MODULE 7 · STUDY NOTES

# Considering Security, Governance, and Compliance

*Securing generative AI with defense in depth, governing AI initiatives, and meeting compliance obligations on AWS*

Generative AI raises the security stakes: models sit inside critical applications and decision-making, process sensitive data, embody their owner's competitive advantage, and expose entirely new attack surfaces — prompt injection, data poisoning, model theft. This module frames the response as a triad — security (confidentiality, integrity, availability), governance (principles, policies, processes), and compliance (adherence to regulations and standards) — and then tours the AWS toolbox for each: a defense in depth strategy informed by the OWASP Top 10 for LLMs; data and access services (KMS, S3, Macie, IAM, CloudWatch) on purpose-built AI infrastructure; the Generative AI Security Scoping Matrix for judging your share of the work; the AWS CAF governance perspective; and the compliance trio of Amazon Inspector, AWS Artifact, and AWS Audit Manager with its generative AI best-practices framework.

LEARNING OBJECTIVES

- Discuss the security, privacy, and governance considerations for AI systems
- Describe key aspects of a defense in depth approach and how AWS supports this approach
- Identify AWS services that assist with applying governance controls and achieving compliance objectives in AI solutions

## 7.1 Defining security, governance, and compliance for generative AI

Security, governance, and compliance are **interconnected but distinct** functions that every organization deploying AI must address. They serve different goals, yet they are interdependent — good governance helps produce a secure and compliant system, and compliance is impossible if security has gaps. Together they form a comprehensive framework for secure, ethical, and compliant AI, and organizations must address **all three** to mitigate risks, build trust, and unlock the full potential of generative AI.

**TABLE 7.1** Three keys for mitigating risk and building trust

| FUNCTION | WHAT IT ENSURES | WHAT IT INVOLVES |
|---|---|---|
| **Security** | The **confidentiality, integrity, and availability** of data, models, and infrastructure | Protecting systems and data against threats and adversarial attacks |
| **Governance** | Principles, policies, and processes for managing AI systems **throughout their lifecycle** | Ethical considerations, risk management, and decision-making processes |
| **Compliance** | Adherence to relevant **regulations, industry standards, and internal policies** | Mechanisms for auditing, monitoring, and demonstrating responsible AI practices |

**Why security weighs more here.** AWS considers security *job zero* — nothing is more critical to cloud-based systems. Generative AI heightens the risk-management burden for four reasons: gaining business value means integrating AI models into **critical applications and decision-making processes**; AI models often **process sensitive data**; the data models use can provide a **competitive advantage** to its owners; and AI systems present **new fronts for adversarial attacks**.

KEY GOVERNANCE FOCUS FOR GENERATIVE AI

- **Manage, optimize, and scale** the organizational AI initiative — the core of the governance perspective, instrumental for building trust and deploying AI at scale.
- **Drive consistency and manage risk** — align AI use with organizational goals and ensure ethical, effectively managed technology.
- **Implement frameworks** governing ethics, data quality and usage, and regulatory compliance.
- **Manage the cost patterns** of AI workloads — scalable processes and standards let initiatives expand across business units for long-term business value.

KEY COMPLIANCE FOCUS AREAS FOR GENERATIVE AI

- **Ethical practices** — fairness, transparency, and privacy protection across the AI lifecycle: mitigate bias, provide explainability, safeguard user data.
- **High technical standards** — maintain data quality and integrity, define and measure model performance metrics, implement robust security against vulnerabilities and attacks.
- **Regulatory adherence** — data-protection law such as **GDPR** (EU data and privacy law) plus emerging AI-specific rules: the **EU AI Act**, the **NIST AI Risk Management Framework**, and sector-specific regulations in finance, healthcare, and other industries.
- **Risk processes** — risk assessment and management, auditing and monitoring mechanisms, and continuous improvement for responsible, accountable AI development and deployment.

