

Microsoft

Exam Questions AI-103

Developing AI Apps and Agents on Azure



NEW QUESTION 1

- (Topic 1)

You need to recommend a solution to assess the responses generated by Agent1 when the agent uses the product information stored in storage1. The solution must meet the technical requirements.

- A. a Retrieval Augmented Generation (RAG) evaluator
- B. a coherence evaluator
- C. a custom guardrail
- D. model fine-tuning

Answer: A

NEW QUESTION 2

- (Topic 1)

You need to configure Agent1 to answer customer questions about only the Contoso products. The solution must meet the business requirements. What should you do?

- A. Apply top-p sampling.
- B. Modify the system message instructions.
- C. Add few-shot examples.
- D. Increase the value of the temperature parameter.

Answer: B

NEW QUESTION 3

HOTSPOT - (Topic 1)

You need to configure the model deployment for Agent1 to meet the technical requirements.

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

Deployment type:

Version update policy:

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Deployment type:

Version update policy:

NEW QUESTION 4

HOTSPOT - (Topic 1)

You need to ensure that Agent1Dev Team can access Agent1. The solution must meet the security and compliance requirements.

How should you complete the Python code? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

```
from azure.identity import DefaultAzureCredential
from azure.ai.projects import AIProjectClient
from azure.core.credentials import AzureKeyCredential
myEndpoint = "https://contoso.services.ai.azure.com/api/projects/project1"
project_client = AIProjectClient(
    endpoint=myEndpoint,
    credential=  ,
)
myAgent = "Agent1"
agent = project_client.agents.  (agent_name=myAgent)
print(f"Retrieved agent: {agent.name}")
```

create_version

get

get_version

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area

```
from azure.identity import DefaultAzureCredential

from azure.ai.projects import AIProjectClient

from azure.core.credentials import AzureKeyCredential

myEndpoint = "https://contoso.services.ai.azure.com/api/projects/project1"

project_client = AIProjectClient(

    endpoint=myEndpoint,

    credential=  ,

)

myAgent = "Agent1"

agent = project_client.agents.  (agent_name=myAgent)

print(f"Retrieved agent: {agent.name}")
```

create_version

get

get_version

NEW QUESTION 5

- (Topic 2)

You have an Azure subscription that contains a Microsoft Foundry resource. You need to build an app that will suggest product names from a given product description. Which Foundry model should you use?

- A. GPT-4
- B. DALL-E
- C. Whisper
- D. embeddings

Answer: A

NEW QUESTION 6

- (Topic 2)

You have a Microsoft Foundry project.

You plan to build a customer support solution that contains an agent. The solution must meet the following requirements:

- Provide accurate, context-aware responses grounded in internal product documentation stored in Azure AI Search.
- Require deep, multi-step reasoning across long contexts.
- Generate detailed natural language responses.

Which type of model should you use to power the agent?

- A. a multimodal model
- B. a key phrase extraction model
- C. a small language model (SLM)
- D. a large language model (LLM)

Answer: D

NEW QUESTION 7

HOTSPOT - (Topic 2)

You have a Python application that uses Azure OpenAI structured outputs to extract fields from unstructured receipt text and support ticket text. The application uses the following schema.

```
receipt_response_format = {
  "type": "json_schema",
  "json_schema": {
    "name": "ReceiptExtraction",
    "strict": True,
```

Answer Area		
Statements	Yes	No
The schema requires that a receipt has a discount_code value.	☐	☐
The schema can be used as is for Azure OpenAI structured outputs.	☐	☐
The generated JSON object will preserve the schema property order of merchant, order_number, discount_code, line_items, and total.	☐	☐

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area		
Statements	Yes	No
The schema requires that a receipt has a discount_code value.	☐	☒
The schema can be used as is for Azure OpenAI structured outputs.	☐	☒
The generated JSON object will preserve the schema property order of merchant, order_number, discount_code, line_items, and total.	☒	☐

NEW QUESTION 8

- (Topic 2)
You have an invoice-processing application named App1 that uses Azure Content Understanding in Foundry Tools. You are building a new Content Understanding pipeline named Pipeline1 that must meet the following requirements:
? Compare an invoice to its related purchase order.
? Validate the invoice against static vendor contract documents.
? Return a single structured output that includes discrepancy findings.
You need to configure Pipeline1 and expose the pipeline as a single analyzer endpoint. What should you configure?

A. A single-file task in standard mode that uses confidence scores enabled for the extracted fields
B. A multi-file task in pro mode that uses the vendor contract files as reference data
C. A multi-file task in standard mode that uses the invoice and purchase order as input to the analyzer
D. A single-file task in standard mode that uses the vendor contract provided as an additional document during analysis

Answer: B

NEW QUESTION 9

- (Topic 2)
You have a chat app in a Microsoft Foundry project and an Azure AI Search vectorized index. You need to connect to the index to meet the following requirements:
• Complex questions must retrieve information from multiple chunks.
• Multi-turn conversations must influence retrieval planning.
• Retrievals must run in parallel to reduce latency. Which retrieval approach should you use?

A. classic Retrieval Augmented Generation (RAG)
B. chain of thought
C. agentic Retrieval Augmented Generation (RAG)
D. iterative retrieval

Answer: C

NEW QUESTION 10

- (Topic 2)

You have a Microsoft Foundry project that uses Azure AI Search to ground an agent in internal documentation. After a recent content update, users report that the agent's answers have become less accurate. You need to identify whether the retrieved content is negatively influencing the model's generated responses. Which observability signal should you review?

- A. prediction drift metrics
- B. groundedness evaluation metrics
- C. latency breakdown traces
- D. indexer status and failure history

Answer: B

NEW QUESTION 10

- (Topic 2)

You have a Microsoft Foundry project that contains an agent. The agent generates summaries from retrieved policy documents. You need to improve response completeness. The solution must be implemented in the logic of the application code before responses are returned. What should you do?

- A. Decrease the value of the temperature parameter.
- B. Add a reflection pass before the responses are returned.
- C. Switch to Retrieval Augmented Generation (RAG).
- D. Enable response streaming.

Answer: B

NEW QUESTION 13

DRAG DROP - (Topic 2)

You have a Microsoft Foundry project that contains an agent. You need to enable long-term memory to ensure that the agent can recall user preferences across separate conversations. Stored memories must be isolated per authenticated user without the client application manually generating user IDs. How should you complete the Python code? To answer, drag the appropriate values to the correct targets. Each value may be used once, more than once, or not at all.
NOTE: Each correct selection is worth one point.

Values

⋮ "session"

⋮ "{{\${conversationId}}}"

⋮ "{{\${userId}}}"

⋮ [mem_store_name]

⋮ [memory_tool]

⋮ MemorySearchTool("support_mem_store")

Answer Area

```
from azure.ai.projects.models import MemorySearchTool, PromptAgentDefinition

mem_store_name = "agent_mem_store"

memory_tool = MemorySearchTool(
    memory_store_name=mem_store_name,
    scope= 

Value

 ,
)

agent_def = PromptAgentDefinition(
    model="gpt-5.2",
    instructions="You are a customer support assistant.",
    tools= 

Value


)
```

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Values

⋮ "session"

⋮ "{{ \$conversationId }}"

⋮ "{{ \$userId }}"

⋮ [mem_store_name]

⋮ [memory_tool]

⋮ MemorySearchTool("support_mem_store")

Answer Area

```
from azure.ai.projects.models import MemorySearchTool, PromptAgentDefinition

mem_store_name = "agent_mem_store"

memory_tool = MemorySearchTool(
    memory_store_name=mem_store_name,
    scope= ⋮ "{{ $userId }}" ,
)

agent_def = PromptAgentDefinition(
    model="gpt-5.2",
    instructions="You are a customer support assistant.",
    tools= ⋮ [memory_tool]
)
```

NEW QUESTION 15

HOTSPOT - (Topic 2)

You have an Azure AI Search resource named Search1 that is used by multiple apps hosted in Azure. You need to secure Search1. The solution must meet the following requirements:

- Prevent access to Search1 from the internet.
- Limit the access of each app to query specific indexes.

What should you do? To answer, select the appropriate options in the answer area. NOTE: Each correct answer is worth one point.

Answer Area

To prevent access from the internet:

▼

Configure an IP firewall.
Create a private endpoint
Use Azure roles.

To limit access to query specific indexes:

▼

Create a private endpoint.
Use Azure roles.
User key authentication

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

Answer Area

To prevent access from the internet:

▼

Configure an IP firewall.

Create a private endpoint

Use Azure roles.

To limit access to query specific indexes:

▼

Create a private endpoint.

Use Azure roles.

User key authentication

Passing Certification Exams Made Easy

visit - <https://www.surepassexam.com>

NEW QUESTION 16

HOTSPOT - (Topic 2)

You have a Microsoft Foundry project that contains an agent. You use a GitHub Actions workflow for CI/CD. You need to configure the workflow to automatically evaluate the agent when a pull request (PR) is created and prevent branches from merging if the evaluation results do NOT meet the defined thresholds. How should you configure the workflow? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.

Authentication method:

A personal access token (PAT)

A user-assigned managed identity

An Azure Login action that uses OpenID Connect (OIDC)

If the evaluation results are NOT met, configure the workflow to:

Fail

Lock the target branch

Send an alert

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Authentication method:

A personal access token (PAT)

A user-assigned managed identity

An Azure Login action that uses OpenID Connect (OIDC)

If the evaluation results are NOT met, configure the workflow to:

Fail

Lock the target branch

Send an alert

NEW QUESTION 21

DRAG DROP - (Topic 2)

You have a Docker host named Host1 that contains a container base image. You have an Azure subscription that contains a custom speech-to-text model named model1. You need to run model 1 on Host1. Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

Actions

Retrain the model.

Configure disk logging.

Export model1 to Host1.

Run the container.

Request approval to run the container.

Answer Area

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Actions

⋮ Retrain the model.

⋮ Configure disk logging.

⋮ Export model1 to Host1.

⋮ Run the container.

⋮ Request approval to run the container.

Answer Area

⋮ Request approval to run the container.

⋮ Export model1 to Host1.

⋮ Run the container.

NEW QUESTION 25

- (Topic 2)
You have a Microsoft Foundry project that contains a model deployment.
You have an application that calls the deployment by using the Azure OpenAI v1 API and DefaultAzureCredential.
The developers at your company receive HTTP 403 errors when they send inference requests, even after running az login.
You need to ensure that the developers can perform model inference. The solution must follow the principle of least privilege.
Which role-based access control (RBAC) role should you assign to the developers?

- A. Cognitive Services OpenAI User
- B. Cognitive Services Data Reader
- C. Cognitive Services User
- D. Contributor

Answer: A

NEW QUESTION 28

HOTSPOT - (Topic 2)
You need to recommend a plan to create a customer support agent by using the Microsoft Foundry Agent Service. The agent must meet the following requirements:

- Retain user preferences across multiple conversations.
- Enable users to provide contextual grounding by directly uploading documents during a chat.

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

To retain user preferences across conversations, use:

Agent memory that uses persistent storage

Conversation history

Orchestration-managed session context

To enable users to provide contextual grounding during chats, use the:

Azure AI Search tool

Code interpreter tool

File search tool

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

To retain user preferences across conversations, use:

Agent memory that uses persistent storage

Conversation history

Orchestration-managed session context

To enable users to provide contextual grounding during chats, use the:

Azure AI Search tool

Code interpreter tool

File search tool

NEW QUESTION 29

- (Topic 2)
You have an app named App1 that uses a Microsoft Foundry multimodal model deployment.
App1 runs optical character recognition (OCR) on uploaded images and appends the OCR output to the prompt as additional context.
Some uploaded images contain embedded text.
You need to prevent potentially malicious instructions from being processed by the model. What should you use?

- A. protected material text
- B. prompt shields for user prompts
- C. image moderation
- D. prompt shields for documents

Answer: D

NEW QUESTION 33

- (Topic 2)
You have a Microsoft Foundry project that contains an agent. The agent uses Azure AI Search for Retrieval Augmented Generation (RAG). You plan to ingest and index PDF product manuals. You need to build a solution that supports semantic similarity matching. The solution must ensure that the agent retrieves relevant data when user questions use different wording than the product manuals.

- A. vector search
- B. semantic ranking
- C. suggesters
- D. analyzers

Answer: A

NEW QUESTION 35

- (Topic 2)
Note: This section contains one or more sets of questions with the same scenario and problem. Each question presents a unique solution to the problem. You must determine whether the solution meets the stated goals. More than one solution in the set might solve the problem. It is also possible that none of the solutions in the set solve the problem.
After you answer a question in this section, you will NOT be able to return. As a result, these questions do not appear on the Review Screen.
You have a multimodal AI generative model that accepts image uploads and uses extracted image text to generate responses.
You discover that users can upload unsafe images and embed hidden instructions into images to manipulate the model.
You need to implement controls to mitigate the risk. