overhead is significant. **Common mistake:** trying to certify before the underlying risk-management work is real. ISO 42001 audits the *system*; if the system is paper-only, you'll fail the audit and burn budget. Land NIST AI RMF first, then layer ISO on top.

FRAMEWORK 5 · OWASP ASVS

Use it as the verification spine your AI controls plug into

Use when: defining acceptance criteria for new code, structuring pen-tests, or measuring AppSec maturity. **Don't use when:** you need AI-specific guidance — ASVS is not AI-aware in its base form. Use it for the surrounding application; use OWASP LLM Top 10 for the AI-specific layer. **Common mistake:** picking ASVS Level 3 across the board. L3 is for high-assurance systems; most products should target L2 with L3 only on critical paths. Avoid blanket-L3 mandates.

FRAMEWORK 6 · MICROSOFT AI RED TEAM

Use it as the operational playbook for red-team engagements

Use when: standing up an internal AI red team, scoping a third-party AI red-team engagement, or training existing red-teamers on AI specifics. **Don't use when:** you need an exhaustive attack catalogue — that's MITRE ATLAS. Microsoft's methodology is operational ("how we red-team") more than encyclopedic. **Common mistake:** running an AI red team without responsible-disclosure and safety guardrails. Many findings are dual-use; without a clear handling process, you ship attacker capability rather than reducing it.

How the 6 fit together (in one paragraph)

OWASP LLM Top 10 anchors the **application-layer review**. MITRE ATLAS describes the **adversary** you're defending against. NIST AI RMF gives you the **enterprise risk spine**. ISO/IEC 42001 turns that spine into a **certifiable management system**. OWASP ASVS provides the **verification grammar** for code-level controls. Microsoft AI Red Team gives you the **operational playbook** for offensive testing. You stack — you don't pick.

The 6-Layer Stack Diagram (Printable)

Every AI system you'll secure sits on a 6-layer stack. Each layer has its own primary frameworks. Print this page; it's the second desk reference, paired with the matrix on page 3.

6	**Layer 6 · Governance & Audit** Board reporting, risk register, model inventory, third-party assessments, regulator-facing artefacts. **Primary: ISO/IEC 42001 · Supporting: NIST AI RMF**
5	**Layer 5 · Enterprise Risk Management** Govern / Map / Measure / Manage. Risk taxonomy, control library, lifecycle gates, oversight committees. **Primary: NIST AI RMF · Supporting: ISO/IEC 42001**
4	**Layer 4 · Adversary & Threat Modelling** Who attacks, how they attack, what evidence appears in your telemetry. Detection-engineering targets. **Primary: MITRE ATLAS · Supporting: MS AI Red Team**
3	**Layer 3 · Application Risks (LLM-specific)** Prompt injection, sensitive info disclosure, excessive agency, supply-chain weaknesses, model DoS. **Primary: OWASP LLM Top 10 · Supporting: OWASP ASVS**
2	**Layer 2 · Application Security Verification** Authentication, authorisation, input/output validation, session handling, cryptography, error handling, logging. **Primary: OWASP ASVS · Supporting: OWASP LLM Top 10**
1	**Layer 1 · Operational Red-Team Testing** Active offensive testing of the live system. Engagement scoping, technique selection, evidence capture, remediation hand-off. **Primary: MS AI Red Team · Supporting: MITRE ATLAS · OWASP LLM Top 10**

How to read the stack

- **Layers 1–3** are operational — what you do *to* the system (test, verify, find risks).
- **Layers 4–6** are organisational — what you do *about* the system (model the adversary, manage risk, govern).
- The frameworks are **layered, not exclusive**. Mature programmes touch every layer.
- Beginning programmes typically start at Layer 3 (OWASP LLM Top 10) and Layer 5 (NIST AI RMF) — the two highest-leverage entry points.

How to Combine the Frameworks · 4 Patterns

Different organisational starting points need different framework stacks. These four patterns cover ~90% of practitioner scenarios. Pick the closest, adapt at the edges.

PATTERN A · STARTUP / SCALE-UP

OWASP LLM Top 10 + MS AI Red Team

You are: Series B–E, 50–500 people, shipping an AI product. **You need:** defensible security posture without compliance overhead. **The minimum stack:** bake the LLM Top 10 into your SDL; commission a quarterly AI red-team engagement using the Microsoft methodology as scope. Skip ISO 42001 until customers demand it. Skip ATLAS as a primary unless you have a dedicated threat-intel function.