COMMON PITFALL

The three functions are not interchangeable on the exam. *Security* wording centers on **confidentiality, integrity, availability**; *governance* on **principles, policies, processes, lifecycle**; *compliance* on **regulations, standards, internal policies**. Watch which vocabulary the question uses.

## 7.2 Security of AI systems on AWS

The **Open Worldwide Application Security Project (OWASP)** is a nonprofit foundation working to improve software security; its Top 10 is the standard awareness document for web-application risk. OWASP has published a comparable **Top 10 for LLM applications and generative AI**, grouping the vulnerabilities by the component of an LLM application they most affect. Understanding and mitigating them is critical to secure, responsible deployment — and a defense in depth strategy is how you address them.

**TABLE 7.2** OWASP Top 10 vulnerabilities for LLM applications

| # | VULNERABILITY | WHAT GOES WRONG |
|---|---|---|
| **1** | Prompt injection | Malicious user inputs manipulate the behavior of the language model through crafted prompts |
| **2** | Insecure output handling | Insufficient validation, sanitization, and handling of LLM outputs before they pass downstream to other components and systems |
| **3** | Training data poisoning | Manipulation of pre-training, fine-tuning, or embedding data to introduce vulnerabilities, backdoors, or biases that compromise security, effectiveness, or ethics |
| **4** | Model denial of service | An attacker's interactions consume exceptionally high resources — quality of service declines for everyone, and resource costs spike |
| **5** | Supply chain vulnerabilities | Weaknesses in the software, hardware, or services used to build or deploy a model; third-party pre-trained models and training data are susceptible to tampering and poisoning |
| **6** | Sensitive information disclosure | Outputs reveal sensitive information, proprietary algorithms, or confidential details — unauthorized data access, IP loss, privacy violations |
| **7** | Insecure plugin design | Model-driven plugins accept free-text input without validation or type checking, enabling malicious requests up to remote code execution |
| **8** | Excessive agency | A model granted too much autonomy or capability performs damaging actions in response to unexpected or ambiguous outputs |
| **9** | Overreliance | Over-dependence on the model: erroneous output delivered authoritatively (hallucination) is trusted without oversight — breaches, misinformation, legal and reputational damage |
| **10** | Model theft | Unauthorized access and exfiltration of proprietary models — weights and parameters copied or extracted to create a functional equivalent |

**Defense in depth** is the strategy for countering these risks: implement multiple **redundant defenses** across every layer of the AI/ML environment, so that if one control fails, additional layers exist to isolate threats and *prevent, detect, respond, and recover* from security events. **Secure everything** — apply the strategy to generative AI workloads, data, and information, and protect the AWS accounts and assets behind the solution. Finally, **combine options to optimize**: evaluate a mix of services and solutions at each layer to improve security and resiliency.

**TABLE 7.3** Securing every layer

| LAYER | WHAT IT INVOLVES |
|---|---|
| **Data protection** | Protect data **at rest, in transit, and in use** |
| **Identity & access management** | Only authorized users, applications, or services can access and interact with the cloud infrastructure and its services |
| **Application protection** | Secure coding, input validation, access controls, vulnerability management — prevents application-level attacks such as SQL injection and cross-site scripting (XSS) |
| **Network & edge protection** | Firewalls, web application firewalls (WAFs), intrusion detection/prevention (IDS/IPS), secure network configurations — mitigates unauthorized access, DDoS, network exploits |
| **Infrastructure protection** | Guards against unauthorized access, data breaches, system failures, and natural disasters |
| **Threat detection & incident response** | Identify and address potential security threats or incidents |
| **Compliance & governance** | Policies, procedures, and mechanisms that ensure adherence to regulations and standards |

AWS has implemented robust data governance, encryption protocols, and access controls for **more than 15 years**; its core infrastructure is built to satisfy the security requirements of the **military, global banks, and other high-sensitivity organizations**. Under the **shared responsibility model** you focus on securing your applications and data while AWS secures the infrastructure — and with AWS services you *inherit* comprehensive compliance controls incorporating global and regional standards and certifications. This attention to security extends to the AWS AI infrastructure and generative AI services: managed services such as **Amazon Q, Amazon Bedrock, and Amazon SageMaker AI** carry these protections built in.

