

PRACTICE QUESTIONS

GIAC / SANS

GAIPS

GIAC AI Platform Security

SHARED BY [u/vividgecko1980](#) on aiexamcerts

EDITION Free Set 2

QUESTIONS 10

SHARED October 18, 2026

EXAM GIAC / SANS GAIPS, GIAC AI Platform Security

Shared by [u/vividgecko1980](#) 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. What does the tokenizer do in a language model pipeline?

- A Converts input text into the integer token IDs the model reads
- B Stores the vector index that holds the knowledge base for later similarity search
- C Chooses which GPU runs each layer of the network during a serving request
- D Encrypts the prompt before it is transmitted over the network to the endpoint

Answer: A

The tokenizer maps text to the discrete IDs the model reads and maps generated IDs back to text. Vector storage, GPU placement and encryption are handled by other parts of the serving stack. Encryption is the distractor a security reader is most likely to pick, but transport protection happens at the TLS layer.

2. Increasing chunk overlap in a retrieval index always improves answer accuracy, because repeated text gives the model more evidence to work with.

- A True
- B False

Answer: B (False)

Overlap can help keep sentences that straddle a boundary, but it also duplicates content, adds index size and can fill the context with near-identical passages. The benefit depends on the corpus and the task, so 'always' is wrong. The statement is false.

3. What is the main purpose of a moderation classifier placed on model output in a generative AI application?

- A Compressing the output so that it is delivered to the user more quickly on slow networks
- B Flagging or blocking policy-violating content before it reaches users
- C Translating the output into other languages for users in different regions
- D Ranking the output by relevance to a search query before it is shown on the page

Answer: B

A moderation classifier checks the generated text against policy categories and can block or flag it, which adds a control that does not depend on the main model's refusal behavior. Compression, translation and relevance ranking are unrelated to policy enforcement. Ranking is the distractor a search engineer might suggest because the output is also scored.

4. Which pattern is most appropriate when one agent delegates work to another agent that runs in a separate trust domain?
- A Sharing the delegating agent's full credentials with the other agent for the length of the task
 - B Passing raw conversation logs as the only interface between the two agents in the workflow
 - C Authenticating each agent, scoping delegated permissions and validating returned results
 - D Allowing the sub-agent to modify the parent agent's system prompt when it finds a better approach

Answer: C

Mutual authentication, scoped delegation and validation of returned results keep a sub-agent from exceeding its mandate and stop its output from being trusted blindly. Sharing full credentials gives the sub-agent the parent's entire authority. Letting a sub-agent rewrite the parent's system prompt hands it control over the parent's behavior.

5. The practice of recording where each training dataset came from and under what license it may be used is called data _____.

Answer: provenance (also accepted: lineage)

Provenance and lineage both describe the documented origin and processing history of a dataset, which underpins licensing compliance and poisoning investigations. The accepted answers are the two standard terms for this record. Dataset versioning is related but does not by itself record the source and license.

6. What is the main risk of an LLM application that installs third-party packages without pinning versions or verifying them?
- A Slower model responses caused by larger dependency packages loaded into the serving process
 - B Higher GPU memory use caused by the extra libraries that load alongside the model weights
 - C Reduced embedding quality caused by a library that changes the default normalization method
 - D Malicious or typosquatted dependencies installed without detection

Answer: D

Unpinned and unverified dependencies let an attacker publish a look-alike or hijacked package that the build installs automatically, which is a supply chain compromise. Latency, memory use and embedding quality are performance or accuracy effects and are not security outcomes. Embedding quality is the tempting distractor because library updates can change model behavior.

7. Which two measures reduce the exposure of a model serving endpoint? (Choose two.)
- A Binding the server to all network interfaces so that monitoring tools can reach it directly
 - B Running the model server with root privileges so it can read any file it needs on disk
 - C Placing the endpoint behind a gateway that enforces authentication and rate limiting
 - D Disabling request size limits so that clients can send large documents in one call
 - E Restricting outbound connections from the serving container to the destinations it requires

Answer: C, E

An authenticating gateway with rate limits controls who can call the endpoint and how often, and egress restrictions limit what a compromised server can contact. Binding to all interfaces, running as root and removing size limits each increase exposure or abuse potential. Binding to all interfaces is the tempting option because it looks like a convenience for monitoring.

8. An agent reads a web page that contains hidden instructions to send internal data to an external address. Which architectural measure most directly reduces the impact?
- A Separating untrusted page content from the agent's instructions and limiting outbound actions
 - B Raising the agent's sampling temperature so it is less likely to follow a single instruction exactly
 - C Storing fetched web pages in a larger cache so they can be reviewed later by an administrator
 - D Running the agent on a faster CPU so that it finishes reading each page before any injection can act

Answer: A

Keeping untrusted content out of the instruction channel and limiting outbound actions means a hidden instruction cannot easily cause data to leave. Temperature and CPU speed do not change whether the agent obeys injected text. Caching pages stores the malicious content for longer and does not stop the agent from acting on it.

9. What does a software bill of materials for an AI system usually aim to list?
- A The employees who labeled each batch of training examples during the project
 - B System components and dependencies, such as libraries and model artifacts
 - C The office locations where the training team and the support desk are based
 - D The GPU purchase prices and invoice numbers paid for each hardware order

Answer: B

An SBOM inventories the components that make up a system, including libraries, images and model artifacts, so vulnerabilities and license issues can be traced. Labeler names, office locations and purchase records are administrative details that the inventory does not need. Labeler records are closer to data provenance than to a bill of materials.

10. If the retriever checks permissions only after the model has already seen a document, the generated answer can reveal content that the user is not authorized to read.
- A True
 - B False

Answer: A (True)

Once a document has been placed in the model's context, the model can draw on it in the answer even if a later filter hides the source. Authorization must be enforced before content reaches the model. The statement is true.