PATTERN B · ENTERPRISE NEW TO AI

NIST AI RMF + OWASP LLM Top 10 + OWASP ASVS

You are: a Fortune 1000 firm beginning AI adoption, with existing AppSec maturity. **You need:** a defensible governance posture and product-level controls. **The minimum stack:** NIST AI RMF as the enterprise spine, LLM Top 10 + ASVS for product-level controls. ATLAS as a stretch goal in Year 2 once your SOC catches up. ISO 42001 only if your audit committee asks for certification.

PATTERN C · REGULATED ENTERPRISE

ISO/IEC 42001 + NIST AI RMF + OWASP LLM Top 10 + MITRE ATLAS

You are: banking, pharma, insurance, healthcare, public sector, or an EU-presence enterprise. **You need:** certifiable controls, regulator-defensible governance, and live threat intelligence. **The minimum stack:** all four. NIST AI RMF as the operating spine, ISO 42001 for the external badge, LLM Top 10 + ATLAS at the operational layer.

PATTERN D · AI VENDOR / FRONTIER LAB

MS AI Red Team + MITRE ATLAS + OWASP LLM Top 10 + ISO/IEC 42001

You are: a foundation-model vendor, AI-native scale-up, or specialist consultancy. **You need:** deep offensive capability plus credible governance posture for your customers. **The minimum stack:** MS AI Red Team as the operational engine, ATLAS as the adversary catalogue, LLM Top 10 as the application-layer baseline, ISO 42001 to win enterprise contracts.

What ties all four together: none of the patterns skips OWASP LLM Top 10. It is the universal entry point. Beyond that, your industry and compliance posture decide which of the other five frameworks join the stack — and in which order.

Artifact Build Sheet 1 · OWASP LLM Top 10

FRAMEWORK · OWASP LLM TOP 10

LLM-Feature Security Review Report

Time: 6–10 hours
Output: 8–12 page report + 1-page exec summary
Tooling: Threat-model template, test prompt library
Used in: AppSec review board, interviews

1. Pick one shipped (or in-design) LLM feature and write a one-paragraph description of its purpose, data sources, and user surface.
2. For each of the 10 risk categories, write 2–3 sentences: is the feature exposed, how, and what evidence backs that claim.
3. Run a short test prompt library against the feature (or a clone). Capture screenshots/transcripts for any successful inducement.
4. For each non-trivial finding, propose a mitigation in the format "control · owner · cost-band · acceptance criteria."
5. Score the feature L1 / L2 / L3 on each category and explain the bands in a single sentence.
6. Write the 1-page exec summary last — three bullets on top risks, one paragraph on residual risk after the proposed mitigations.

Defending this artefact in interview / audit

- **Walk the worst finding** end-to-end: how you found it, how it would be exploited, how the proposed mitigation closes it.
- **Defend one judgement call** — typically a finding you *didn't* raise. Show that you considered it and have a reason it sits below the line.
- **Have a numeric residual-risk view**: which categories remain orange/red after mitigations, and the rationale for accepting that residual.

Common scoring mistakes

- Marking every category as "applicable" without evidence — looks thorough, reads as un-prioritised.
- Confusing severity (impact) with exposure (likelihood). Score both.
- Mitigations without owners. Without a named owner, the mitigation is a wish.

Artifact Build Sheet 2 · NIST AI RMF

FRAMEWORK · NIST AI RMF

Enterprise AI Risk Programme Skeleton

Time: 12–20 hours
Output: Programme charter + control matrix + RACI
Tooling: RMF function template, control library
Used in: Board / risk committee briefings, CISO interview rounds

1. Write a 1-page programme charter — purpose, scope, executive sponsor, success metrics, decision rights.
2. For each of the four RMF functions (Govern · Map · Measure · Manage), list 3–5 categories with named owners and review cadence.
3. Build a control matrix: rows = your AI use-cases (model inventory), columns = the RMF categories. Each cell holds the control state (Not Started / In Progress / Operational / Validated).
4. Write a one-paragraph RACI for the top 5 controls: Responsible / Accountable / Consulted / Informed.
5. Define 5 KPIs the board will see monthly — three operational (e.g., # high-risk models without conformity assessment) and two strategic (e.g., audit-finding closure rate).
6. Draft a 6-month roadmap with three quarterly checkpoints and a clear definition of "complete v1."

