GSDC.
Global Skill Development Council

AI CYBERSECURITY JOBS BRIEF | CGAIC CERTIFICATION | PRINTABLE PDF

The full AI cybersecurity jobs brief.

The 9-module syllabus, the 5-role salary table, the employer placement map, the AI cybersecurity jobs market data, and the sample exam.

5	9	2.5L+
HIRING ROLES	MODULES	CERTIFIED PROS

Inside the toolkit:

- **9 module syllabi · verbatim**
- **5-role salary table · USA 2026**
- **Employer placement map**
- **Sample exam · printable checklist**
- **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.

AI Cybersecurity Jobs Market — 2026 Snapshot

Five distinct AI cybersecurity roles dominate hiring in 2026. They sit across SOC, threat intel, ML engineering, red team, and platform engineering. One CGAIC credential opens all five. Five headline numbers from analysing 3,200+ live AI-security job postings between Jul 2024 and Dec 2025.

FINDING 01 · DEMAND

AI cybersecurity postings grew 187% year-over-year

Active job postings tagged with AI/ML security skills tripled from Jul 2024 to Dec 2025. The five roles in this brief make up ~82% of those postings. The remaining ~18% are hybrid titles where AI security is a stated responsibility but not the primary focus.

FINDING 02 · SUPPLY GAP

~4 open AI-security roles per qualified candidate

For every actively-hiring AI-security role, there are roughly 0.25 qualified candidates in the market. The supply gap is widest at Mid and Senior bands — where employers most need 3–6 years of security experience plus demonstrable AI fluency. Certified candidates fill this gap directly.

FINDING 03 · MEDIAN PAY

AI Security Engineer median $186K · GenAI Red Teamer $215K

USA tier-1 metro median total compensation. AI Security Engineer leads in volume; GenAI Red Teamer leads in pay; ML Security Engineer at $195K sits between. AI SOC Analyst is the highest-volume entry-level path; AI Threat Intel Analyst is the strongest pivot for traditional threat-intel professionals.

FINDING 04 · CERTIFICATION EFFECT

Certified candidates close offers 2.1× faster

Among 814 placements tracked, candidates holding a vendor-neutral AI-cybersecurity certification reached final-stage offer in a median of 5.0 weeks vs 10.6 weeks for non-certified peers applying to the same role family. Recruiter shortlist time shrank 49%.

FINDING 05 · GEOGRAPHIC SPREAD

63% of open roles are remote-eligible or hybrid

Unlike traditional SOC roles (28% remote-eligible), AI security roles skew heavily remote-friendly — the work is research, build, and audit-heavy rather than facility-bound. Tier-1 US metros still anchor pay, but candidates outside them can compete for ~63% of openings without relocating.

5-Role Salary Table · USA 2026

All five hiring roles on one page. Total comp = base salary + target bonus + equity at vest, normalised to a 4-year vest. USA tier-1 metro baseline. Use the regional multipliers in the salary report to localise.

Role	Junior	Mid Median	Senior	Lead/Staff
AI Security Engineer	$148K	$186K	$248K	$298K
GenAI Red Teamer	$152K	$215K	$268K	$305K
ML Security Engineer	$152K	$195K	$262K	$298K
AI Threat Intel Analyst	$138K	$172K	$225K	$278K
AI SOC Analyst	$129K	$165K	$215K	$278K

Median total compensation by role (mid-career, USA tier-1)

Sources triangulated: Glassdoor (n = 1,240), ZipRecruiter (n = 880), Levels.fyi (n = 410), GSDC partner panel of 12 employers. Outliers winsorised at 1st/99th percentile.

Role 1 · AI Security Engineer

1

AI Security Engineer

Builds and hardens production AI systems — guardrails, MLOps security, secure RAG pipelines, model-registry hardening. The platform-engineer track.

Mid-career median · $186K · USA tier-1 baseline · highest postings volume

WHAT YOU DO

Threat-model and harden production AI features. Ship LLM guardrails, build hardened RAG pipelines, secure agentic workflows, configure model-registry signing and provenance, integrate audit logging into the model gateway. The role sits between platform engineering and security; you're the engineer the SOC asks for help, and the engineering org asks for review.

TOOL STACK YOU'LL USE

Python (heavy), one cloud (AWS / Azure / GCP), Guardrails AI or NeMo Guardrails, LangChain or Llama-Index for RAG audits, a model registry (MLflow / SageMaker / Vertex), pytest for regression, terraform for IaC. SIEM access for detection integration. Most candidates have ~70% of this; the AI-specific 30% is what certification fills.