| AWS – SECURITY OF THE CLOUD | CUSTOMER – SECURITY IN THE CLOUD |
|---|---|
| *the underlying cloud infrastructure* | *workloads, data, and applications on top* |
| ▪ Hardware, software, networking, and facilities that run AWS Cloud services<br>▪ Purpose-built AI infrastructure — secure by design<br>▪ Security machinery integrated into managed generative AI services | ▪ Configuring security groups and implementing access controls<br>▪ Encrypting data; maintaining software updates and patches<br>▪ Exact share varies with the AWS services chosen and how they are integrated |

Mapped onto the defense in depth layers, the key services for securing a generative AI solution are: **Amazon CloudWatch** for threat detection and incident response, **AWS purpose-built AI infrastructure** for infrastructure protection, **IAM** for identity and access management, and **AWS KMS, Amazon S3, and Amazon Macie** for data protection. (The module focuses on data at rest; protecting data in transit matters equally — the Well-Architected Framework security pillar covers it.)

**TABLE 7.4** Key AWS services for securing generative AI

| SERVICE | ROLE FOR GENERATIVE AI | DETAILS WORTH KEEPING |
|---|---|---|
| **AWS KMS** | Secure, centralized key management — all data including model artifacts and storage volumes encrypted | Integrates with **Amazon Bedrock and SageMaker AI**; automatic and manual key rotation aids regulatory compliance |
| **Amazon S3** | Secure object storage for AI workflows; robust encryption; fine-grained IAM access policies (least privilege) | Encrypted **by default** with server-side encryption (AWS managed or customer managed keys); stores model invocation logs, common prompts, artifacts |
| **Amazon Macie** | Fully managed data security/privacy service — ML + pattern matching to automatically discover and help protect **sensitive data in S3** | Inspects bucket inventory and all objects; single dashboard; integrates with EventBridge and Security Hub for centralized remediation |
| **IAM** | Fine-grained roles with least-privilege permissions; role-based control across dev/test/prod environments | Model-specific permissions — control who can use which **Bedrock** foundation models for inference or customization |
| **CloudWatch** | Real-time monitoring for unusual patterns or breaches; customizable alerts for risks such as **prompt injection** or sensitive-information disclosure | Encrypted logging with customer managed keys; **Observability Access Manager** centralizes monitoring across accounts |

### PURPOSE-BUILT AI INFRASTRUCTURE

- AI infrastructure is **secure by design** — nobody at AWS can access the workloads or data running on it.
- Customers can completely isolate their AI data, **even from their own users and applications**, using integrated encryption and access services.
- Data is **encrypted by default**; the same granular encryption and access controls used throughout the AWS Cloud protect data and models.
- Regardless of the service used, **no customer data is used to train Amazon or third-party foundation models**.

### THREE APPS, THREE SECURITY POSTURES

How much security you owe depends on the use case and the data. A **movie finder app** built on a public service such as PartyRock (an Amazon Bedrock playground): you secure your application, but no company data is at risk and you are not responsible for the security of the service itself — low data-security concerns, narrow breadth. An **enterprise customer-service app** raises both. A **healthcare app using your own proprietary model** built on patient data and internal business data sits at the top: you must protect highly sensitive patient data *and* internal data, and secure **every element of the ML lifecycle** — the training environment, the model itself, and the production deployment.

The **Generative AI Security Scoping Matrix** turns that intuition into a framework for assessing and implementing security controls across the AI lifecycle. For each of five scopes of generative AI use it sets out considerations across five security disciplines: **governance and compliance** (policies, procedures, reporting that empower the business while minimizing risk), **legal and privacy** (regulatory and privacy requirements), **risk management** (threat identification and mitigations), **controls** (implementing the security controls that mitigate risk), and **resilience** (architecting for availability and business SLAs).

