Practical DevSecOps

CONSULTING GUIDE | NOVEMBER 2026

KINGDOM OF SAUDI ARABIA | NCA AICG-1:2026

AI Security Readiness in Saudi Arabia

October 2026 showed how fast AI changes the attack. This guide turns the NCA draft AI Cybersecurity Guidelines into controls, a 20-point checklist and a 90-day plan you can start on Monday.

PREPARED BY

Practical DevSecOps Consulting Team

BOOK A 30-MIN CALL →

AI is shipping faster than AI security. October proved it.

Saudi organisations are putting generative AI and agents into customer service, software delivery and operations. Here is what matters in four lines, and where to find the detail.

01 **The risk is already priced in**

SAR 27M is the average breach in the Kingdom. Public-facing applications, which now include AI chat and APIs, are the costliest way in at SAR 34.7M.

02 **October changed the clock**

PHP2Shell was exploited about five hours after its details leaked and over six hours before the fix shipped. AI makes finding and weaponising bugs cheap.

03 **The expectation is written down**

The NCA draft AICG-1:2026 sets 42 guidelines across 4 domains and names generative and agentic AI in scope. It reuses the ECC structure you know.

04 **DevSecOps is the cheapest fix**

IBM links a DevSecOps approach, EDR and encryption to the lowest breach costs in Saudi Arabia. Build AI controls into the pipeline once, reuse forever.

Inside this guide

		PAGE
01	October 2026: the vulnerability storm	03
02	AI found the bugs. Attackers moved faster	04
03	What a breach costs in Saudi Arabia	05
04	AI in the attack, and the defence	06
05	The regulatory picture	07
06	Mindmap of the draft AICG-1:2026	08
07	Anatomy of an AI system	09
08	Ten guidelines that need new engineering	10
09	From guideline to control and evidence	11
10	20-point readiness checklist	13
11	Score and 90-day plan	14
12	How we help	15

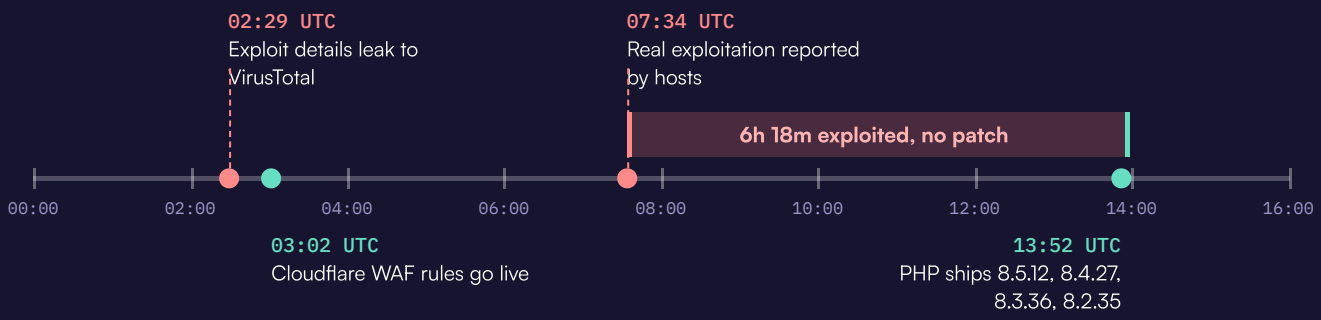
One month, one lesson: attackers now move in hours.

A record Patch Tuesday, firewall zero-days, a PHP flaw exploited before its fix, and a wave of bugs in AI agents and gateways. All in 31 days.

PHP2Shell: 21 October, hour by hour

CVE-2026-109618 | CVE-2026-109619 | CVSS 9.2

Two memory bugs in PHP itself, present since PHP 5.4. Every server running PHP 5.4 or later was exposed.

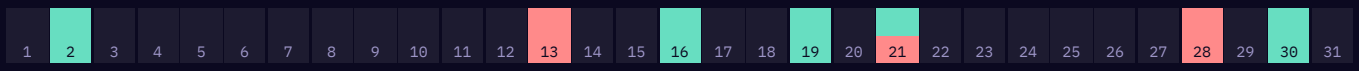


~70%
of websites run PHP

33 min
from leak to the first WAF rules

2 countries
France and the Netherlands took government sites offline to patch

The rest of the month



2 OCT **AWS Loom agent platform** CVE-2026-103956
Full admin to any network client when no identity provider was set

13 OCT **Microsoft Patch Tuesday** CVE-2026-96951
Record 1,411 flaws. Windows Shell MotW bypass exploited as a zero-day

16 OCT **Claude Agent SDK** CVE-2026-96620
Crafted session ID injects an MCP server and runs commands past permissions

19 OCT **Envoy AI Gateway and Agent Mesh** CVE-2026-103091
Any OIDC token holder could self-enrol and call every MCP tool

