

PRACTICE QUESTIONS

AWS

AIP-C01

AWS Certified Generative AI Developer - Professional

SHARED BY [u/Calm-Jaguar-3655](#) on aiexamcerts

EDITION Free Set 1

QUESTIONS 10

SHARED October 15, 2026

EXAM AWS AIP-C01, AWS Certified Generative AI Developer - Professional

Shared by [u/Calm-Jaguar-3655](#) on aiexamcerts

10 practice questions for the AWS AIP-C01 exam, AWS Certified Generative AI Developer - Professional, each followed by its answer and a short explanation. More free sets and the full practice exam: www.aiexamcerts.com/aws/aip-c01. Spotted a mistake? Email support@aiexamcerts.com with the exam code AIP-C01 and the question number.

1. A team must show auditors which source passages support each generated answer. What should the application return with every answer?
- A Only the final answer text, with the retrieved passages discarded to save storage
 - B Source attribution and citation metadata for the retrieved passages
 - C The query text stored alone in a separate table
 - D A single shared document reused for every tenant

Answer: B

Returning citations and attribution metadata lets auditors trace each claim to the passage that supported it. Discarding passages is tempting to save storage, but it removes the evidence. Storing queries alone and reusing one shared document do not establish traceability.

2. The Bedrock unified API for conversational models that accepts messages and tool configurations is called the _____ API.

Answer: Converse (also accepted: Converse API)

The Converse API standardizes message and tool handling across supported models, which simplifies moving between them. Streaming variants are tempting, but they use the same unified request model. InvokeModel is a lower-level call whose request body depends on the model.

3. Which Bedrock Agents setting hands the proposed action and its parameters back to the application so the application runs it?
- A Session attributes
 - B Prompt override
 - C Knowledge base retrieval
 - D Return of control

Answer: D

Return of control passes the requested action back to the application, which can run it under its own rules and then report the result. Session attributes are tempting because they carry state, but they store values rather than handing over actions. Prompt override changes instructions and knowledge base retrieval fetches documents.

4. Step Functions retries a failed task only when the state definition includes a retry policy.
- A True
 - B False

Answer: A (True)

Retries are configured in the state definition with retry policies that name error types, intervals, and maximum attempts. The statement is true. Teams often assume retries happen automatically, which is the tempting error.

5. Which two actions support human oversight of an agent that can change customer accounts? (Choose two.)
- A Letting the agent approve its own changes after a self-review
 - B Giving the agent administrator rights so it is never blocked
 - C Requiring confirmation before any account change is executed
 - D Disabling audit trails to reduce storage costs
 - E Logging each proposed action with the user and the reason

Answer: C, E

Confirmation gates and action logs keep a person in control of consequential changes and leave a record of decisions. Self-approval removes the human check, which is the tempting shortcut for automation. Administrator rights and disabled audit trails both increase risk.

6. Which Guardrails feature detects personal identifiers such as names or account numbers and can mask or block them?
- A Denied topics
 - B Contextual grounding check
 - C Model invocation logging of requests
 - D Sensitive information filters

Answer: D

Sensitive information filters detect personal identifiers and either mask or block them according to configuration. Denied topics is tempting because it also filters responses, but it blocks subjects rather than identifiers. Grounding checks test source support, and invocation logging records requests without filtering them.

7. A nightly job classifies a very large document set and no answer is needed in real time. Which Bedrock option is most appropriate?
- A Provisioned Throughput reserved for the full day
 - B Streaming responses to the user's browser
 - C A real-time endpoint using the largest model available
 - D Batch inference

Answer: D

Batch inference processes large offline sets at lower cost than real-time calls, which suits overnight work. Provisioned Throughput is tempting when volume is high, but it reserves capacity for steady demand and costs more when idle. Streaming and the largest real-time model target interactive use.

8. Reserving model capacity for predictable, steady traffic is the purpose of Provisioned _____.

Answer: Throughput

Provisioned Throughput reserves model units for predictable throughput and latency, which suits steady high-volume traffic. On-demand is tempting because it avoids reservations, but it bills per request without reserved capacity. Batch processing handles offline work instead.

9. What is a golden dataset used for in a generative AI release pipeline?
- A Training the foundation model from scratch
 - B Replacing the production database
 - C Encrypting customer prompts at rest across every application and environment
 - D Regression testing of answers against expected results before release

Answer: D

A golden dataset of representative inputs with expected outputs or criteria lets a team detect regressions before a change ships. Training from scratch is tempting because datasets usually feed training, but a golden set is held back for evaluation. Database replacement and encryption are unrelated to the test set.

10. A Bedrock agent gives wrong answers in production. What should the team examine first?
- A The agent trace for the failing session, showing its steps, tools, and passages
 - B Raising the max tokens setting for all users
 - C Deleting the agent and creating a new one from scratch
 - D Disabling the knowledge base so the agent relies only on what it learned in training

Answer: A

Agent traces show each step, tool call, and retrieved passage, which locates whether the problem is retrieval, orchestration, or instructions. Raising max tokens is tempting as a quick fix, but it does not explain the error. Deleting the agent or disabling knowledge removes evidence and capability.