WHAT INTERVIEWS PROBE

Architecture review on a RAG system (expect cross-tenant leak and prompt-injection scenarios); guardrail design trade-offs; one live coding round on input/output filters; threat-modelling a vendor LLM integration. **Two artifacts hiring managers ask to see:** a working guardrail kit and a hardened RAG repo with three test corpora.

HOW TO POSITION YOURSELF

If you're a backend or platform engineer: lead with your production systems work, layer AI security on top. If you're a traditional security engineer: lead with threat modelling and audit, show one working AI-tooling integration. Cert + 1–2 working artifacts beats a master's degree from a non-AI background.

Role 2 · GenAI Red Teamer

2

GenAI Red Teamer

Leads offensive AI engagements — prompt injection, jailbreak chains, model extraction, agentic exploitation. Reports map to OWASP LLM Top 10 and MITRE ATLAS.

Mid-career median · $215K · highest pay band of the 5 · scarcest talent pool

WHAT YOU DO

Run sanctioned offensive engagements against AI features — prompt injection (direct + indirect), jailbreak chains, RAG poisoning, training-data extraction, model-inversion attempts, tool-call hijacks in agentic systems. Deliverable is a structured engagement report mapping findings to OWASP LLM Top 10 and MITRE ATLAS with reproducible PoCs.

TOOL STACK YOU'LL USE

garak, promptmap, custom payload generators, Burp Suite or self-built fuzzers, Python for harness work, Git for PoC repos, screen recording for demos. Lower code burden than AI Security Engineer; higher creative-attack burden. Strong overlap with traditional red-team craft, augmented with AI-specific attack patterns.

WHAT INTERVIEWS PROBE

"Show your most interesting jailbreak" — they want the full chain, not just the final payload. Discussion of methodology vs trophy hunting. One live exercise against a sandboxed LLM. Discussion of responsible disclosure for AI vendor findings. **The artefact:** a polished engagement report (OWASP-aligned) with at least 5 reproducible findings is what closes the offer.

HOW TO POSITION YOURSELF

Strongest pivot from web/app pentesting. If you have OSCP-level offensive credentials plus 3–5 reproducible AI exploits, you're competitive at Senior. Frontier AI labs, Big Tech AI red teams, and Big-4 offensive practices are the highest-pay employers; sovereign-linked engagements pay tax-free at parity.

Distinguishing GenAI Red Teamer from traditional Red Teamer

- Traditional red team: lateral movement, persistence, exfil. **GenAI red team:** model-level attacks, agent hijacks, retrieval poisoning.
- Traditional pay band: $145–230K mid-career. **GenAI red team:** $172–268K mid-career. The gap is the AI-attack-pattern specialisation.
- Many hires are dual-track — traditional offensive base + AI-specific layer. CGAIC fills the AI layer; OSCP/OSWE fill the base.

Role 3 · ML Security Engineer

3

ML Security Engineer

Secures the ML/AI training and deployment pipeline — model integrity, training-data security, supply chain, MLOps security, drift and rollback. The MLOps-meets-security track.

Mid-career median · $195K · USA tier-1 baseline · second-highest pay band

WHAT YOU DO

Secure the full ML lifecycle — training data integrity, dataset signing, model-registry security, signed adapters and LoRA review, supply-chain controls on public hubs, drift monitoring with security signals, rollback procedures. Detect and respond to model-level supply-chain attacks like LoRA backdoors and poisoned training data.

TOOL STACK YOU'LL USE

MLflow / Weights & Biases, model-signing tooling (cosign-equivalent for ML), DVC or LakeFS for data versioning, terraform/pulumi for infra, SBOM generation for ML artifacts, drift monitors (Evidently / Arize / WhyLabs), one cloud (AWS / Azure / GCP). Heavy on supply-chain hygiene tooling.

WHAT INTERVIEWS PROBE

Threat-model the model-training pipeline — expect questions on training-data poisoning, dataset access controls, dependency hygiene for ML packages, third-party model hub trust assumptions. One scenario question on responding to a LoRA backdoor disclosure. **Artefact to lead with:** a signed-and-provenanced model registry pipeline with at least one rollback runbook.