**TABLE 7.5** The five scopes of generative AI use

| SCOPE | WHAT YOU ARE DOING | EXAMPLE |
|---|---|---|
| 1 | Using **public** generative AI services | PartyRock; gen-AI apps on your phone or the web |
| 2 | Using a **third-party** service for an enterprise application | SaaS products for sales management, scheduling, or customer support |
| 3 | Building an application on a **versioned model** | App calling a pre-trained foundation model |
| 4 | **Customizing** a model with your data | Fine-tuning on company data |
| 5 | **Training** a model from scratch with your data | Fully proprietary model |

COMMON PITFALL

The matrix does not say lower scopes are risk-free — Scope 1 guidance applies even to generative AI apps you use on your phone. What changes is **how much of the stack you are responsible for**: by Scope 5 you own security for the training environment, the model, and the deployment.

## 7.3 Governing AI systems on AWS

**Cloud governance** is a set of rules, processes, and reports that aligns your AWS Cloud use with business objectives. It covers **security** (multi-account strategies, continuous monitoring, control policies), **compliance** (automated checks, reporting, remediation), **operations** (enterprise-wide controls), **identity** (centralized identity and access management at scale), **cost** (usage reports and policy enforcement), and **resilience** (assessments and testing integrated into development and deployment pipelines). Building an AI governance practice starts by identifying all key stakeholders and assembling a team representing **multiple business units**, closely aligned with the organization's AI strategy.

KEY GOVERNANCE ACTIVITIES

- **Define governance goals**, including ethical goals.
- **Identify areas of potential risk.**
- **Develop policies and guidelines** for responsible AI and compliance.
- **Define monitoring mechanisms** for AI systems, and determine actions based on predefined thresholds.
- **Continuously revise** results and existing policies — governance is an ongoing, active process, balancing innovation and control.

The **AWS Cloud Adoption Framework for AI, ML, and generative AI** extends the AWS CAF to the specific concerns of AI systems, identifying foundational AI capabilities across six perspectives: **business, people, governance, platform, security, and operations**. Two companion tools help you measure yourself: the **AWS Well-Architected Tool** provides a consistent process for measuring your architecture against AWS best practices (including a machine learning lens for ML workloads), and **AWS Trusted Advisor** inspects your environment and recommends where you can save money, improve system availability and performance, or close security gaps.

**TABLE 7.6** CAF governance perspective – four focus areas

| FOCUS AREA | GOAL |
|---|---|
| **Cloud financial management** | Plan, measure, and optimize the cost of AI in the cloud |
| **Data curation** | Create value from data catalogs and products |
| **Risk management** | Mitigate and manage the risks inherent to AI |
| **Responsible AI** | Foster continuous AI innovation through responsible AI practices (covered in depth in the previous module) |

### CLOUD FINANCIAL MANAGEMENT IN PRACTICE

- AI cost follows a **zig-zag trajectory** — high or low data-preparation costs up front, a volatile proof-of-concept phase (especially with model training or retraining), then generally low inference costs post-deployment.
- Use **purpose-built AI hardware** — Amazon EC2 **Trn1** and **Inf2** instances — and AI services during resource-intensive phases; **AWS Inferentia** architecture further optimizes inference cost.
- **Starting from foundation models in Amazon Bedrock or SageMaker AI reduces development cost** versus building from scratch.
- Prioritize cost visibility: **tag resources**, track and analyze data, training, and inference costs over time; align AI projects with business goals and calculate cost-versus-performance trade-offs for a positive ROI.
- Budget for responsible AI early — an often-underestimated cost that reduces risk and expense downstream.

### DATA CURATION IN PRACTICE

- Treat data as a **valuable asset and a product** — democratize it and make it discoverable across the organization; acquiring, labeling, cleaning, and processing data speeds time-to-value and boosts model performance.
- Assign **direct owners or data stewards** to datasets; implement easy-to-use, human-readable **repositories, catalogs, and dictionaries** so teams of all skill levels can discover, understand, and collaborate on data.
- Apply **data quality assessments and governance rules** carefully — done well they accelerate data use; done poorly they hinder innovation.
- Expand data resources deliberately: buy external data, create **synthetic data** with ML algorithms, crowdsource labeling of internal data, or automate data generation and capture — with governance for when to use each.
- These recommendations align to the **AWS modern data architecture**, which leans on services such as Amazon S3 for data management and access.