21 OCT **LiteLLM and agentic intrusions** Supply chain
Malware slipped into LiteLLM release packages. CrowdStrike reports four actors using AI agents in intrusions

21 OCT **PHP2Shell** CVSS 9.2
Unauthenticated remote code execution in PHP, exploited before the patch

28 OCT **Cisco ASA and FTD firewalls** CVE-2026-76418
Zero-day exploited to crash SSL VPN devices

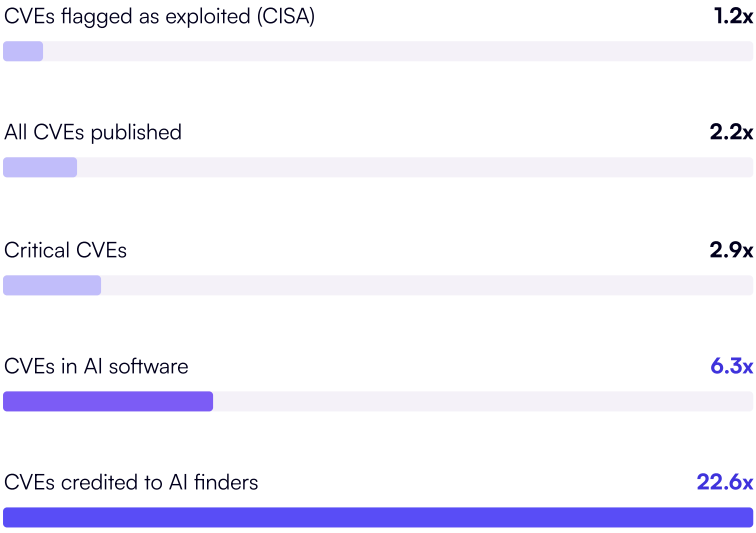
30 OCT **TAKO AI Agent for Okta** CVSS 9.8
Unauthenticated first-run admin takeover of an AI agent

AI found the bugs. Attackers moved faster.

Disclosures have doubled and AI-credited finds are up more than 20 times. Exploitation grew far more slowly. What really changed is speed, and AI software itself has become a target.

Growth, 2025 to 2026

Same window, 1 January to 23 October. Multiplier shown.



86,292

CVEs published by 23 October, more than all of 2025 (48,151)

34 to 769

CVEs crediting an AI system or AI lab as the finder

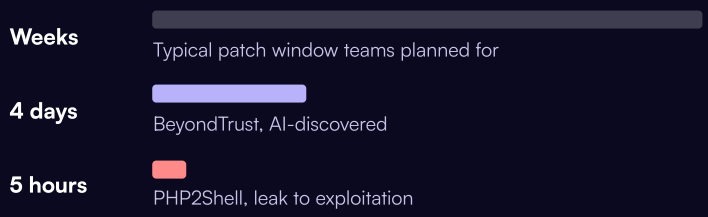
50% vs 26%

Share of likely AI-found flaws that lead to remote code execution, against other flaws

DISCLOSURE TO EXPLOIT

4 days

An AI-found BeyondTrust flaw (CVE-2026-1731) was exploited by one cluster within four days, and five more within a week.



What it means for an AI estate in the Kingdom

01 AICG 2-1

Know what you run

3,191 CVEs hit AI software this year. Half sit in agent orchestration tools. An AI inventory is step one.

02 AICG 2-6-3

Patch at machine speed

You need a kill switch for vulnerable AI features and a WAF rule path that ships in minutes.

03 AICG 2-2

Guard the gateway

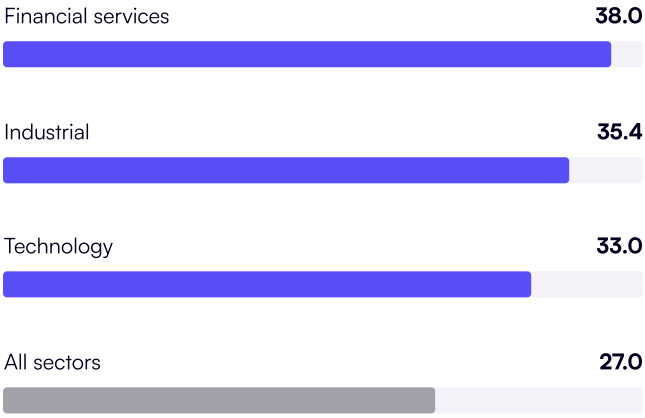
An AI gateway holds every provider key behind it. Vault the keys, scope the tokens, rate-limit.

What a breach costs in Saudi Arabia

IBM and the Ponemon Institute studied 602 breached organisations worldwide, including Saudi Arabia, for incidents between March 2025 and February 2026.

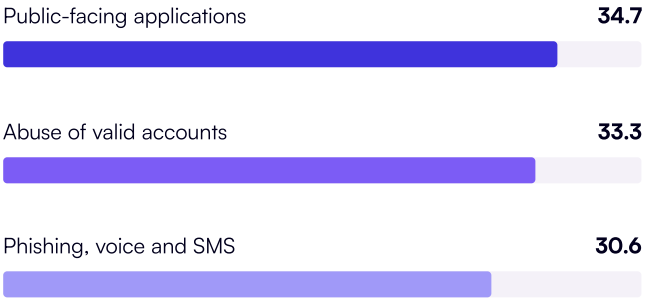
By sector

SAR MILLION, AVERAGE PER BREACH



By attack vector

SAR MILLION, AVERAGE PER BREACH



AI chat interfaces and APIs are public-facing applications. They sit in the most expensive category.

The 200-day line

Under 200 days: **SAR 22M** | Over 200 days: **SAR 32M**



CUTS THE COST

1. A DevSecOps approach
2. Endpoint detection and response
3. Encryption of sensitive data

RAISES THE COST

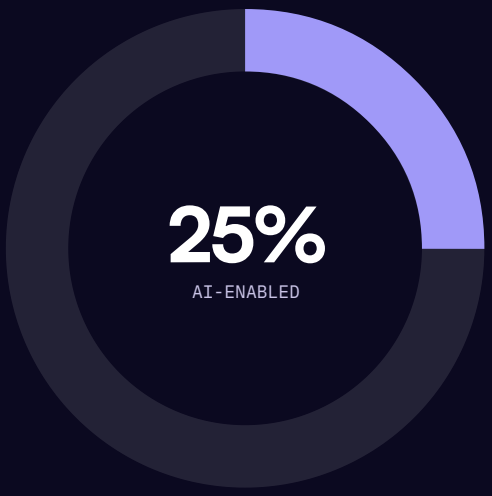
1. Mismanaged secrets and keys
2. Missing security skills
3. Poor threat prioritisation

37%

Only 37% of breached Saudi organisations encrypted sensitive data at rest and in transit.

AI is in the attack. It is also the cheapest defence.

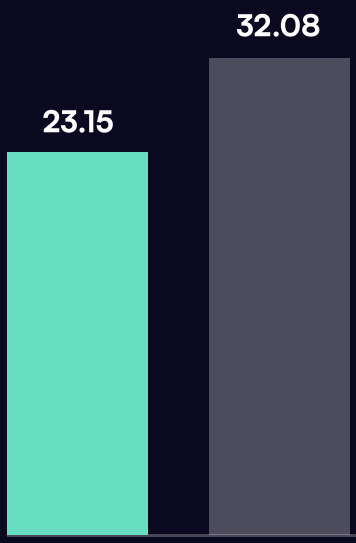
Attackers use AI to move faster. Defenders who use AI and automation well spend far less per breach. Every AI system you deploy is also a new asset to protect.



Of malicious breaches in Saudi Arabia used AI.
Middle East: 26%, with another 11% unsure.

The automation dividend

Average breach cost in the Kingdom, SAR million



SAR 8.9M
saved with extensive use. Yet 25% of Saudi organisations use none.

No. 1
cost driver in the Middle East: mismanaged secrets and keys

55%
of regional organisations with a SOC run AI agents inside it

16%
of regional breaches began in the supply chain, at USD 8.45M each

Why this matters for AI systems

Agents run on service accounts, API keys and tokens. The region's top cost drivers are exactly the controls agents depend on. Teams secure non-human identities mainly through a machine identity inventory (57%) and by extending zero trust to AI-driven processes (43%).

The rules are converging on AI

The AI guidelines apply requirements the NCA already sets elsewhere to AI systems. A working ECC programme can be extended rather than rebuilt.



Why act on a draft?

The draft calls the guidelines recommended, not mandatory.

01 It restates NCA requirements you may already be measured on.

02 PDPL enforcement already targets weak technical security in AI data flows.

03 Clients and regulators will ask AI security questions shaped by this text.

04 Controls built now become evidence later.

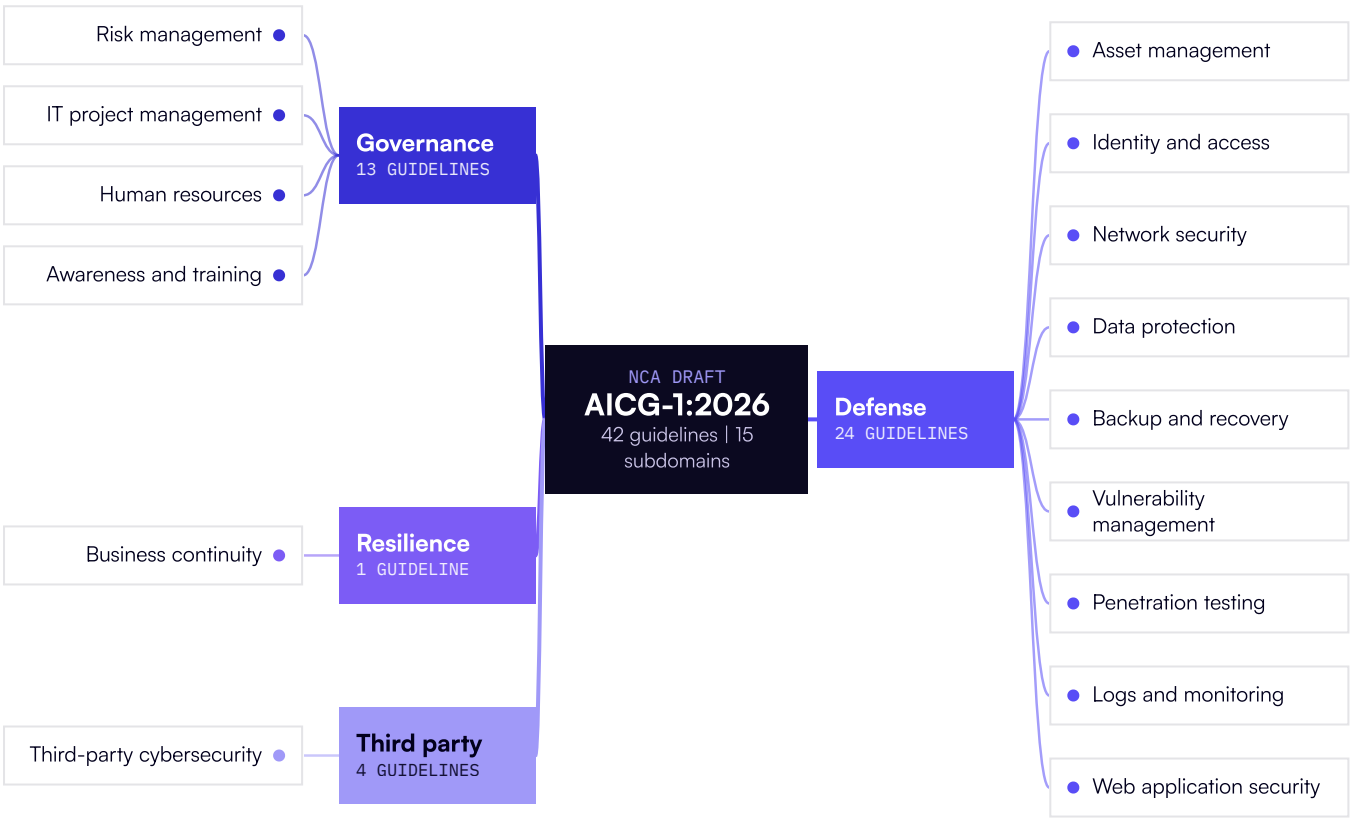
How the frameworks stack

Each layer builds on the one below it.



42 guidelines on one page

Four domains, fifteen subdomains, the same shape as the ECC. Defense carries more than half of the work.





SHARE OF 42 GUIDELINES BY DOMAIN

RESILIENCE 1

Lifecycle in scope

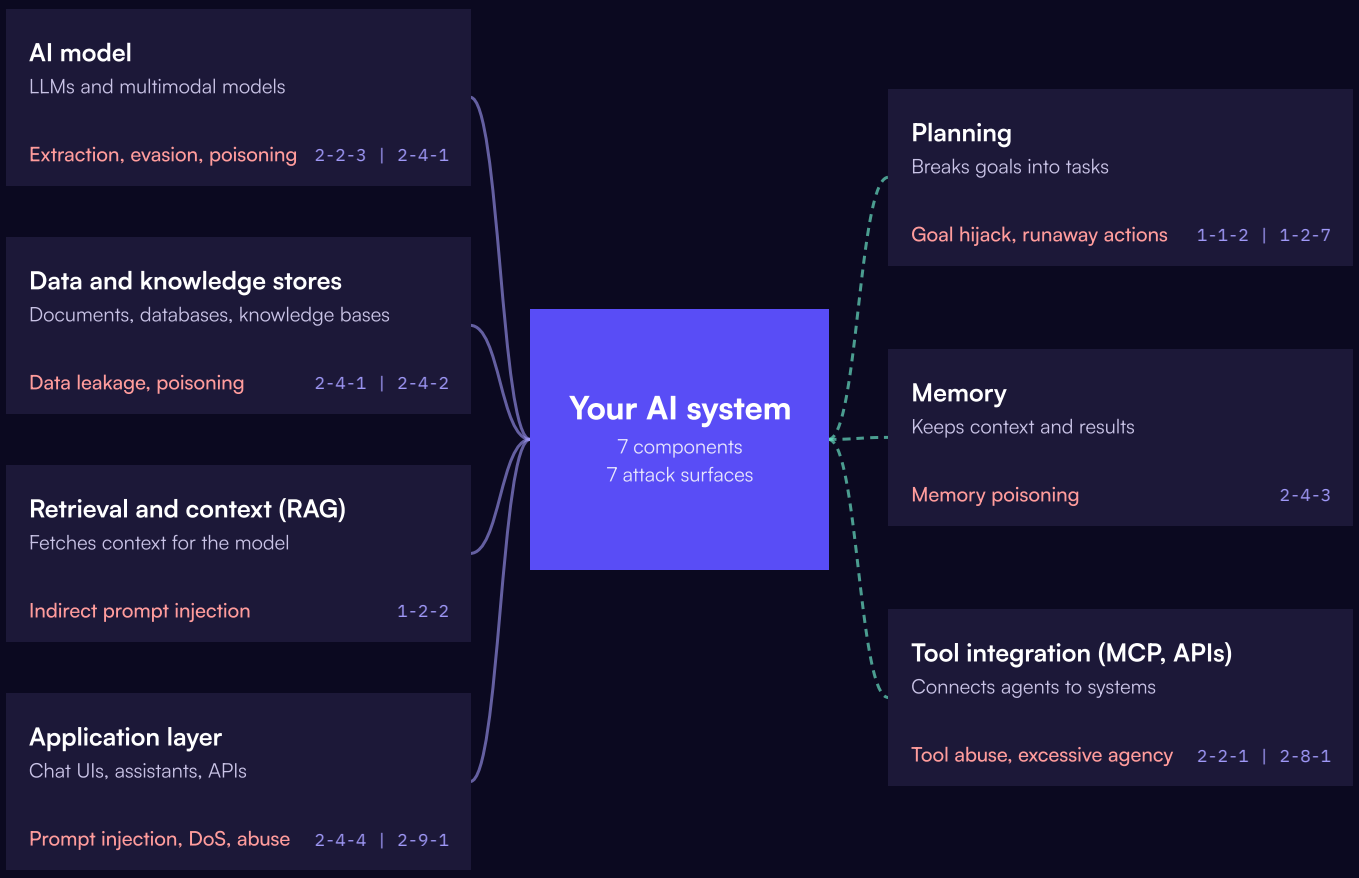


Seven components. Seven ways in.

Four components make a generative AI system. Agents add three more that let it plan and act. Each maps to specific draft guidelines.

GENERATIVE AI CORE

AGENTIC EXTENSIONS



October 2026, mapped to the diagram

AI AGENT CVES

8 OCT

Pydantic AI web chat

Any website the developer visits can run the agent and its approval-gated tools

APPLICATION LAYER

16 OCT

Claude Agent SDK

Crafted session ID injects an MCP server and bypasses permissions

TOOL INTEGRATION

19 OCT

Envoy AI Gateway

Any OIDC token holder can call every MCP tool

IDENTITY AND TOOLS

30 OCT

TAKO AI Agent for Okta

First-run admin takeover, CVSS 9.8

APPLICATION LAYER

The agent risk in one line

Data from tools and memory can carry instructions. Treat it as untrusted, validate it, and log every tool call. October's Claude Agent SDK, Envoy AI Gateway and TAKO agent CVEs all sit on this boundary.

Ten guidelines that need new engineering

Most guidelines extend controls you already run. These ten need new designs, tests or pipeline gates. They take the longest, so start here.

01 AICG 1-1-2

Score agent actions by impact

Classify what each agent can do by impact, likelihood and reversibility. Match safeguards to the score.

02 AICG 1-2-2

Treat tool and memory data as untrusted

Validate anything entering the context window from tools, memory or third-party APIs.

03 AICG 1-2-3

No uncontrolled learning in production

Do not retrain live models unless the risk is assessed and controls are in place.

04 AICG 1-2-5

Review AI-generated code

Security review, testing and approval before AI-written code reaches any environment.

05 AICG 1-2-7

Graduated autonomy

Grow agent independence step by step, with human oversight of security decisions.

06 AICG 2-2-3

Rate limits against extraction

Usage thresholds and rate limits to reduce model extraction and data leakage.

07 AICG 2-4-3

Protect agent memory

Block unauthorised changes to memory and validate stored content against poisoning.

08 AICG 2-4-4

Tested guardrails

Detect malicious input, unauthorised actions and disclosure. Test before go-live and after changes.

09 AICG 2-7-2

Red team and prove the override

Adversarial evaluation and red teaming, plus proof that human override works.

10 AICG 3-1-1

Safe shutdown and rollback

Rollback, safe shutdown and manual alternatives for AI systems during an incident.

New design	New test	New pipeline gate
01, 02, 05, 07	08, 09, 10	03, 04, 06

From guideline to control and evidence

For each subdomain: the controls that satisfy it, the evidence an auditor will ask for, and the related OWASP Top 10 for LLM Applications 2025 risks.

SUBDOMAIN	WHAT TO IMPLEMENT	EVIDENCE | OWASP LLM
D1 Cybersecurity Governance		
1-1 **Risk management**	AI risk register per use case. Agent action catalogue scored by impact, likelihood and reversibility. Threat model per AI system.	Risk register, threat models, action matrix LLM06
1-2 **IT project management**	Security gates in the AI pipeline. Context validation for RAG, tools and memory. AI-generated code scanned and reviewed. Approval for high-impact AI.	Pipeline configs, gate results, approvals LLM01 LLM05
1-3 **Human resources**	Vetting for privileged AI access. Acceptable use and confidentiality covering prompts and outputs.	Signed policies, vetting records LLM02
1-4 **Awareness and training**	Role-based AI security training for builders. Safe prompting awareness for all staff.	Training plan, completion data LLM02
D2 Cybersecurity Defense, part 1		
2-1 **Asset management**	Inventory of models, datasets, vector stores, prompts, agents, MCP servers and AI vendors, with owners.	AI asset inventory, AI-BOM LLM03
2-2 **Identity and access**	Least-privilege roles for agents. Short-lived tokens, vaulted secrets. Rate limits and quotas at the gateway.	Role matrix, access reviews, gateway policy LLM06 LLM10
2-3 **Network security**	Segment AI components. Egress allow-lists for model, tool and retrieval traffic.	Network diagrams, egress rules LLM06

OWASP Top 10 for LLM Applications 2025

LLM01	Prompt injection	LLM02	Sensitive information disclosure
LLM03	Supply chain	LLM04	Data and model poisoning
LLM05	Improper output handling	LLM06	Excessive agency
LLM07	System prompt leakage	LLM08	Vector and embedding weaknesses
LLM09	Misinformation	LLM10	Unbounded consumption

SUBDOMAIN	WHAT TO IMPLEMENT	EVIDENCE | OWASP LLM
D2 Cybersecurity Defense, part 2		
2-4 **Data protection**	Encrypt prompts, embeddings, weights, logs and outputs. Classify data before it enters a model. Memory integrity checks. Input, output and action guardrails.	Encryption settings, guardrail tests LLM02 LLM04 LLM08
2-5 **Backup and recovery**	Versioned backups of models, prompts, configs and vector stores. Tested restore and rollback.	Backup policy, restore logs LLM04
2-6 **Vulnerability management**	Scan model files, libraries and containers. Track AI advisories. Kill switch for vulnerable features.	Scan reports, advisory log LLM03
2-7 **Penetration testing**	AI pen test before go-live and periodically. Red teaming. Human override drills.	Pen test and red team reports LLM01 LLM07
2-8 **Logs and monitoring**	Log prompt config changes, dataset access, tool calls and AI-initiated actions. AI detections in the SOC.	SIEM use cases, retention LLM06
2-9 **Web application security**	TLS, rate limiting, DoS protection, controlled egress and unsafe-channel filters for AI apps and APIs.	WAF and API gateway config LLM10
D3 Cybersecurity Resilience		
3-1 **Business continuity**	Rollback, safe shutdown and manual fallback per AI system. AI scenarios in BCM and DR tests.	Runbooks, DR test results LLM06
D4 Third-Party Cybersecurity		
4-1 **Third party**	Review AI vendors, models and plugins before use and on a schedule. Contract clauses for AI security, exit and disposal.	Vendor assessments, contracts LLM03

20-point readiness checklist

Answer for your most important AI system. Yes = 2, Partly = 1, No = 0. The Excel version that comes with this guide adds up your score for you.

		YES	PART	NO
D1	**Governance**			6 QUESTIONS
01	AI security risks are assessed across the full lifecycle, from design to retirement. 1-1-1	☐	☐	☐
02	Agent actions are classified by impact and reversibility, with safeguards for high-impact actions. 1-1-2	☐	☐	☐
03	Data from tools, memory, RAG and third-party APIs is treated as untrusted and validated. 1-2-2	☐	☐	☐
04	AI-generated code is security reviewed and tested before use. 1-2-5	☐	☐	☐
05	High-impact AI needs a formal security approval before production. 1-2-6	☐	☐	☐
06	Teams that build or operate AI get role-based AI security training. 1-4-1	☐	☐	☐
D2	**Defense**			10 QUESTIONS
07	We keep an inventory of all AI components with owners and lifecycle status. 2-1-1/2	☐	☐	☐
08	AI agents and service accounts have least-privilege, role-based access. 2-2-1	☐	☐	☐
09	Tokens, keys and AI service accounts are reviewed on a fixed schedule. 2-2-2	☐	☐	☐
10	Rate limits and usage thresholds protect against extraction and data leakage. 2-2-3	☐	☐	☐
11	AI components are segmented and egress is limited to approved destinations. 2-3-1	☐	☐	☐
12	Prompts, embeddings, weights, logs and outputs are encrypted and access-restricted. 2-4-1	☐	☐	☐
13	Guardrails are tested before release and after every material change. 2-4-4	☐	☐	☐
14	AI systems are pen tested and red-teamed before go-live and periodically. 2-7-1/2	☐	☐	☐
15	Prompt config changes, tool calls and AI-initiated actions are logged and monitored. 2-8-1/2	☐	☐	☐
16	Vulnerable AI features can be disabled quickly until a fix is validated. 2-6-3	☐	☐	☐
D3	**Resilience**			2 QUESTIONS
17	Each AI system has a defined rollback, safe shutdown and manual fallback. 3-1-1	☐	☐	☐
18	Backup, recovery and rollback for AI systems are tested periodically. 2-5-2	☐	☐	☐
D4	**Third party**			2 QUESTIONS
19	AI vendors, third-party models and plugins are assessed before use and monitored. 4-1-2	☐	☐	☐
20	Contracts set AI security duties, exit, portability and secure disposal terms. 4-1-3/4	☐	☐	☐

Your total _____ **/ 40**

Read your result on page 14

Read your score. Then run 90 days.



0 TO 19

Exposed

Start with inventory, access and guardrails.

20 TO 31

Developing

Close gaps in testing, logging, agents and vendors.

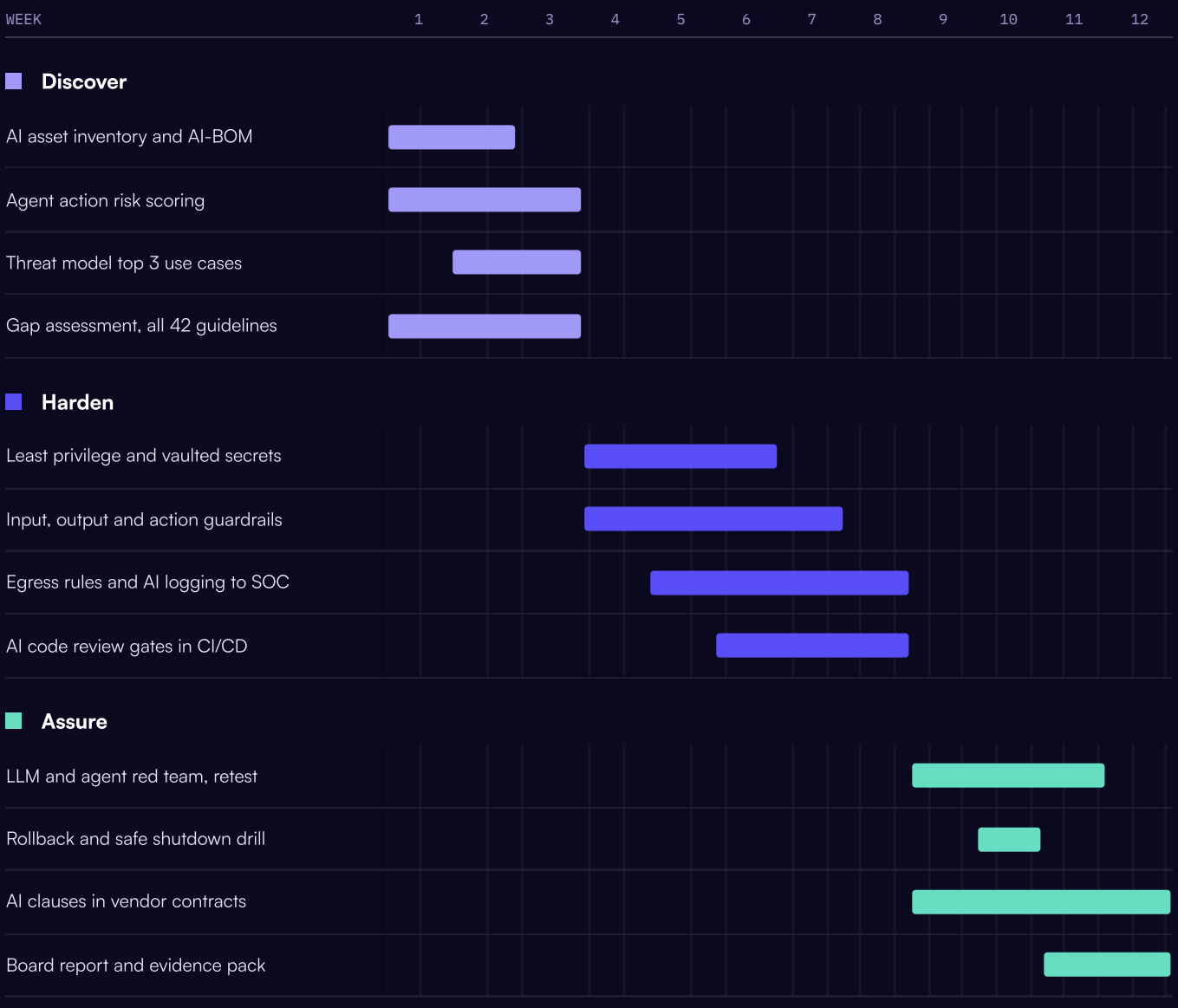
32 TO 40

Advanced

Red team, evidence and continuous assurance.

Bands are a Practical DevSecOps heuristic for prioritising work, not an NCA rating.

A 90-day plan, week by week



Bring your score to a free 30-minute call. We review your answers and your AI use cases, and agree the quick wins for the next 30 days.

BOOK A CALL →

EMAIL US →

Fixed scope. Fixed timeline. Evidence at the end.

The Practical DevSecOps Consulting Team runs hands-on engagements for security and engineering teams. Each one ends with evidence you can show a regulator, a client or your board.

3 TO 4 WEEKS — BEST STARTING POINT

AI Security Readiness Assessment

- ✓ Gap report against all 42 draft guidelines
- ✓ AI asset inventory and risk register
- ✓ Prioritised 90-day remediation plan

2 TO 3 WEEKS

LLM and Agent Red Team

- ✓ Prompt injection, tool abuse and MCP testing
- ✓ Data leakage and model extraction checks
- ✓ Human override validation and retest

6 TO 8 WEEKS

Secure AI SDLC

- ✓ Guardrails and AI code review in CI/CD
- ✓ Agent identity, secrets and egress controls
- ✓ AI detections and logging for the SOC

2 WEEKS PER USE CASE

Agentic AI Threat Modelling

- ✓ Component and data flow mapping
- ✓ Agent action risk scoring
- ✓ Controls mapped to AICG and OWASP

How an engagement runs

1. **Scoping call** — 30 minutes. We agree systems, goals and timeline.
2. **Assess** — Hands-on testing and review against AICG-1:2026.
3. **Fix with you** — Pair with your engineers on the priority gaps.
4. **Evidence pack** — Report, retest results and audit-ready evidence.

Built by the team behind the training

Practical DevSecOps builds hands-on certifications in AI security, MCP security and threat modelling. The consulting team brings the same practical methods to every engagement.



Sources

Every figure in this guide comes from the sources below. Numbers in brackets follow the order they appear.

[1] **NCA.** Draft AI Cybersecurity Guidelines (AICG-1:2026), Istitlaa consultation, 5 July to 5 August 2026.

[2] **IBM / Ponemon.** Cost of a Data Breach Report 2026, Saudi Arabia and Middle East findings, 3 August 2026.

[3] **M. Vu.** "AI and the October 2026 vulnerability storm", DEV Community, 30 October 2026, citing CyberHire, GTIG, Recorded Future, Cloudflare, VulnCheck and CrowdStrike.

[4] **BleepingComputer.** Microsoft October 2026 Patch Tuesday, 13 October 2026; Cisco firewall zero-day, 29 October 2026.

[5] **NVD / Strix.** CVE-2026-63377 and CVE-2026-63378, TAKO AI Agent for Okta, 30 October 2026.

[6] **awesome-ai-agent-attacks.** GitHub timeline of AI agent incidents, October 2026 entries.

[7] **Al Tamimi & Co; Clyde & Co.** Analyses of NCNICC-1:2025 (Oct 2026) and ECC-2:2024 (Jan 2025).

[8] **OWASP.** Top 10 for Large Language Model Applications, 2025.

NOTES

AICG-1:2026 is a draft. Guideline references and counts follow the published draft and may change, and wording is paraphrased. Breach figures are IBM averages. CVE figures are third-party counts. This guide is general guidance, not legal advice.

Practical DevSecOps

NEXT STEP

Secure your AI before it becomes the breach.

Book 30 minutes with the Practical DevSecOps Consulting Team. Bring your checklist score and your top AI use cases. Leave with your gaps and the first steps to close them.

BOOK A 30-MIN CALL →

EMAIL THE TEAM →

Practical DevSecOps Consulting Team