

# Reliable CSPAI Test Topics & CSPAI Valid Learning Materials



P.S. Free 2026 SISA CSPAI dumps are available on Google Drive shared by Pass4training: [https://drive.google.com/open?id=1cA\\_kacIqG7kz4McTqTakyq53y1yiuTT](https://drive.google.com/open?id=1cA_kacIqG7kz4McTqTakyq53y1yiuTT)

We are going to promise that we will have a lasting and sustainable cooperation with customers who want to buy the CSPAI study materials from our company. We can make sure that our experts and professors will try their best to update the study materials in order to help our customers to gain the newest and most important information about the CSPAI Exam. If you decide to buy our study materials, you will never miss any important information. In addition, we can promise the updating system is free for you.

## SISA CSPAI Exam Syllabus Topics:

Topic	Details
Topic 1	<ul style="list-style-type: none"> <li>Evolution of Gen AI and Its Impact: This section of the exam measures skills of the AI Security Analyst and covers how generative AI has evolved over time and the implications of this evolution for cybersecurity. It focuses on understanding the broader impact of Gen AI technologies on security operations, threat landscapes, and risk management strategies.</li> </ul>

Topic 2	<ul style="list-style-type: none"> <li>Securing AI Models and Data: This section of the exam measures skills of the Cybersecurity Risk Manager and focuses on the protection of AI models and the data they consume or generate. Topics include adversarial attacks, data poisoning, model theft, and encryption techniques that help secure the AI lifecycle.</li> </ul>
Topic 3	<ul style="list-style-type: none"> <li>Using Gen AI for Improving the Security Posture: This section of the exam measures skills of the Cybersecurity Risk Manager and focuses on how Gen AI tools can strengthen an organization's overall security posture. It includes insights on how automation, predictive analysis, and intelligent threat detection can be used to enhance cyber resilience and operational defense.</li> </ul>

>> **Reliable CSPAI Test Topics** <<

## CSPAI Valid Learning Materials, CSPAI Exam

CSPAI valid study test give you an in-depth understanding of the contents and help you to make out a detail study plan for CSPAI preparation. All the questions are edited according to the analysis of data and summarized from the previous test, which can ensure the high hit rate. You just need take the spare time to study CSPAI Training Material, the effects are obvious. You will get a high score with the help of SISA CSPAI study pdf.

## SISA Certified Security Professional in Artificial Intelligence Sample Questions (Q20-Q25):

### NEW QUESTION # 20

Fine-tuning an LLM on a single task involves adjusting model parameters to specialize in a particular domain. What is the primary challenge associated with fine tuning for a single task compared to multi task fine tuning?

- A. Single-task fine-tuning is less effective in generalizing to new, unseen tasks compared to multi-task fine-tuning.
- B. Single-task fine-tuning requires significantly more data to achieve comparable performance to multi-task fine tuning.
- C. Single-task fine-tuning introduces more complexity in managing different versions of the model compared to multi-task fine-tuning.
- D. Single-task fine-tuning tends to degrade the model's performance on the original tasks it was trained on.

**Answer: A**

Explanation:

Single-task fine-tuning specializes the LLM but risks overfitting, limiting generalization to novel tasks unlike multi-task approaches that promote transfer learning across domains. This challenge requires careful regularization in SDLC to balance specificity and versatility, often needing more resources for version management. Exact extract: "Single-task fine-tuning is less effective in generalizing to new tasks compared to multi-task fine-tuning." (Reference: Cyber Security for AI by SISA Study Guide, Section on Fine-Tuning Challenges, Page 115-118).

### NEW QUESTION # 21

In a scenario where Open-Source LLMs are being used to create a virtual assistant, what would be the most effective way to ensure the assistant is continuously improving its interactions without constant retraining?

- A. Reducing the amount of feedback integrated to speed up deployment.
- B. Implementing reinforcement learning from human feedback (RLHF) to refine responses based on user input.
- C. Training a larger proprietary model to replace the open-source LLM
- D. Shifting the assistant to a completely rule-based system to avoid reliance on user feedback.

**Answer: B**

Explanation:

For continuous improvement in open-source LLM-based virtual assistants, RLHF integrates human evaluations to align model outputs with preferences, iteratively refining behavior without full retraining. This method uses reward models trained on feedback to guide policy optimization, enhancing interaction quality over time. It addresses limitations like initial biases or suboptimal responses by leveraging real-world user inputs, making the system adaptive and efficient. Unlike full retraining, RLHF is parameter-efficient and

scalable, ideal for production environments. Security benefits include monitoring feedback for adversarial attempts. Exact extract: "Implementing RLHF allows continuous refinement of the assistant's interactions based on user feedback, avoiding the need for constant full retraining while improving performance." (Reference: Cyber Security for AI by SISA Study Guide, Section on AI Improvement Techniques in SDLC, Page 85-88).

#### NEW QUESTION # 22

In ISO 42001, what is required for AI risk treatment?

- A. Delegating all risk management to external auditors.
- B. Ignoring risks below a certain threshold.
- C. Identifying, analyzing, and evaluating AI-specific risks with treatment plans.
- D. Focusing only on post-deployment risks.

**Answer: C**

Explanation:

ISO 42001 mandates a systematic risk treatment process, involving identification of AI risks (e.g., bias, security), analysis of impacts, evaluation against criteria, and development of treatment plans like mitigation or acceptance. This ensures proactive management throughout the AI lifecycle. Exact extract: "ISO 42001 requires identifying, analyzing, and evaluating AI risks with appropriate treatment plans." (Reference: Cyber Security for AI by SISA Study Guide, Section on Risk Treatment in ISO 42001, Page 270-273).

#### NEW QUESTION # 23

What is a potential risk of LLM plugin compromise?

- A. Reduced model training time
- B. Better integration with third-party tools
- C. Unauthorized access to sensitive information through compromised plugins
- D. Improved model accuracy

**Answer: C**

Explanation:

LLM plugin compromises occur when extensions or integrations, like API-connected tools in systems such as ChatGPT plugins, are exploited, leading to unauthorized data access or injection attacks. Attackers might hijack plugins to leak user queries, training data, or system prompts, breaching privacy and enabling further escalations like lateral movement in networks. This risk is amplified in open ecosystems where plugins handle sensitive operations, necessitating vetting, sandboxing, and encryption. Unlike benefits like accuracy gains, compromises erode trust and invite regulatory penalties. Mitigation strategies include regular vulnerability scans, least-privilege access, and monitoring for anomalous plugin behavior. In AI security, this highlights the need for robust plugin architectures to prevent cascade failures. Exact extract: "A potential risk of LLM plugin compromise is unauthorized access to sensitive information, which can lead to data breaches and privacy violations." (Reference: Cyber Security for AI by SISA Study Guide, Section on Plugin Security in LLMs, Page 155-158).

#### NEW QUESTION # 24

How does ISO 27563 support privacy in AI systems?

- A. By mandating the use of specific encryption algorithms.
- B. By limiting AI to non-personal data only.
- C. By focusing on performance metrics over privacy.
- D. By providing guidelines for privacy-enhancing technologies in AI.

**Answer: D**

Explanation:

ISO 27563 offers practical guidance on implementing privacy-enhancing technologies (PETs) in AI, such as differential privacy or federated learning, to protect data while maintaining utility. It addresses risks like inference attacks, ensuring compliance with privacy regulations. Exact extract: "ISO 27563 supports privacy in AI by providing guidelines for privacy-enhancing technologies." (Reference: Cyber Security for AI by SISA Study Guide, Section on ISO 27563 for Privacy, Page 265-268).

