Boost up Your Certification Score

SAP

C_AIG

SAP Certified - SAP Generative AI Developer

For More Information – Visit link below:

https://www.examsboost.com/

Product Version

- ✓ Up to Date products, reliable and verified.
- ✓ Questions and Answers in PDF Format.

Latest Version: 4.0

1. Micro Skill Drill Exam
2. Unified Scenario Exam

Topic: 1
Micro Skill Drill Exam

Question: 1

A distribution-center assistant was introduced in a platform-based SAP Generative AI Hub environment to help supervisors explain stock exceptions. The rollout is considered successful for complex exception cases, but platform monitoring now shows a growing share of requests comes from automated nightly health checks that ask the assistant to summarize fixed threshold violations already available in deterministic dashboards.

Runtime budget is tightening, and peak-hour supervisors are starting to see slower response times for the complex cases they actually need help with. The operations team will not accept removing the assistant entirely because it has improved handling for ambiguous stock discrepancies. The budget cannot be increased this quarter, and the team must preserve clean-core separation rather than embedding generative logic into the core warehouse application.

What is the best next action?

Response:

A. Lower the model temperature so nightly health-check requests consume less variability during batch execution.

B. Replace the current model with a smaller one for all stock-exception and nightly health-check requests.

C. Keep the current design and shorten all generated summaries so the shared runtime budget stretches further.

D. Route deterministic nightly health-check explanations to a non-generative path and reserve the generative route for ambiguous supervisor-facing exception cases.

Answer: D

Explanation:

Feedback:

The problem is a use-case boundary issue, not a general model failure. The dependency chain is: request-type classification → route eligibility → budget consumption and latency → preserved value for ambiguous cases. Moving fixed threshold explanations to a deterministic path protects platform budget and supervisor response time while keeping generative support where it adds real decision value.

Question: 2

A supplier-risk assistant in a platform-based SAP Generative AI Hub environment compares two governed prompt variants during controlled evaluation before productive routing is updated. The refreshed evaluation component now returns a quality score where higher values indicate better grounding and actionability. After the update, reviewers notice that productive recommendations are becoming less helpful even though evaluation coverage has improved.

The comparison trace shows the selection logic still sorts candidate prompt variants in ascending order and automatically promotes the lowest-scoring variant when both pass the minimum threshold. The business wants to keep automated prompt-variant selection because it speeds controlled improvement cycles, but it cannot continue promoting weaker variants into production.

What is the best next action?

Response:

A. Lower the minimum evaluation threshold so more prompt variants qualify and the selector has a broader choice set during promotion.

B. Add more few-shot examples to both prompt variants while leaving the current ascending selection logic unchanged.

C. Normalize the promotion logic to select the highest-scoring eligible variant under the current evaluation semantics, then validate selection outcomes on representative comparisons.

D. Lower the model temperature for the weaker prompt variant so its recommendations appear more stable before the next comparison cycle.

Answer: C

Explanation:

Feedback:

The issue is not missing evaluation data. It is misinterpreted score semantics in the promotion logic. The dependency chain is: evaluation-score meaning → selection ordering logic → chosen prompt variant → productive recommendation quality. Updating the selector to choose the highest-scoring eligible variant aligns the automation with the current evaluation model.

Question: 3

A legal-review extension in a platform-based SAP Generative AI Hub environment now uses streamed responses so reviewers can see long clause analyses sooner. After a proxy update, the AI runtime still emits all chunks successfully, but some saved analyses contain jumbled paragraphs or repeated fragments. The transport trace shows each streamed chunk now includes an explicit sequence number, while the consuming application still appends chunks in arrival order.

Under stable connections the issue is rare, but during peak network variation later chunks sometimes arrive first. The business team cannot disable long-form analysis because it is required for complex reviews, and they cannot move back to manual copy-paste because volume is too high. They need the best correction at the streaming integration layer.

What is the best next action?

Response:

A. Increase the application timeout so streamed chunks have more time to arrive before the final analysis is saved.

B. Reduce the number of concurrent legal-review requests so chunk order is less likely to vary during busy periods.

C. Add stronger prompt instructions asking the model to repeat key points in each streamed chunk so out-of-order arrivals are less damaging.
D. Reassemble streamed output by sequence number rather than arrival order, then validate saved analyses under delayed and reordered chunk conditions.

Answer: D

Explanation:
Feedback:
The runtime is producing the full response, so the defect lies in stream reassembly. The dependency chain is: chunk sequence metadata → consumer reconstruction logic → final saved analysis order → reviewer-facing correctness. Using sequence numbers instead of arrival order fixes the real contract change introduced by the proxy and preserves the streaming design.

