



Accelerate Your Certification with Microsoft AI-300 Practice Questions

Last chance to prepare smart! Get your hands on free Microsoft Operationalizing Machine Learning and Generative AI Solutions exam PDF questions. Study real AI-300 dumps with verified answers and fast-track your certification success with [PrepBolt](https://prepbolt.com/AI-300.html) AI-300 exam pdf questions and answers.

Thank you for Downloading AI-300 exam PDF Demo

<https://prepbolt.com/AI-300.html>

QUESTIONS & ANSWERS
DEMO VERSION
(LIMITED CONTENT)

Question 1

Question Type: MultipleChoice

Case Study: Mix Questions

Mix Questions

AI-300 Mix Questions IN THIS CASE STUDY

A Retrieval-Augmented Generation (RAG) solution returns incomplete answers because relevant content is inconsistently retrieved from the knowledge source.

You need to improve RAG accuracy without changing the embedding model currently in use. You need to achieve this goal while minimizing operational costs.

Which two actions should you perform? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point. Choose two.

Options:

- A- Tune chunk size and overlap to match content structure.
- B- Implement an optimized re-ranker.
- C- Increase token limits for all requests.
- D- Optimize the length of embedding vectors.

Answer:

A, B

Question 2

Question Type: Hotspot

Case Study: Mix Questions

Mix Questions

AI-300 Mix Questions IN THIS CASE STUDY

You manage a Retrieval-Augmented Generation (RAG) system that uses Azure AI Search to retrieve documents from an indexed knowledge base.

The system must support the following retrieval requirements:

Queries that include exact policy identifiers must return matching documents even when semantic similarity is low.

Natural-language questions must prioritize semantically relevant documents even when keywords are not an exact match.

You need to configure the retrieval approach to meet the requirements.

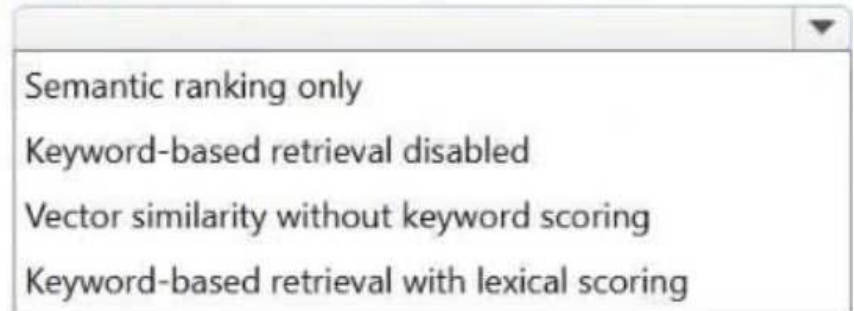
How should you configure the retrieval behavior for each requirement? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.

Hybrid retrieval configuration

Retrieval requirement

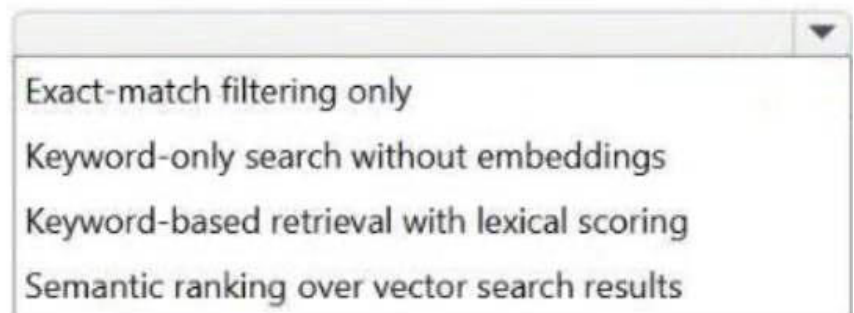
Exact identifier matching

Configuration



A dropdown menu with a downward arrow on the right. The menu is open, showing four options: "Semantic ranking only", "Keyword-based retrieval disabled", "Vector similarity without keyword scoring", and "Keyword-based retrieval with lexical scoring".