Defending this artefact

- **Walk the model inventory** first. If the inventory isn't credible, the rest of the programme isn't credible.
- **Explain how a high-risk model gets governed differently** from a low-risk model. The risk tier must drive concrete control differences.
- **Defend one KPI** the way you'd defend a board OKR — definition, owner, target, why it matters.

What good vs paper-only looks like

- **Good:** every control has a named owner with the authority to act, evidence is captured per quarter, and the board KPIs reflect actual operating reality (not the policy text).
- **Paper-only:** the matrix is full of "In Progress," there's no named owner, and the board KPI is "policy approved" (a one-off, not a recurring metric).

Artifact Build Sheet 3 · MITRE ATLAS

FRAMEWORK · MITRE ATLAS

ATLAS Detection Coverage Heatmap

Time: 8–12 hours
Output: Detection-coverage matrix + 5 detection rules
Tooling: SIEM (Splunk / Sentinel / Elastic), ATLAS navigator
Used in: SOC capability briefings, red/purple-team retros

1. List the tactics from MITRE ATLAS relevant to your AI use-cases. Don't try to cover every tactic — pick the ones that match your system architecture.
2. For each tactic, list 2–3 techniques and write one-line attacker narratives in your environment. Specificity matters — "prompt injection" is vague, "RAG-document poisoning via vendor-portal upload" is useful.
3. Build a heatmap: rows = tactics, columns = data sources you already collect (proxy, EDR, model gateway, vector DB audit). Mark each cell as Green/Amber/Red for detection feasibility.
4. Write 5 high-leverage detection rules in SIEM-portable pseudocode. Aim for one rule per tactic at first.
5. Document the known false-positive patterns for each rule and the tuning plan.
6. Brief your team on three "intentional gaps" — places you've decided not to invest yet and why. Honesty about gaps is a credibility multiplier.

Defending the heatmap

- **Walk one detection rule**: how it fires, what it would miss, how you tested it.
- **Defend a Red cell**: why you can't detect that tactic today and what the plan is to move it to Amber.
- **Show the residual**: what an attacker could still do that you'd not see, and your compensating controls.

Mapping ATLAS back to your existing ATT&CK detections

Many AI attacks chain into classical MITRE ATT&CK behaviour once the initial AI-specific tactic succeeds (e.g., post-exploitation lateral movement). Build an ATLAS → ATT&CK bridge table so your SOC isn't surprised when an "AI attack" turns into "credential dumping" two hops later.

Artifact Build Sheet 4 · ISO 42001 + ASVS

FRAMEWORK · ISO/IEC 42001 + OWASP ASVS

AIMS Readiness Pack with ASVS Verification

Time: 14–22 hours

Output: Readiness assessment + control evidence index

Tooling: ISO 42001 clause map, ASVS L2 checklist

Used in: Internal audit, pre-certification audit, vendor due diligence

1. Build a clause-by-clause readiness assessment against the ISO 42001 Annex. For each clause, capture the current state, evidence reference, and gap.
2. Map each clause to existing controls from ISO/IEC 27001 or NIST AI RMF where they overlap — avoid rebuilding the wheel.
3. For software systems in scope, run an ASVS L2 verification. Capture pass/fail per chapter, with a remediation backlog for failures.
4. Build a single evidence index — every clause/control has one or more pointer(s) to artefacts (policies, tickets, screenshots, audit logs).
5. Write a 2-page exec summary: readiness percentage, top three risks before audit, recommended timeline to certification (typically 6–12 months for prepared orgs).
6. Identify the AI-specific control gaps that the underlying ISMS/NIST work doesn't already cover — those are where ISO 42001 actually adds work.

Why ISO + ASVS together

ISO 42001 audits the *system*; ASVS verifies the *code*. An ISO auditor will accept an AppSec verification standard as evidence for operational security controls. Bundling them avoids two parallel control libraries and gives your developers a clean acceptance checklist they actually use.

What's typically missing on first attempt

- **AI impact assessments** per use-case — not just the policy, the actual completed assessments.
- **Model-card discipline** — a model card per deployed model, refreshed at the cadence the policy claims.
- **Incident response specifically for AI** — generic IR runbooks don't cover prompt-injection or supply-chain weights poisoning.
- **Third-party AI governance** — vendor assessments, data-flow agreements, lock-in/exit terms.