Question: 4

A compensation-policy assistant is supported through a governed console in a cloud-only SAP Generative AI Hub environment. During a control review, the team finds that a temporary rollout setting still allows a broad first-line support group to export productive evaluation results that include case-level scoring and grounded evidence references. The broader group still needs dashboard visibility to confirm whether failures occurred in retrieval, evaluation, or response generation.
However, current governance policy allows export of productive evaluation artifacts only to a smaller approved support role because those exports can expose sensitive policy interpretation details. The assistant cannot be paused because employee traffic is steady, and first-line troubleshooting visibility must remain available.
What is the best next action?
Response:

A. Keep evaluation export available because the first-line support group has not misused it since rollout.
B. Remove all console visibility from the first-line support group so no productive evaluation details remain accessible outside the approved role.
C. Restrict productive evaluation export to the approved support role while preserving view-only diagnostic visibility for first-line support, then verify incident triage coverage.
D. Add a process note telling first-line support to export productive evaluation artifacts only when a severe incident is declared.

Answer: C

Explanation:
Feedback:
The issue is an operational capability-scope mismatch, not an assistant-output problem. The dependency chain is: console export permission → access to productive evaluation artifacts → governance-compliant support boundary → retained first-line troubleshooting capability. Restricting export while preserving view-only diagnostics restores the required control model without removing legitimate monitoring support.

Question: 5

A service-renewal assistant runs in a platform-based, cloud-only SAP Generative AI Hub environment. The governed prompt design includes a required system instruction block, a reusable tone module, and a region-specific compliance insert. In workspace validation, the full assembled prompt produces correct renewal language. In productive use, however, responses for regulated regions now miss the required compliance sentence even though the core renewal summary is still correct.

The route trace shows the productive assembly step references the current system instruction block and tone module, but the compliance insert pointer still targets an archived artifact version that no longer contains the approved text. The assistant must remain live today, and the team cannot hard-code regional wording in the consuming application because the governed prompt lifecycle must remain the control point.

What is the best next action?

Response:

A. Strengthen the tone module so renewal responses sound more formal even when the compliance insert is missing.

B. Update the productive prompt assembly to reference the current approved compliance insert, then validate the full assembled prompt through the live route.

C. Ask users in regulated regions to append the missing compliance wording manually until the next prompt review cycle.

D. Switch all renewal traffic to a generic fallback template that avoids region-specific language entirely.

Answer: B

Explanation:

Feedback:

The issue is a prompt-composition lifecycle mismatch, not a general content-quality problem. The dependency chain is: prompt artifact reference → assembled governed prompt → region-specific compliance text presence → productive output compliance. Updating the live assembly to point to the current approved compliance insert fixes the actual binding defect while preserving the governed prompt architecture.

Question: 6

A policy-drafting assistant in a platform-based, cloud-only SAP Generative AI Hub environment generates clause summaries from approved source documents. After a recent route optimization, users notice that regenerated summaries still reflect the previous document revision even though the updated source file is visible in the workspace.

The execution trace shows the orchestration route is reusing a cached intermediate result keyed only by document ID, while the source document now has a newer revision under the same ID. The team must restore correct behavior today because legal reviewers are already using the assistant, and they cannot disable the optimization entirely for all use cases due to response-time targets. They need the smallest correction that fixes the stale-output problem at the right dependency layer.

What is the best next action?

Response:

A. Update the cache key or invalidation logic so document revision is included in the orchestration path, then validate summary output against the latest source revision.
B. Increase the maximum response length so the model has more space to reflect the updated clause language in regenerated summaries.
C. Add a stronger prompt instruction telling the model to prioritize the newest wording whenever multiple clause versions are present.
D. Ask legal reviewers to manually clear and rerun affected summaries until the next route redesign is scheduled.

Answer: A

Explanation:
Feedback:
The artifact shows a lifecycle-state mismatch between current source content and cached orchestration output. The dependency chain is: source revision state → cache key or invalidation behavior → reused intermediate result → final summary correctness. Including document revision in cache handling corrects the upstream dependency instead of treating the stale answer as a wording problem.

Question: 7

A benefits-policy assistant is live in a cloud-only SAP Generative AI Hub environment and is supported through a governed operational console. During a routine control review, the team finds that a temporary rollout setting still allows a broad analyst group to duplicate productive prompt artifacts into a sandbox workspace for troubleshooting. The broader group still needs read-level visibility into productive artifact versions and execution states for first-line support.
However, current lifecycle policy allows duplication of productive artifacts only to a smaller approved engineering role because copied artifacts can be modified outside the primary governed path. The assistant cannot be paused because employee traffic is steady, and troubleshooting visibility must remain available.
What is the best next action?
Response:

A. Keep duplication available because the broader analyst group has not misused the copied artifacts since rollout.
B. Remove all productive-console visibility from the broader analyst group so no artifact actions remain possible outside the approved role.
C. Restrict productive-artifact duplication to the approved engineering role while preserving read-only artifact and execution visibility for the broader analyst group.
D. Add a process note instructing the broader analyst group to duplicate productive artifacts only after getting approval in a separate channel.

Answer: C

Explanation:

Feedback:
The issue is an operational capability-scope mismatch, not an assistant-output defect. The dependency chain is: console role capability → ability to duplicate governed productive artifacts → lifecycle-control integrity → compliant first-line support model. Restricting duplication while preserving read-only visibility restores the required control boundary without weakening operational monitoring.

Question: 8

A maintenance-guidance assistant is running in a platform-based, cloud-only SAP Generative AI Hub environment. The team is in a phased modernization from an older grounded document library to a newly reprocessed library. Current-product questions return accurate grounded answers, but technicians working on older assets now receive generic responses or irrelevant procedures.
The retrieval trace shows live semantic search is querying both libraries together, while the legacy library still uses an older embedding configuration and is returning low-relevance fallback snippets.
Operations cannot pause the assistant because field teams rely on it daily, and the migration must continue incrementally rather than through a one-time cutover. The team needs the best corrective action at the retrieval compatibility layer.
What is the best next action?
Response:

A. Increase the number of retrieved snippets so older maintenance procedures have a better chance of appearing in the final context.
B. Exclude the non-reindexed legacy library from live semantic retrieval until it is reindexed under the active embedding configuration, then validate mixed-product queries.
C. Add prompt instructions telling the model to prefer older procedures whenever an asset appears to belong to an earlier product generation.
D. Keep both libraries active and lower response variability so irrelevant grounded content affects the final answer less often.

Answer: B

Explanation:
Feedback:
The issue is not weak prompting or random generation. It is a coexistence compatibility problem between active retrieval sources. The dependency chain is: embedding configuration alignment→ comparable similarity scoring across libraries → eligible grounded evidence → maintenance answer quality. Keeping an incompatible legacy library in live semantic retrieval introduces low-quality evidence. Excluding it until reindexing restores a valid retrieval basis during incremental modernization.

Question: 9

A reimbursement-guidance assistant is supported through a governed operational console in a cloud-only SAP Generative AI Hub environment. During a quarterly access review, the team discovers that a temporary rollout setting still allows a broad first-line support group to approve activation of archived prompt variants in a replay-only test lane that mirrors productive behavior closely. The broader group still needs read-level visibility into variant history, replay outcomes, and execution-state diagnostics for incident triage.

However, current governance policy allows archived-variant activation only to a smaller approved engineering role because replay activation can regenerate controlled policy interpretations and influence support decisions if used incorrectly. The assistant cannot be paused because employee usage is steady, and first-line troubleshooting visibility must remain available.

What is the best next action?

Response:

A. Keep archived-variant activation available because the first-line support group has not misused it since rollout.

B. Remove all console visibility from the first-line support group so no archived-variant actions remain possible outside the approved role.

C. Restrict archived-variant activation to the approved engineering role while preserving read-only variant history, replay visibility, and diagnostic access for first-line support.

D. Add a process note telling first-line support to activate archived prompt variants only after separate engineering approval during severe incidents.

Answer: C

Explanation:

Feedback:

The issue is an operational capability-scope defect, not an assistant-output problem. The dependency chain is: console capability → ability to activate archived governed variants → lifecycle-control integrity → compliant first-line support model. Restricting activation to the approved engineering role restores the proper control boundary while preserving legitimate read-only visibility for incident triage.

Question: 10

A product-return assistant is active in a platform-based SAP Generative AI Hub environment and grounds its answers on approved regional return-policy documents. After a phased content migration, users asking about bundle returns now receive accurate answers for current bundles but inconsistent answers for retired bundle packages that still require historical policy support.

The retrieval trace shows retired bundle documents were ingested successfully, but their package identifiers were stored with hyphens removed during the migration, while live user queries still send the formatted identifier with hyphens. The assistant must remain live because fulfillment teams use it daily, and the team cannot disable retired-package retrieval because historical returns still occur and must follow the approved archived policy.

What is the best next action?

Response:

A. Normalize package-identifier handling between migrated source metadata and live retrieval queries, then validate both current and retired bundle-return scenarios.

B. Increase the number of retrieved snippets so archived bundle-return policies have a better chance of appearing with current policy results.

C. Add a prompt instruction telling the model to interpret formatted and unformatted package identifiers as equivalent whenever results seem incomplete.

D. Remove the retired-package filter temporarily so users at least receive a general return answer during the migration period.

Answer: A

Explanation:
Feedback:
The problem is not missing archived content. It is an identifier-format mismatch at the retrieval-eligibility layer introduced during migration. The dependency chain is: migrated source identifier format → live query filter match → eligible archived or current policy subset → correct grounded bundle-return answer. Normalizing identifier handling restores dependable access to both current and historical policy content.

Topic: 2
Unified Scenario Exam

Question: 11

Unified Scenario: Retail Service Launch Readiness with SAP Generative AI Hub
CHALLENGE 1 — Reusable Prompt Assets Across Multiple Entry Points
During cutover rehearsal, service agents submit the same claim-explanation request from two different application entry points. The outputs differ in tone and structure even though the business intent is the same. Which action should the development team take first to support first-wave consistency and later reuse?
Response:

A. Maintain separate prompt templates in each application so teams can refine them faster during rollout
B. Route both entry points to the same governed prompt asset in the shared prompt management path
C. Allow each application owner to keep a local prompt copy until the seasonal peak has passed
D. Increase the document collection used for grounding so the style differences become less visible

Answer: B

Explanation:
Feedback:
Using the same governed prompt asset across both entry points addresses the most direct cause of inconsistent output behavior in this scenario. It also supports the company's later goal of reusing prompt assets across additional applications without multiplying maintenance paths.

Question: 12

Unified Scenario: Retail Service Launch Readiness with SAP Generative AI Hub
CHALLENGE 1 — Reusable Prompt Assets Across Multiple Entry Points
A developer proposes a quick workaround: keep the shared prompt registry for governance, but let each consuming application apply small local edits to the prompt text before runtime execution. What is the strongest reason this is a poor choice for the pilot launch?

Response:

A. It prevents grounded responses from using refreshed enterprise documents
B. It removes the need for prompt templates altogether
C. It weakens controlled reuse because the shared asset no longer defines the actual runtime behavior consistently
D. It forces the team to widen model access during the pilot

Answer: C

Explanation:
Feedback:
This is a partial-fix trap. It keeps the appearance of central governance but breaks runtime consistency because local edits change the actual executed prompt behavior per entry point. That undermines both launch stability and later reuse.

Question: 13

Unified Scenario: Retail Service Launch Readiness with SAP Generative AI Hub
CHALLENGE 2 — Controlled Model Access for Launch Window Stability
During final rehearsal, developers want broad model switching enabled in the live pilot so they can compare latency and output quality while agents are already using the assistant. Platform operations prefers a narrower approved runtime set for the first wave. Which decision best fits the scenario?
Response:

A. Expose all available models in production for the pilot because broader comparison will shorten later optimization cycles
B. Restrict live usage to an approved runtime set and keep broader comparisons in controlled evaluation activity outside the launch path
C. Delay the launch until every available model has been benchmarked under live business traffic
D. Let service agents choose the model at runtime so real usage patterns determine the best option quickly

Answer: B

Explanation:
Feedback:
This balances performance evaluation with operational control. The scenario emphasizes launch stability, support ownership, and repeatable validation, so live runtime behavior should remain narrow while experimentation continues in a separate controlled path.

Question: 14

Unified Scenario: Retail Service Launch Readiness with SAP Generative AI Hub
CHALLENGE 2 — Controlled Model Access for Launch Window Stability

The team notices one alternative model produces faster troubleshooting drafts, but that model has not yet been included in the approved launch runtime path. Which is the best next step for the pilot window?
Response:

A. Add the faster model directly to the live runtime path because performance should outweigh governance in a seasonal launch
B. Keep the approved live runtime path intact and continue evaluating the faster model outside the pilot's controlled production scope
C. Replace prompt templates with free-text prompting so the faster model can be used more flexibly
D. Expand retrieval collections first, then decide whether the faster model should become the default in production

Answer: B

Explanation:
Feedback:
This is the best performance-versus-governance weighting in the scenario. It allows continued optimization work without destabilizing the governed runtime boundary needed for the first production wave.

Thank You for Trying Our Product

For More Information – Visit link below:

https://www.examsboost.com/

15 USD Discount Coupon Code:

G74JA8UF

FEATURES

- ✓ 90 Days Free Updates
- ✓ Money Back Pass Guarantee
- ✓ Instant Download or Email Attachment
- ✓ 24/7 Live Chat Support
- ✓ PDF file could be used at any Platform
- ✓ 50,000 Happy Customer

60 DAY MONEY BACK GUARANTEE 100% SATISFACTION GUARANTEED