PRACTICE QUESTIONS

Linux Foundation / CNCF

SAIA

Secure AI Associate

PUBLISHED BY aiexamcerts

EDITION	Free Set 1
QUESTIONS	10
ISSUED	November 8, 2026
EXAM	Linux Foundation / CNCF SAIA, Secure AI Associate

Shared by **aiexamcerts** on aiexamcerts

10 practice questions for the Linux Foundation / CNCF SAIA exam, Secure AI Associate, each followed by its answer and a short explanation. More free sets and the full practice exam: **aiexamcerts.com/linux-foundation-cncf/saia**. Spotted a mistake? Email support@aiexamcerts.com with the exam code SAIA and the question number.

---

1. A system prompt instruction to ignore malicious instructions is by itself a sufficient defense against prompt injection.

A True

B False

Answer: B (False)

Model-level instructions can be overridden, so injection defense needs architectural controls like privilege limits, input handling, and output validation.

---

2. Which two practices help detect poisoned training data? (Choose two.)

A Statistical outlier and anomaly detection on incoming samples

B Raising the learning rate

C Reviewing the provenance and trust level of each data source

D Training for additional epochs

E Deleting dataset checksums

Answer: A, C

Screening samples and vetting sources catch poisoning before it is learned. Removing checksums eliminates the integrity evidence needed for detection.

---

3. Why is loading an untrusted pickle-serialized model file dangerous?

A Unpickling can execute arbitrary code embedded in the file

B It permanently disables GPU acceleration

C It converts the weights to a smaller precision

D It always produces a lower-accuracy model

Answer: A

Pickle can run attacker-supplied code during deserialization. Formats like safetensors store only tensors and avoid that risk.

---

4. When a model is retired, its credentials should be revoked and its inventory status updated.

A True

B False

Answer: A (True)

Leaving access and stale records behind creates orphaned risk, so retirement includes closing both.

---

5. Which two practices secure an LLM agent that can call external tools? (Choose two.)

A Give the agent one broad administrator token

B Issue scoped, short-lived credentials for each tool

C Let the model decide its own permissions at runtime

D Require human approval before irreversible actions

E Disable logging of tool calls to protect privacy

Answer: B, D

Scoped credentials and human approval limit blast radius. Letting the model set its own permissions removes the control entirely.

---

6. Version 7 of a model shows a regression after promotion. What practice allows a prompt, safe recovery?

A Keep prior versions immutable in the registry for rollback

B Retrain from scratch before any rollback is allowed to take place

C Delete the earlier versions as soon as possible in order to save space

D Overwrite the old version with the new one under the same version identifier

Answer: A

Immutable, staged versions make rollback fast and auditable. Overwriting the old version removes the recovery point.

---

7. A critical CVE is published in a library used by some of your inference services. What is the first step?

A Delete the model registry so the affected artifacts can no longer be pulled

B Retrain every model from scratch so the new weights avoid the vulnerable code

C Use SBOM or AI-BOM data to identify which services include the affected version

D Wait until attackers begin exploiting the library before taking any action

Answer: C

Bill-of-materials data lets you scope exposure quickly. Retraining models does not patch a vulnerable library.

---

8. Which NIST AI RMF function covers establishing policies, accountability structures, and a risk-aware culture across the organization?

A Manage

B Measure

C Map

D Govern

Answer: D

Govern is the cross-cutting function for policies, roles, and culture. Map is the tempting choice, but it is about establishing context and identifying risks for a specific system.

---

9. A dataset is licensed for non-commercial research only, but the team plans to train a commercial product on it. What should they do?

A Review the license terms and obtain permission or use a dataset with suitable terms

B Proceed with training because the data is publicly downloadable from the website

C Remove the license file from the downloaded archive before training begins

D Train first and review the license terms only after the product has launched

Answer: A

Permitted use is set by the license, not by availability. Public downloadability, the tempting argument, does not grant commercial rights.

---

10. The framework of build-integrity levels called Supply-chain Levels for Software Artifacts is abbreviated _____.

Answer: SLSA

SLSA defines progressively stronger guarantees around source, build, and provenance.