9-Module Syllabus · CGAIC (Verbatim)

All 9 modules of the Certified Generative AI in Cybersecurity program. Each framework above is taught inside one or more modules; the mapping is on page 13.

MODULE 01
Foundations · LLMs for Security Pros
How LLMs work end-to-end at the depth a security professional needs. Tokenisation, attention, RAG, agents, tool calls.

MODULE 02
AI Threat Landscape
MITRE ATLAS taxonomy, OWASP LLM Top-10, attacker motivations, AI-specific kill chain. Maps to traditional MITRE ATT&CK.

MODULE 03
Gen-AI Phishing & Social Engineering
AI-generated phishing, deepfake voice/video, BEC variants, detection signatures, user-side defences.

MODULE 04
AI-Augmented Malware
Polymorphic payloads, AI-generated obfuscation, prompt-injection-based C2, defender techniques.

MODULE 05
Prompt Injection & LLM Exploitation
Direct + indirect injection, jailbreak chains, model extraction, training-data leakage, embedding attacks.

MODULE 06
Secure-by-Design for AI Systems
Guardrails, input/output filters, scope-limiting agents, threat modelling for AI features. Heavy on ASVS.

MODULE 07
MLOps Security & Supply Chain
Model registry, signing & provenance, supply-chain attacks, monitoring, rollback, secret scanning.

MODULE 08
AI Governance, Risk & Compliance
NIST AI RMF, ISO/IEC 42001, EU AI Act, NYC LL 144, board reporting, vendor governance.

MODULE 09
Capstone · Defend & Certify
Pick your pathway, build the three artefacts, defend in front of an evaluator, earn the CGAIC credential.

Total program time: 90 days, 6–8 hours per week. Exam format: 40 multiple-choice questions, 90 minutes, free retake on first failure.

🚀 NEXT COHORT STARTING SOON

Join the next CGAIC cohort with this guide in hand

You've now seen the syllabus. The next cohort uses this exact framework stack — applying now earns the launch window discount.

Join The Next Cohort →

Module-to-Framework Mapping

Which CGAIC module covers which framework, at what depth. Use this to plan your study path if you're prioritising a specific framework.

Module	OWASP LLM 10	ATLAS	NIST AI RMF	ISO 42001	ASVS	MS Red Team
M01 · Foundations	Intro	Intro	Intro	Intro	Intro	Intro
M02 · Threat Landscape	Heavy	Core	—	—	—	Med
M03 · GenAI Phishing	Med	Med	—	—	—	Med
M04 · AI-Augmented Malware	Med	Heavy	—	—	—	Med
M05 · Prompt Injection	Core	Heavy	—	—	Med	Heavy
M06 · Secure-by-Design	Heavy	Med	Med	Med	Core	Med
M07 · MLOps Security	Med	Med	Heavy	Med	Heavy	Med
M08 · Governance, Risk & Compliance	—	—	Core	Core	Med	—
M09 · Capstone	Defend	Defend	Defend	Defend	Defend	Defend

How to read this

- **Core** · the framework is the primary teaching target of that module.
- **Heavy** · the framework is a primary supporting reference, covered in depth.
- **Med** · the framework appears in context with working knowledge expected.
- **Intro** · introductory coverage only; foundations established for later modules.
- **Defend** · the framework features in your capstone evidence.

If you only have time for half the program and you want framework breadth: prioritise Modules 2, 5, 6, 7, 8 — these cover all six frameworks at Med or higher.

30+ Learn-by-Doing Labs · Catalog (1–16)

Each lab is a time-boxed, evaluator-reviewed exercise tied to one or more frameworks. You finish each lab with an artefact you can show in interviews or use on the job.

