

PRACTICE QUESTIONS

GIAC / SANS

GAIPS

GIAC AI Platform Security

SHARED BY [u/gentleyak1995](#) on aiexamcerts

EDITION Free Set 1

QUESTIONS 10

SHARED October 18, 2026

EXAM GIAC / SANS GAIPS, GIAC AI Platform Security

Shared by [u/gentleyak1995](#) on aiexamcerts

10 practice questions for the GIAC / SANS GAIPS exam, GIAC AI Platform Security, each followed by its answer and a short explanation. More free sets and the full practice exam: aiexamcerts.com/giac-sans/gaips. Spotted a mistake? Email support@aiexamcerts.com with the exam code GAIPS and the question number.

1. A company runs an inference service on Kubernetes. Which setting most directly limits what a compromised model container can reach on the cluster network?
 - A Enabling verbose application logs for every request handled by the pod
 - B Increasing the number of pod replicas behind the service for higher availability
 - C A NetworkPolicy that restricts egress to only the services the model needs
 - D Setting a higher CPU request so the pod is scheduled on larger nodes in the cluster

Answer: C

A NetworkPolicy that limits egress is the control that restricts lateral movement from a compromised container to the services it does not need. Logging helps detection but does not restrict reach. Replicas and CPU requests affect availability and scheduling, not what a container can connect to.

2. The learned numerical values inside a neural network that are adjusted during training are called _____.

Answer: weights (also accepted: parameters)

Training updates the weights, which together with biases are the parameters that define the model's behavior. They are stored in the model artifact and are what gets signed, versioned or stolen. Weights and parameters are used interchangeably in this context.

3. Which NIST AI Risk Management Framework function is concerned with cultivating a risk-aware culture, with roles, accountability and policies across the organization?
 - A Map, which establishes the context in which an AI system operates and the risks it presents
 - B Measure, which assesses identified risks with metrics, testing and monitoring over time
 - C Manage, which decides how to respond to and prioritize risks once they are assessed
 - D Govern, which sets policies, accountability and a risk-aware culture

Answer: D

Govern is the cross-cutting function that sets policies, roles and accountability, and it applies throughout the other three. Map, Measure and Manage each handle a more specific step in the risk lifecycle. Map is the tempting distractor because it also involves context, but it addresses the system's context rather than organizational culture.

4. Which two design choices make an agent's tool use safer? (Choose two.)
- A Scoping each tool to the minimum permissions its task requires
 - B Letting the agent generate new tools at runtime and run them immediately without review
 - C Passing the full conversation history to every tool on every call
 - D Sharing one API token across all tools so that configuration stays simple
 - E Requiring explicit human approval before irreversible actions such as payments or deletions

Answer: A, E

Least-privilege tool scopes and approval gates limit what a manipulated agent can do and make consequential steps visible to a person. Runtime tool generation, full history sharing and one shared token each widen the blast radius. The shared token option is the tempting distractor because simple configuration sounds like good engineering.

5. Which artifact best supports tracing a training dataset's origin, license and transformation history?
- A Lineage records of each dataset's source, license and transformations
 - B A larger count of GPU nodes assigned to the training cluster for each run
 - C A faster network switch installed between the storage and compute tiers
 - D A shorter model name chosen to make the release easier to mention in meetings

Answer: A

Lineage records connect each dataset to its source and the processing steps that produced the training set, which is what audits and incident response need. Hardware and naming choices have no bearing on where the data came from. A faster network switch is the distractor a performance-focused engineer might choose first.

6. Direct preference optimization fine-tunes a policy directly on preference pairs, without training a separate reward model first.
- A True
 - B False

Answer: A (True)

DPO reformulates preference learning so the policy is optimized from chosen and rejected response pairs in a single stage, removing the separate reward model used in classic RLHF. The statement is true.

7. Which property of a vector index most affects how quickly approximate nearest-neighbor search returns results on a large corpus?
- A The file extension of the source document that the chunk was taken from originally
 - B Whether the index uses an approximate method such as HNSW
 - C The display name of the person who uploaded the document to the shared portal
 - D The time of day the document was uploaded and whether it was edited afterward

Answer: B

The indexing method and its tuning determine the trade-off between search speed and recall for approximate nearest-neighbor search. File extensions, uploader names and upload times are metadata that the search algorithm does not use to rank vectors. The file extension is the tempting distractor because it is stored alongside each chunk.

8. A chatbot framework that builds prompts by string concatenation can let user input overwrite the developer's _____.

Answer: instructions

When user text is concatenated into the prompt without separation, it can contradict or replace the developer's intended instructions. Structured message roles, input validation and output checks reduce the risk. The accepted answers name the same part of the prompt.

9. Which control best prevents an LLM application from rendering unsafe model output in a web page?
- A Raising the maximum output token limit so responses are never cut off mid-sentence
 - B Caching each response for a full day so repeated questions are answered quickly
 - C Encoding or sanitizing model output for the context where it is rendered, such as HTML
 - D Asking the model to rewrite its answer in a formal tone before the response is displayed

Answer: C

Context-aware encoding or sanitization stops model output from being interpreted as markup or script in the browser. Token limits and caching do not change how the output is rendered. Asking for a formal rewrite is a style request, and it does not remove markup or script from the text.

10. Under the EU AI Act, which risk tier is reserved for practices that are prohibited outright, such as certain forms of social scoring?
- A Minimal risk, which covers systems that carry little or no regulated obligation
 - B Limited risk, which covers systems that carry specific transparency obligations for users
 - C High risk, which covers systems subject to conformity assessment and ongoing obligations
 - D Unacceptable risk, which covers practices that are prohibited outright

Answer: D

The EU AI Act places certain practices, including some forms of social scoring, in the unacceptable-risk tier, where they are banned. High-risk systems are allowed but face strict requirements, which is the tempting distractor for anyone who thinks of social scoring as merely regulated. Limited and minimal risk tiers carry lighter obligations.