**Risk management.** The non-deterministic nature of AI models makes specific outcomes hard to guarantee, creating financial risks and potential sunken costs not seen in traditional software development. Mitigations include **model cards** — technical artifacts documenting a model so developers can design around its strengths and inform users of its weaknesses — and lightweight builds: **proofs of concept (POCs), minimum viable products (MVPs), and minimum loveable products (MLPs)** that let stakeholders try solutions without building a fully featured application. Beyond finances, AI poses **legal and ethical risks** — including risks classified by local legislatures such as the European Union, and those inherent to AI's open-ended, non-deterministic outputs. Develop **safeguards and architectures that constrain AI systems** to keep subsystem failures from propagating downstream, manage risk at the level of whole processes, and plan for long-term challenges: **data and concept drift** (the model and its training data no longer matching reality, causing poor predictions) and security against bad actors.

| AWS CONFIG | AWS CLOUDTRAIL |
|---|---|
| *what resources are, and how they are set up*<br>▪ Continuously monitors and records resource configurations and their changes over time<br>▪ Detailed inventory of resources, current configuration, and relationships between resources<br>▪ Rules and evaluation mechanisms support compliance auditing, security analysis, change tracking, and troubleshooting | *events that occur in the account*<br>▪ Detailed audit trail of **all API calls** made in the account<br>▪ Tracks create, read, update, delete (CRUD) actions across services, including Amazon Bedrock and Amazon S3<br>▪ Integrates with CloudWatch to monitor and raise alarms — detect threats and anomalies in real time |

**COMMON PITFALL**

Certification tip straight from the deck: remember the **type of data** each service deals with. **AWS Config** is about *what and how resources are set up*; **CloudTrail** is about *events that occur* in an AWS account. Questions are engineered to make you swap them.

## 7.4 Compliance for AI systems on AWS

AWS takes a comprehensive approach to compliance so that organizations can meet regulatory obligations while using AWS services. Compliance matters most for **regulated workloads** — industries with high regulatory or industrial demands such as **financial services, healthcare, and aerospace** — where failure to comply could create a major health or safety risk.

**TABLE 7.7** How AWS supports compliance

| COMPONENT | WHAT IT MEANS |
|---|---|
| **Shared responsibility model** | AWS is responsible for compliance **of** the cloud; customers are responsible for compliance of their **use** of AWS services |
| **Compliance by design** | Compliance considerations are built into AWS services — built-in controls and features that help meet regulatory requirements |
| **Compliance frameworks & tools** | **AWS Audit Manager, AWS Artifact, AWS Config** — monitor compliance posture, automate reporting, verify resource configuration |
| **Certifications & reports** | AWS maintains certifications across standards such as **ISO 27001, SOC 1/2/3, and PCI DSS**; AWS Artifact provides on-demand access |
| **Compliance assistance** | Partnerships with third-party auditors and consulting firms — understand obligations, implement controls, prepare for audits |
| **Training & resources** | Documentation, webinars, and hands-on training on compliance best practices |

| AMAZON INSPECTOR | AWS ARTIFACT |
|---|---|
| *assesses your resources*<br>▪ Automated security assessments of AWS resources and applications — EC2 instances checked against predefined rules packages<br>▪ Identifies vulnerabilities, exposures, and deviations from best practices<br>▪ Detailed findings **prioritized by severity**; continuous monitoring with updates from AWS security researchers<br>▪ Integrates with development and deployment pipelines — security built in from the ground up, not an afterthought | *documents AWS's compliance*<br>▪ On-demand downloads of AWS security and compliance documents (ISO, SOC, PCI DSS)<br>▪ Certifications from accreditation bodies validating AWS security controls<br>▪ Security and compliance documents for **independent software vendors (ISVs)** selling on AWS Marketplace; tracks agreements across accounts<br>▪ Documents can be submitted to auditors or regulators as **audit artifacts** |