HOW TO POSITION YOURSELF

Strongest pivot from MLOps engineer or DevSecOps with ML exposure. Combine traditional supply-chain skills (SLSA, signing, attestation) with ML-specific knowledge (LoRA, fine-tuning, vector-store hygiene). Big Tech AI infra teams, frontier AI lab security teams, and large model platforms are highest-pay employers; F500 enterprises with serious AI deployments pay slightly below.

Role 4 · AI Threat Intel Analyst & Role 5 · AI SOC Analyst

4

AI Threat Intel Analyst

Tracks AI-attacker TTPs, maps to MITRE ATLAS, briefs the SOC, and feeds detection-engineering with AI-specific signatures. The strongest pivot for traditional threat-intel professionals.

Mid-career median · $172K · USA tier-1 baseline

WHAT YOU DO & HOW YOU GET INTERVIEWED

Translate emerging AI-attacker research (red-team disclosures, academic papers, frontier-lab safety publications) into detection rules and threat briefings. Build an internal ATLAS coverage heatmap, run quarterly purple-team validation, brief the CISO on AI-attacker trajectory. **Interviews probe:** "Walk me through one ATLAS technique you'd build detection for, end-to-end" — typical answer combines a paper-derived TTP, a SIEM detection rule, and a purple-team validation plan. **Tools:** SIEM (Splunk / Sentinel / Elastic), MITRE ATLAS Navigator, OSINT for vendor research, Markdown for briefings.

5

AI SOC Analyst

Detection & triage of AI-powered attacks at L2/L3 — AI phishing, deepfake voice/video BEC, agentic SOC triage operation. Highest postings volume; cleanest entry-level path.

Mid-career median · $165K · USA tier-1 baseline · 28% of all AI-sec postings

WHAT YOU DO & HOW YOU GET INTERVIEWED

Investigate AI-related alerts at L2/L3 — AI-generated phishing campaigns, voice-clone vishing, AI-augmented BEC, agentic-system triage. Operate AI-augmented SOC tools (auto-enrichment agents, agentic triage pipelines). **Interviews probe:** walking through one alert end-to-end, distinguishing AI-generated phishing from a sophisticated human one, judging when to trust the agentic triage recommendation vs override. **Tools:** SIEM/SOAR, EDR, AI-augmented triage pipelines (e.g. LangGraph builds), email-gateway integration. Strongest entry-level role for security analysts pivoting into AI; lowest barrier to first interview.

Employer Placement Map · Where the Roles Go

Where certified candidates actually placed across 814 tracked placements. Six employer categories cover ~90% of the market. Each card shows the role mix and the named anchors that frequently appear in postings.

Big Tech & Hyperscalers

Pay index **1.28×** · placement share **22%** · strongest for AI Security Engineer + ML Security Engineer + GenAI Red Teamer

- Microsoft · Azure AI security org
- Amazon · Bedrock & AI safety
- Apple · AI security platform
- Google · DeepMind safety
- Meta · GenAI safety
- NVIDIA · model risk org

Frontier AI Labs · safety/security teams

Pay index **1.25×** · placement share **11%** · strongest for GenAI Red Teamer + ML Security Engineer

- Anthropic · red-team & eval
- Cohere · model safety
- xAI · adversarial robustness
- OpenAI · red-team & trust
- Mistral · safety org
- Inflection · alignment safety

Banking, Insurance & Financial Services

Pay index **1.05×** · placement share **20%** · strongest for AI SOC Analyst + AI Threat Intel + AI Security Engineer

- JP Morgan · global AI security
- HSBC · AI risk & controls
- Citi · AI governance
- Goldman Sachs · GenAI controls
- BlackRock · model-risk security
- Mastercard / Visa · AI security ops

Pharma · Healthcare · Life Sciences

Pay index **0.95×** · placement share **9%** · strongest for AI Security Engineer + AI Threat Intel

- Pfizer · AI risk & safety
- Roche · model risk
- HCA · AI SOC ops
- Novartis · GenAI controls
- UnitedHealth · AI threat intel
- Mayo / Cleveland Clinic · AI risk

Big-4 / Consulting · AI Security Practices

Pay index **1.00×** · placement share **14%** · all 5 roles · strongest for GenAI Red Teamer + AI Threat Intel

- Deloitte · AI risk & cyber
- PwC · AI trust & security
- Accenture · cyber AI
- EY · GenAI assurance
- KPMG · AI risk advisory
- Bain & BCG · AI risk practices

MSSP / Pure-Play Cyber + Government

Pay index **0.85×** · placement share **14%** · highest volume for AI SOC Analyst

- CrowdStrike · MDR AI ops
- Mandiant (Google) · AI threat intel
- US Federal civilian (cleared)
- Palo Alto · Unit 42 AI
- Wipro / TCS / Infosys GCC
- MITRE · AI risk research

Hiring Signals · How to Read AI-Security JDs

Most AI-security postings are still written by recruiters new to the space. The signals below tell you what the role *actually* is, beyond the title. Use this when triaging which postings to apply to.

If the JD says…	…it likely means	Map to role
"AI Security Engineer" + cloud requirement	Platform-engineering track; expect IaC/Python and one cloud at depth.	AI Security Engineer
"GenAI Red Team" + OSCP	Offensive engagement role; expect production-grade exploit-chain demos in interviews.	GenAI Red Teamer
"AI/ML Security" + MLflow / model registry	MLOps-meets-security; supply-chain and pipeline hygiene focus.	ML Security Engineer
"AI Threat Intel" + ATLAS / Mandiant style	Research-heavy; SIEM-detection bridge; quarterly briefing rhythm.	AI Threat Intel Analyst
"SOC Engineer · AI" + Splunk / Sentinel	L2/L3 detection + AI-augmented triage; volume-hiring path.	AI SOC Analyst
"AI Risk" + NIST AI RMF / EU AI Act	Governance-leaning; less hands-on engineering, more policy + audit.	Hybrid · Threat Intel + Governance
"AI Trust & Safety" at Big Tech	Mix of red team + policy + product; senior bands; high cross-functional load.	Hybrid · Red Team + Governance

Red flags in AI-security JDs (skip or push back)

- **"5+ years AI security experience"** — the discipline is younger than that. Push back; the requirement is decorative, not enforced.
- **No mention of OWASP LLM Top 10 or MITRE ATLAS** — the role is more buzzword than substance. Apply with caution; ask in interview which framework the team uses.
- **"Build LLM defences from scratch with no tooling budget"** — the role is set up to fail. Ask what budget exists for LLM API costs and red-team tooling.
- **Title says "AI" but JD body is 90% traditional security** — the role is being relabelled for budget. Pay band will be traditional security, not AI.

> *JDs lie. Hiring managers don't. Get to a hiring manager call before applying to anything you're unsure about.*

9-Module CGAIC Syllabus (Verbatim)

All 9 modules of the Certified Generative AI in Cybersecurity program. One credential maps to all 5 hiring roles above. Each role's pre-interview prep concentrates on different modules — see the role-to-module mapping on page 11.

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, GAN/VAE anomaly detection, 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, agentic security, threat modelling for AI features.

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 3 artifacts, defend in front of an evaluator, earn the CGAIC credential. The deliverable hiring managers ask about.

Total program time: 90 days · 6–8 hours per week. Exam format: 40 MCQ, 90 min, free retake.

Role-to-Module Mapping Matrix

All five roles share the same nine modules — but pre-interview depth differs sharply. Use this matrix to know which modules to over-prepare for your target role.

Module	AI Sec Eng	GenAI Red Team	ML Sec Eng	AI Threat Intel	AI SOC Analyst
M01 · Foundations · LLMs	Core	Core	Core	Core	Core
M02 · Threat Landscape	Medium	Heavy	Medium	Heavy	Heavy
M03 · GenAI Phishing	Aware	Medium	Aware	Heavy	Heavy
M04 · AI-Augmented Malware	Medium	Medium	Heavy	Heavy	Heavy
M05 · Prompt Injection & LLM Exploitation	Heavy	Core	Medium	Medium	Medium
M06 · Secure-by-Design	Core	Medium	Heavy	Aware	Medium
M07 · MLOps Security	Heavy	Aware	Core	Medium	Aware
M08 · Governance & Compliance	Medium	Medium	Medium	Heavy	Medium
M09 · Capstone	Defend	Defend	Defend	Defend	Defend

How to read this

- **Core** · the module is the primary skill anchor for the role. Spend the most hands-on time.
- **Heavy** · the module supports the role at depth; expect interview questions here.
- **Medium** · working knowledge expected; can be probed but not the focal point.
- **Aware** · understand the vocabulary and structure. Don't be the person who's never heard of it.
- **Defend** · the module's labs become your capstone defence.

Quick "where to focus" by role

- **AI Security Engineer:** M05 + M06 + M07 (build/harden side).
- **GenAI Red Teamer:** M02 + M05 (offensive depth, threat-landscape fluency).
- **ML Security Engineer:** M06 + M07 (secure-by-design + supply chain).
- **AI Threat Intel Analyst:** M02 + M03 + M04 + M08 (research + governance).
- **AI SOC Analyst:** M02 + M03 + M04 (detection + threat-pattern fluency).

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

Each lab is a time-boxed, evaluator-reviewed exercise. You finish each lab with an artefact you can show in interviews or reuse on the job. The list below is the first half; the second half is on page 13.

- 01 LLM Tokeniser & Embedding Lab
- 02 Prompt-Injection Attack Lab (basic)
- 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 Build
- 08 SIEM Rule · AI Phishing Pattern
- 09 SIEM Rule · Prompt-Injection C2
- 10 Deepfake-Voice Detection Triage
- 11 GAN Anomaly Detector · Train
- 12 GAN Detector · Eval Harness
- 13 VAE Behaviour Engine · Train
- 14 RL Responder · Gym Setup
- 15 RL Responder · Reward Shaping
- 16 AI Incident Response Runbook

Which labs map to which interview

- **AI Security Engineer interview:** Labs 03, 04, 18, 19 (RAG injection, filter bypass, guardrail build, hardened RAG).
- **GenAI Red Teamer interview:** Labs 02, 03, 04, 26, 27 (injection chain, RAG poisoning, jailbreak chain, OWASP-format report).
- **ML Security Engineer interview:** Labs 21, 22, 23, 24 (signing/provenance, LoRA backdoor lab, registry hardening, drift monitoring).
- **AI Threat Intel interview:** Labs 05, 06, 07, 08, 09 (threat-model workshop, ATLAS bridge, coverage heatmap, SIEM rules).
- **AI SOC Analyst interview:** Labs 08, 09, 10, 16, 28 (SIEM rules, deepfake triage, response runbook, agentic SOC).

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

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

- **17** ASVS L2 Verification Sprint
- **18** Guardrail Kit · Working Code
- **19** Hardened RAG Architecture Build
- **20** Vector-DB Security Audit
- **21** MLOps · 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 Format
- **28** Agentic SOC Triage · LangGraph Build
- **29** Vendor Governance Assessment
- **30** EU AI Act · Risk Classification Lab
- **31** Board-Pack One-Pager · AI Risk
- **32** Capstone Build & Defence

Lab → portfolio artefact mapping

- **Lab 11 + 12** → GAN anomaly detector (with eval harness) — strong for ML Security Engineer interview.
- **Lab 18** → Working LLM guardrail repo on GitHub — strong for AI Security Engineer interview.
- **Lab 19** → Hardened RAG architecture with three test corpora — strong for AI Security Engineer architecture round.
- **Lab 27** → A full OWASP-format red-team engagement report — the single artefact that closes GenAI Red Teamer offers.
- **Lab 28** → Agentic SOC triage pipeline (LangGraph) — strong for AI SOC Analyst senior pivot.
- **Lab 32** → Capstone defence — your three chosen artifacts presented to an evaluator.

The minimum-viable interview portfolio

Six labs cover every role at a defensible depth: **Lab 11–12** (GAN train + eval), **Lab 18** (guardrail kit), **Lab 19** (hardened RAG), **Lab 27** (OWASP red-team report), **Lab 28** (agentic SOC), and **Lab 32** (capstone). Most candidates ship 8–12 labs; the six above are the must-haves.

The AI-Security Interview Process

The interview loop is consistent across the five roles. Five stages, ~10–14 calendar days end-to-end for fast movers. Knowing the stages is half the prep.

STAGE 1 · RECRUITER SCREEN (30 MIN)

Hits one of three failure modes for most candidates

Recruiter screens budget, geography, certifications, and one technical anchor. Failures usually come from: (1) salary expectation not anchored to USA tier-1 mids, (2) no certification flagged on resume, (3) no link to a portfolio artefact in the first response. Bring all three.

STAGE 2 · HIRING MANAGER (45 MIN)

"Walk me through one of your artefacts in 5 minutes"

Pick the artefact most closely matched to the role — guardrail kit for AI Security Engineer, OWASP report for GenAI Red Teamer, signed registry for ML Security Engineer. Walk it top-to-bottom in 5 minutes. The remaining 40 minutes are questions; the artefact framing controls them.

STAGE 3 · TECHNICAL DEEP DIVE (60–90 MIN)

One live exercise per role, plus 30 minutes of probing

Engineer roles: architecture review or live coding. Red team: live jailbreak attempt against a sandboxed LLM. ML Security: threat-model the training pipeline. Threat intel: pick an ATLAS technique, design detection. SOC: triage a live alert. Practice the exercise pattern, not just the underlying skill.

STAGE 4 · CROSS-TEAM PANEL (60 MIN)

Tests the working-relationship dimension

2–3 cross-functional peers (engineering, product, GRC) probe how you communicate, document, escalate, and disagree. The single biggest fail point: candidates who can build but not *describe*. If you've defended a capstone in front of an evaluator, you've practised this.

STAGE 5 · OFFER + NEGOTIATION (3–10 DAYS)

Where the salary-table data on page 3 matters

Anchor at the 75th-percentile cell for your role and level. Separate base / equity / signing. Negotiate title before number where possible — title sets the next 3 years of band moves. Ask for AI-tooling budget in writing. Most certified candidates land 12–18% above their first offer through standard negotiation.

Sample Exam — Part 1 of 2

Six representative questions from the CGAIC exam. The real exam is 40 MCQ in 90 minutes with a free retake on first failure. Answers at the end of part 2.

Q1 · MODULE 05 · PROMPT INJECTION

A customer-support chatbot ignores its system prompt 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 · MODULE 06 · SECURE-BY-DESIGN

A RAG application retrieves documents from a vector store that any tenant can write to. The single highest-impact mitigation to ship first is:

(a) Add a profanity filter on the LLM response.

(b) Add a per-tenant retrieval scope so retrieval only returns documents owned by the requesting tenant.

(c) Increase the LLM temperature to add response variety.

(d) Cache responses for 1 hour.

Q3 · MODULE 02 · 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.

(b) Initial Access via Supply Chain Compromise.

(c) Execution via Command-Line Interface.

(d) Discovery via Cloud Service Discovery.

Sample Exam — Part 2 of 2

Q4 · MODULE 07 · MLOPS SECURITY

You discover a LoRA adapter pulled from a public hub introduces a backdoor that activates on a specific trigger phrase. The most appropriate immediate control is:

(a) Block all public LoRA sources at the artifact-registry layer; require signed, internally-reviewed adapters only.

(b) Increase logging granularity on the model gateway.

(c) Add a trigger-phrase regex to the input filter.

(d) Quarantine the affected user.

Q5 · MODULE 04 · GAN ANOMALY DETECTION

Your GAN-based anomaly detector trains stably but converges to a generator that produces only one type of benign sample. The correct diagnosis is:

(a) Discriminator overfitting; add dropout.

(b) Mode collapse; introduce mini-batch discrimination or Wasserstein-GAN with gradient penalty.

(c) Learning rate too low; raise it 10×.

(d) Insufficient training data; the architecture is fine.

Q6 · MODULE 08 · GOVERNANCE

Under the EU AI Act, an LLM-based resume screener used in EU hiring is most accurately classified as:

(a) Minimal risk — no obligations.

(b) Limited risk — transparency obligations only.

(c) High risk — full conformity assessment, registration, and human-oversight obligations apply.

(d) Prohibited — cannot be deployed in the EU.

Answer key

Q1 — a · Q2 — b · Q3 — b · Q4 — a · Q5 — b · Q6 — c

If you scored 5–6 of 6

You already think like an AI-security practitioner. Pick the harder role tracks — GenAI Red Teamer or AI Security Engineer at Senior band.

If you scored 3–4 of 6

Foundations are solid; you have governance and edge-case gaps. The 90-day program is well-paced for you — most candidates in this band land at 90%+ on the real exam.

Pre-Application Checklist · Printable

Tear this page out (or print it). Run this checklist before applying to any of the five roles. Every box you can tick lifts your interview rate; every box you can't is a 60-minute fix before applying.

Role & positioning

- ✓ You've picked **one** of the five roles as your primary target.
- ✓ You've identified one **fallback role** from the other four.
- ✓ Your resume top-line matches the **role title** you've picked.
- ✓ Your LinkedIn headline matches the role title (recruiters search by exact phrase).

Artefacts ready (link in resume header)

- ✓ **One technical artefact** matched to the role (guardrail / red-team report / GAN detector / ATLAS heatmap / agentic SOC).
- ✓ **One model card or runbook** — proves you ship docs alongside code.
- ✓ **One audit-style artefact** — OWASP audit kit, vendor evaluation, or red-team report.
- ✓ GitHub link is in your resume header and confirmed working.

Credentials

- ✓ CGAIC **verification ID** in resume header.
- ✓ LinkedIn certifications section updated with the credential.
- ✓ One traditional-security credential (Security+ / CISSP / OSCP) if you have it — list under CGAIC.
- ✓ One cloud credential (AWS / Azure / GCP fundamentals) if you have it.

Application logistics

- ✓ Salary anchor pre-decided using the page 3 table at the 75th-percentile cell.
- ✓ Geography filter set — remote/hybrid/on-site picked before applying.
- ✓ Three companies pre-targeted from the employer placement map (page 8).
- ✓ One warm intro pursued in parallel — never just cold-apply for senior bands.

Interview prep ready

- ✓ You can walk your strongest artefact in **under 5 minutes**.
- ✓ You can name one failure mode and one improvement you'd make in a v2.
- ✓ You've run through the stage-by-stage interview structure on page 14.
- ✓ You have one peer or mentor who'll do a mock loop with you.

> *Most candidates fail at the recruiter screen, not the technical round. Get the resume header, credential, and one artefact link right — recruiter screen goes from 16% pass-through to 33%.*

Glossary & About This Brief

Glossary

- **CGAIC:** Certified Generative AI in Cybersecurity — GSDC's vendor-neutral AI-security certification.
- **Total comp:** Base salary + target bonus + equity at vest, normalised to a 4-year vest. Excludes one-off signing bonuses.
- **USA tier-1 metro:** San Francisco, New York, Seattle, DC, Boston — the highest-pay USA metros where the baseline salary table on page 3 applies.
- **Time-to-offer:** Calendar weeks from first recruiter contact to written offer. Measured for tracked placements where both candidate and recruiter confirmed dates.
- **MITRE ATLAS:** The Adversarial Threat Landscape for AI Systems — MITRE's tactics-and-techniques framework for AI attacks.
- **OWASP LLM Top 10:** The current OWASP top-10 application-security risks specific to LLM applications.
- **LoRA:** Low-Rank Adaptation — a parameter-efficient fine-tuning method; LoRA adapters can carry backdoors if sourced from untrusted hubs.
- **SBOM:** Software Bill of Materials — emerging analogue for ML artifacts; provides supply-chain provenance.
- **Recruiter screen:** The first 30-minute call. ~16% non-certified vs ~33% certified pass-through rate.
- **Capstone defence:** 30-minute live evaluation where you walk an evaluator through your 3 chosen artifacts.

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 on 3 artifacts, 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 Full AI Cybersecurity Jobs Brief · On One Page

Market snapshot (page 2)

Postings grew 187% YoY. ~4 open roles per qualified candidate. AI Security Engineer median $186K · GenAI Red Teamer $215K. 63% of postings are remote-eligible or hybrid.

The 5 hiring roles (pages 3–7)

AI Security Engineer · GenAI Red Teamer · ML Security Engineer · AI Threat Intel Analyst · AI SOC Analyst. Different pay bands, different tool stacks, one certification (CGAIC) opens all five.

Employer placement map (page 8)

6 employer categories cover ~90% of placements: Big Tech 22% · F-Services 20% · Big-4 14% · MSSP+Gov 14% · Frontier AI Labs 11% · Pharma 9%. Pay indices from 0.85× (MSSP) to 1.28× (Big Tech).

Hiring signals & JD decoder (page 9)

How to read AI-security JDs beyond the title. Red flags: "5+ years AI security experience," no OWASP/ATLAS reference, no tooling budget, AI-prefix on traditional security body.

9-module CGAIC syllabus (pages 10–11)

Foundations · Threat Landscape · GenAI Phishing · AI-Augmented Malware · Prompt Injection · Secure-by-Design · MLOps · Governance · Capstone. Role-to-module depth matrix on page 11.

30+ LBD labs (pages 12–13)

Each lab is 2–4 hours, evaluator-reviewed, ships a reusable artefact. 6-lab minimum-viable portfolio: 11, 12, 18, 19, 27, 28 + capstone.

Interview process (page 14)

5-stage loop — recruiter screen · hiring manager · technical · cross-team panel · offer. Most candidates fail at stage 1, not stage 3. Resume header + credential + artefact link fix it.