Natural-language relevance



A dropdown menu with a downward arrow on the right. The menu is open, showing four options: "Exact-match filtering only", "Keyword-only search without embeddings", "Keyword-based retrieval with lexical scoring", and "Semantic ranking over vector search results".

Answer:

See the Answer in the Premium Version!

Question 3

Question Type: Hotspot

Case Study: Mix Questions

Mix Questions

AI-300 Mix Questions IN THIS CASE STUDY

You are reviewing a dataset that will be used for an advanced fine-tuning job in Microsoft Foundry.

The fine-tuning job uses preference comparison data.

You review the following dataset excerpt.

Training record 1:

```
01 {"input":[{"role":"system","content":"You are a policy assistant."},
02         {"role":"user","content":"Summarize Policy X for managers."}],
03  "preferred_output":[{"role":"assistant","content":"Policy X applies to all managers and outlines approval
requirements."}],
04  "non_preferred_output":[{"role":"assistant","content":"Policy X applies to all managers and outlines approval
requirements."}]}
```

Training record 2:

```
01
02 {"input":[{"role":"system","content":"You are a policy assistant."},
03         {"role":"user","content":"What is the retention period?"}],
04  "preferred_output":[{"role":"assistant","content":"The retention period is seven years."}],
05  "non_preferred_output":[{"role":"assistant","content":"The retention period is five years."}]}
```

For each of the following statements, select Yes if the statement is true. Otherwise, select No. NOTE: Each correct selection is worth one point.

Preference-based fine-tuning dataset validation

Statement	Yes	No
The first record provides a usable preference signal for preference-based fine-tuning.	<input type="radio"/>	<input type="radio"/>
The second record provides a usable preference signal for preference-based fine-tuning.	<input type="radio"/>	<input type="radio"/>
Each preferred and non-preferred output in the dataset contains a valid assistant message as required for direct preference optimization.	<input type="radio"/>	<input type="radio"/>
Both records must contain different preferred and non-preferred outputs to produce usable training signals.	<input type="radio"/>	<input type="radio"/>

Answer:

See the Answer in the Premium Version!

Question 4

Question Type: Hotspot

Case Study: Mix Questions

Mix Questions

AI-300 Mix Questions IN THIS CASE STUDY

You manage a Retrieval-Augmented Generation (RAG) system that retrieves internal policy documents from a vector index.

Recent analysis shows that:

Retrieved results frequently include duplicated content from the same document.

Retrieved chunks sometimes span unrelated policy sections.

You review the following retrieval and ingestion configurations:

Configuration 1

Parameter	Setting
Chunk size	1,200
Chunk overlap	400
Similarity threshold	0.80
Retrieval strategy	Vector
Top k	8

Configuration 2

Parameter	Setting
Chunk size	300
Chunk overlap	40
Similarity threshold	0.90
Retrieval strategy	Vector
Top k	8

You need to reduce duplicated retrieval results and improve chunk relevance across policy sections.

For each of the following statements, select Yes if the statement is true. Otherwise, select No. NOTE: Each correct selection is worth one point.

RAG retrieval and chunking configurations

Statement	Yes	No
Configuration 1 produces duplicated content across retrieved chunks.	<input type="radio"/>	<input type="radio"/>
Configuration 2 enforces stricter similarity filtering that can exclude relevant chunks.	<input type="radio"/>	<input type="radio"/>
Lowering the similarity threshold is required to reduce semantic dilution.	<input type="radio"/>	<input type="radio"/>
Switching to keyword-only search will eliminate semantic dilution while preserving semantic relevance.	<input type="radio"/>	<input type="radio"/>

Answer:

See the Answer in the Premium Version!

Question 5

Question Type: Hotspot

Case Study: Mix Questions

Mix Questions

AI-300 Mix Questions IN THIS CASE STUDY

A company is creating an internal tool that summarizes long meeting transcripts and extracts action items.

The model must:

Process text inputs up to 200k tokens long.

Generate concise summaries in seconds.

Support interactive testing before integration into the app.

You need to select, deploy, and test a model that supports summarization with low latency.

How should you complete the configuration plan? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.