**AWS Audit Manager** simplifies and automates compliance work: it automates the **collection of evidence** for compliance assessments and provides **pre-built frameworks** — collections of controls with descriptions and testing procedures, grouped to the requirements of a given standard or regulation — which you can also customize for internal audits. Its **framework for generative AI** (the *AWS Generative AI Best Practices Framework v2*) gives you visibility into how a generative AI implementation on **Amazon Bedrock or Amazon SageMaker AI** performs against AWS-recommended best practices. Its controls were developed with AI experts, compliance practitioners, and security assurance specialists across AWS, with input from **Deloitte**; each automated control maps to an AWS data source from which Audit Manager collects evidence. The module's demo walks through setting up an assessment with this framework.

**TABLE 7.8** The eight principles of the Audit Manager generative AI framework

| PRINCIPLE | WHAT IT ASKS OF YOUR IMPLEMENTATION |
|---|---|
| **Responsible** | Develop and adhere to ethical guidelines for deploying and using generative AI models |
| **Safe** | Establish clear parameters and ethical boundaries to prevent harmful or problematic output |
| **Fair** | Consider and respect how the AI system impacts different sub-populations of users |
| **Sustainable** | Strive for greater efficiency and more sustainable power sources |
| **Resilience** | Maintain integrity and availability mechanisms so the system operates reliably |
| **Privacy** | Protect sensitive data from theft and exposure |
| **Accuracy** | Build AI systems that are accurate, reliable, and robust |
| **Secure** | Prevent unauthorized access to generative AI systems |

**SAMPLE EXAM WALKTHROUGH**

*A company is implementing an AI solution on AWS and needs to* ***track changes to their resources*** *for compliance purposes. Which service is MOST suitable?* Key phrases: *AI solution* and *track changes to resources for compliance*. Amazon Inspector assesses security but does not track resource changes over time. AWS Artifact serves AWS compliance reports — it doesn't watch your resources. CloudWatch collects and tracks **metrics**, not detailed configuration history. **AWS Config** is the answer: it provides a detailed view of resource configurations in the account, how resources relate, and **how they were configured in the past**.

**COMMON PITFALL**

Artifact proves **AWS's** compliance to your auditor; it says nothing about your workload. Evidence about *your* implementation comes from tools like **Audit Manager** (automated evidence against frameworks) and **Config** (configuration compliance).

## Module summary

- Security ensures the confidentiality, integrity, and availability of data, models, and infrastructure; governance establishes principles, policies, and processes for the AI lifecycle; compliance ensures adherence to regulations, standards, and internal policies — interdependent, and all three are required to mitigate risk and build trust.
- Generative AI heightens security risk: models are integrated into critical applications and decisions, process sensitive data, embody competitive advantage, and open new fronts for adversarial attack — cataloged by the OWASP Top 10 for LLMs (prompt injection, insecure output handling, data poisoning, model DoS, supply chain, sensitive information disclosure, insecure plugins, excessive agency, overreliance, model theft).
- Defense in depth = multiple redundant controls at every layer: data protection, identity and access management, application protection, network and edge protection, infrastructure protection, threat detection and incident response, compliance and governance.
- Shared responsibility: AWS secures (and is compliant for) the cloud; you secure and comply for what you run in it. AWS AI infrastructure is secure by design, data is encrypted by default, and customer data is never used to train Amazon or third-party foundation models.
- Key security services: AWS KMS (centralized keys, Bedrock/SageMaker AI integration, rotation), Amazon S3 (default encryption, fine-grained access), Amazon Macie (sensitive-data discovery in S3), IAM (least-privilege roles, dev/test/prod separation, per-model Bedrock permissions), CloudWatch (monitoring, custom alerts, encrypted logs).
- The Generative AI Security Scoping Matrix maps five scopes — public services, third-party enterprise SaaS, building on a versioned model, customizing with your data, training from scratch — against five disciplines: governance & compliance, legal & privacy, risk management, controls, resilience. Responsibility grows with scope.
- Governance is an ongoing loop — define goals (including ethical goals), identify risks, develop policies, monitor against thresholds, revise. The CAF for AI/ML/generative AI governance perspective covers cloud financial management, data curation, risk management, and responsible AI; AWS Config records resource configurations while CloudTrail records API calls.
- Compliance support spans shared responsibility, compliance by design, frameworks and tools, certifications and reports (ISO 27001, SOC 1/2/3, PCI DSS), assistance, and training. Amazon Inspector assesses, AWS Artifact

> documents, and AWS Audit Manager automates evidence collection — including a generative AI framework built on eight principles.

## Review questions

**1. Distinguish security, governance, and compliance for AI systems — and explain why an organization must address all three.**

*Answer:* Security ensures the confidentiality, integrity, and availability of data, models, and infrastructure; governance establishes the principles, policies, and processes for managing AI across its lifecycle (ethics, risk, decision-making); compliance ensures adherence to regulations, industry standards, and internal policies via auditing and monitoring. They are interdependent — good governance produces secure, compliant systems, and gaps in any one undermine the others.

**2. Why does generative AI demand more security risk management than a typical cloud application?**

*Answer:* Four reasons: models are integrated into critical applications and decision-making; they often process sensitive data; their data can be a competitive advantage worth stealing; and they present new fronts for adversarial attacks such as prompt injection and data poisoning.

**3. Name and define four vulnerabilities from the OWASP Top 10 for LLMs.**

*Answer:* Examples: prompt injection — malicious input manipulates model behavior; training data poisoning — tampered training data plants backdoors or bias; excessive agency — too much autonomy enables damaging actions; model theft — weights/parameters exfiltrated to clone the model. (Also: insecure output handling, model DoS, supply chain, sensitive information disclosure, insecure plugin design, overreliance.)

**4. What are the three key practices of a defense in depth approach?**

*Answer:* Use multiple redundant defenses (if one control fails, other layers prevent, detect, respond, recover); secure everything — workloads, data, information, plus the AWS accounts and assets behind them; and combine services and solutions, evaluating at each layer to improve security and resiliency.

**5. A hospital builds a diagnosis assistant on a model trained from scratch on patient data. Which scope of the Security Scoping Matrix applies, and what must the hospital secure?**

*Answer:* Scope 5 — training a model from scratch with your data. It must secure the internal business data, the highly sensitive patient data, and every element of the ML lifecycle: the training environment, the model itself, and the production deployment.

**6. Which service answers "who called the Bedrock API yesterday, and what did they do?" — and which answers "how is this resource configured, and how has that changed?"**

*Answer:* CloudTrail answers the first: an audit trail of all API calls (CRUD actions) including Amazon Bedrock and S3. AWS Config answers the second: it continuously records resource configurations and their changes over time.

**7. List the four focus areas of the CAF for AI/ML/generative AI governance perspective, with one practice for each.**

*Answer:* Cloud financial management — tag and track costs, use purpose-built hardware (Trn1/Inf2), start from FMs in Bedrock/SageMaker AI. Data curation — data stewards, human-readable catalogs, synthetic or purchased data with governance. Risk management — model cards, POCs/MVPs/MLPs, safeguards against failure propagation, plan for drift. Responsible AI — foster innovation through responsible practices.

**8. How does the AWS Audit Manager generative AI framework work, and what are its eight principles?**

*Answer:* Audit Manager automates evidence collection; the framework provides pre-built controls (developed with AWS experts and input from Deloitte), each mapped to an AWS data source, to assess a Bedrock or SageMaker AI implementation against best practices. The principles: responsible, safe, fair, sustainable, resilience, privacy, accuracy, secure.