- 01 LLM Top 10 · Feature Review
- 02 Prompt-Injection Attack Lab
- 03 Indirect-Injection via RAG
- 04 Output-Filter Bypass Bench
- 05 ATLAS Threat-Model Workshop
- 06 ATLAS → ATT&CK Bridge Table
- 07 Detection-Coverage Heatmap
- 08 SIEM Rule · AI Phishing Pattern
- 09 SIEM Rule · Prompt-Injection C2
- 10 Deepfake-Voice Detection Triage
- 11 NIST RMF · Programme Charter
- 12 NIST RMF · Control Matrix Build
- 13 NIST RMF · Board KPI Pack
- 14 ISO 42001 · Readiness Pack
- 15 ISO 42001 · Evidence Index
- 16 AI Incident Response Runbook

How a lab is structured

- A **2–4 hour** time-box with a clear deliverable.
- A guided **prompt + dataset + tool stack** you can replicate in your environment.
- **Evaluator review** on output quality plus written feedback to apply on the next lab.
- A reusable **portfolio artefact**: model card, detection rule, control matrix, audit memo, or playbook.

30+ Learn-by-Doing Labs · Catalog (17–32)

The second half of the labs catalog focuses on architecture, MLOps security, advanced red-team, and capstone-track artefacts.

- **17** ASVS L2 · Verification Sprint
- **18** Guardrail Kit · Working Code
- **19** Secure RAG Architecture Build
- **20** Vector-DB Security Audit
- **21** MLOps Pipeline · Signing & Provenance
- **22** Supply-Chain · LoRA Backdoor Lab
- **23** Model Registry Hardening
- **24** Model Monitoring & Drift Alerts
- **25** MS AI Red Team · Engagement Scope
- **26** Jailbreak Chain · Reproducible PoC
- **27** Red-Team Report · OWASP LLM Format
- **28** Remediation Memo · Builder-Friendly
- **29** Vendor Governance Assessment
- **30** EU AI Act · Risk Classification Lab
- **31** Board-Pack One-Pager · AI Risk
- **32** Capstone Build & Defence

If you only have time for six labs

The minimum portfolio for an AI-security interview loop, in priority order: **Lab 1** (LLM Top 10 review), **Lab 7** (Detection coverage), **Lab 11** (NIST programme charter), **Lab 18** (Guardrail kit), **Lab 25** (Red-team engagement scope), **Lab 32** (Capstone). Together they touch every framework at a defensible depth.

What's not in the lab catalog (and why)

- **Production penetration testing** against third-party systems — out of scope for legal reasons; engagements happen against approved targets only.
- **Vendor-specific tooling certifications** — CGAIC is vendor-neutral by design. You'll touch many tools; you won't earn a tool badge.
- **Pure data-science labs** — model training, fine-tuning. CGAIC focuses on securing AI, not building it.

Sample Exam · 6 Framework-Focused Questions

Six representative questions across the 6 frameworks. The real CGAIC exam is 40 MCQ in 90 minutes with a free retake.

Q1 · OWASP LLM TOP 10

A chatbot exposes administrator capabilities when asked in Pig Latin. The most accurate OWASP LLM Top 10 category for this finding is:

(a) LLM01 · Prompt Injection.
(b) LLM06 · Sensitive Information Disclosure.
(c) LLM08 · Excessive Agency.
(d) LLM10 · Model Theft.

Q2 · MITRE ATLAS

An attacker uploads poisoned documents to a vendor portal that feeds a customer-facing RAG application. In MITRE ATLAS, the most accurate tactic for this initial step is:

(a) Resource Development (preparing attacker infrastructure).
(b) Initial Access via Supply Chain Compromise.
(c) Execution via Command-Line Interface.
(d) Discovery via Cloud Service Discovery.

Q3 · NIST AI RMF

In NIST AI RMF, which function explicitly covers organisational culture, accountability, and policy?

(a) Govern.
(b) Map.
(c) Measure.
(d) Manage.

Q4 · ISO/IEC 42001

ISO/IEC 42001 is best described as:

(a) A technical control catalogue for AI systems.
(b) A certifiable management-system standard for an AI Management System (AIMS).
(c) A US-only regulation for federal AI procurement.
(d) A vulnerability list comparable to OWASP LLM Top 10.

Q5 · OWASP ASVS

For a high-value financial application built on an LLM, the appropriate ASVS verification target for the critical authentication path is:

(a) Level 1.
(b) Level 2 across the board with Level 3 on the critical path.
(c) Level 3 across the entire application by default.
(d) ASVS does not apply to LLM applications.

FAQs · Honest Answers Before You Enrol

Do I need to memorise all 6 frameworks?

No. You need **working fluency in the concepts and structure** of all 6, deep operating proficiency in 2–3, and the ability to pick the right one for a given problem. The exam tests structural understanding; the capstone tests applied skill. Memorising clauses is wasted effort.

Will frameworks change before I finish?

Yes — OWASP LLM Top 10 and MITRE ATLAS update quarterly to annually. The certification is designed around *structural* knowledge that survives version bumps. Lab content tracks the latest published version at cohort start.

How is CGAIC different from CISSP or CISM?

CISSP and CISM cover the classical security and management body of knowledge; neither covers OWASP LLM Top 10, MITRE ATLAS, NIST AI RMF, ISO 42001, OWASP ASVS, or Microsoft AI Red Team in depth. CGAIC is purpose-built for AI security work — designed to complement existing certifications, not replace them.

What's the exam format?

40 multiple-choice questions, 90 minutes, free retake on first failure. Pass mark is ~70%. The capstone is separate — three artefacts plus a 30-minute defence with an evaluator.

How long until I see a salary uplift?

Median time-to-offer for certified candidates in AI security roles is 5–7 weeks after credentialing. Mid-career uplift typically lands at 18–28% on an external move. In-role bumps run lower (8–12%) but cost less to negotiate.

What if my employer pays for it?

Most learning budgets cover vendor-neutral certifications. The enrolment receipt is corporate-tax friendly in most jurisdictions. Forward this guide and the syllabus on page 12 to your L&D lead — most clear in under a week.

Glossary & About This Guide

Glossary

- **CGAIC:** Certified Generative AI in Cybersecurity — the GSDC vendor-neutral AI security credential.
- **OWASP LLM Top 10:** The current OWASP top-10 application-security risks specific to LLM applications.
- **MITRE ATLAS:** The Adversarial Threat Landscape for AI Systems — MITRE's tactics-and-techniques framework for AI attacks.
- **NIST AI RMF:** The U.S. NIST AI Risk Management Framework, with four functions: Govern, Map, Measure, Manage.
- **ISO/IEC 42001:** The international standard for an AI Management System (AIMS) — certifiable, like ISO 27001.
- **OWASP ASVS:** The Application Security Verification Standard, with three verification levels (L1, L2, L3).
- **Microsoft AI Red Team:** Microsoft's published operational methodology for red-teaming generative AI systems.
- **AIMS:** AI Management System — the structured governance system that ISO 42001 audits.
- **RAG:** Retrieval-Augmented Generation — architecture where an LLM is grounded with retrieved documents at query time.
- **Prompt injection:** An attack where the model is induced to ignore its system instructions via crafted input.

About the Global Skill Development Council

GSDC is a global, independent skill-certification body building worldwide credentials for the future of work. The CGAIC program is part of GSDC's portfolio of AI-era professional certifications — designed with practitioners, validated by mentors actively working in the field, and trusted by 2,50,000+ certified professionals across 45+ countries.

Verifying your credential

Once you complete the 40-MCQ assessment and the capstone defence, your CGAIC credential is issued with a unique verification ID. Recruiters and hiring managers can verify the credential directly on the GSDC registry — no third-party validation needed.

The 6-Framework Field Guide · On One Page

The 6 frameworks (pages 2, 4, 5)

OWASP LLM Top 10 · MITRE ATLAS · NIST AI RMF · ISO/IEC 42001 · OWASP ASVS · Microsoft AI Red Team. Stack them, don't pick.

The 8-dimension matrix (page 3)

Scope · Layer · Audience · Format · Update cadence · Certifiable · Regulator weight · Effort. Single-page desk reference comparing all six on every dimension.

The 6-layer stack (page 6)

Governance & Audit · Enterprise Risk · Adversary Modelling · Application Risks · App Security Verification · Operational Red-Team. Every framework lives at a primary layer with supporting roles elsewhere.

The 4 combination patterns (page 7)

Startup/Scale-up · Enterprise new to AI · Regulated enterprise · AI vendor/frontier lab. Pick the closest match and adapt at the edges.

The 4 framework artefacts (pages 8–11)

OWASP LLM Top 10 review report · NIST AI RMF programme skeleton · MITRE ATLAS detection heatmap · ISO 42001 readiness pack with ASVS verification. The four-artefact portfolio is what hiring managers ask about.

How CGAIC fits (pages 12–13)

9 modules · 30+ labs · 40-MCQ exam · capstone defence. Vendor-neutral. 90-day format designed to run alongside a full-time security role.