Deploying and testing models for summarization

Requirement

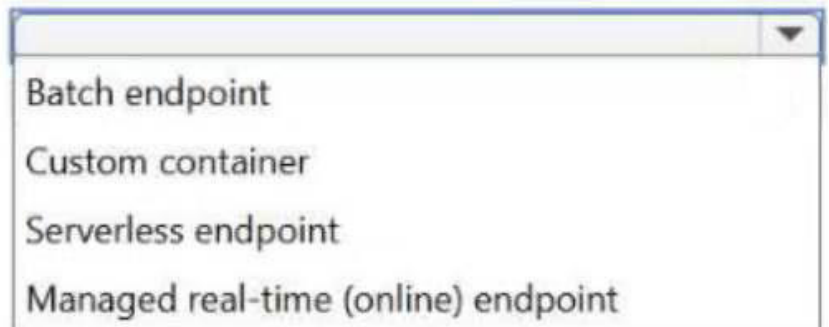
Select a model.

Tool



A dropdown menu with a downward arrow on the right. The menu is open, showing four options: Whisper, Phi-3-Mini, GPT-4.1-Mini, and Text-Embedding-Ada-002.

Select a deployment type.



A dropdown menu with a downward arrow on the right. The menu is open, showing four options: Batch endpoint, Custom container, Serverless endpoint, and Managed real-time (online) endpoint.

Test the model.



A dropdown menu with a downward arrow on the right. The menu is open, showing three options: Prompt flow, Microsoft Foundry Chat playground, and Microsoft Foundry model comparison.

Answer:

See the Answer in the Premium Version!

Question 6

Question Type: Hotspot

Case Study: Mix Questions

Mix Questions

AI-300 Mix Questions IN THIS CASE STUDY

A team is building a generative AI agent by using Retrieval-Augmented Generation (RAG) in Microsoft Foundry.

The team frequently updates prompt content. The team must be able to track changes across contributors while avoiding full application redeployments.

You need to enable rapid prompt iteration with traceability. Applications consuming the agent must be able to use updated prompts without requiring redeployment.

What should you configure for each requirement? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.

Prompt management configuration

Requirement

Track prompt changes across teams.

Configuration



- Git repository
- Azure Blob Storage
- Azure Machine Learning Datastores

Apply prompt updates without redeployment.



- CI/CD pipeline
- Environment variables
- Agent's latest published version reference

Answer:

See the Answer in the Premium Version!

Question 7

Question Type: MultipleChoice

Case Study: Mix Questions

Mix Questions

AI-300 Mix Questions IN THIS CASE STUDY

An organization is deploying several generative AI workloads by using Microsoft Foundry. Each workload must meet different requirements related to data governance, task specialization, and operational cost control.

The organization requires models that meet the following requirements:

Model behavior aligns with the task being performed.

Data handling aligns with internal governance policies.

Operational complexity and cost are justified by workload needs.

You need to select the foundation model options that meet the requirements.

Which three models can you select? Each correct answer presents a complete solution. NOTE: Each correct selection is worth one point. Choose three.

Options:

A- A model that is optimized for conversational reasoning when deploying an interactive assistant

B- The largest available model to simplify operational management

C- The smallest available model to minimize the usage cost

D- A model that supports multiple input types when workloads require combined text and image analysis

E- A model that offers enterprise governance controls when workloads process regulated business data

Answer:

B, C, E

Question 8

Question Type: MultipleChoice

Case Study: Mix Questions

Mix Questions

AI-300 Mix Questions IN THIS CASE STUDY

A team develops and manages a conversational assistant by using Microsoft Foundry.

The team must be able to validate that the assistant does not produce hateful responses before the application is exposed to any users.

You need to evaluate the model output for hateful responses as part of a repeatable validation process.

Which evaluator should you configure first?

Options:

- A- Protected material
- B- Groundedness
- C- Indirect attacks
- D- Content safety

Answer:

D

Thank You for trying AI-300 PDF Demo

To try our AI-300 practice exam software visit
link below

<https://prepbolt.com/AI-300.html>

Start Your AI-300 Preparation

Use Coupon “**SAVE50**” for extra 50% discount on the purchase of Practice Test Software. Test your AI-300 preparation with actual exam questions.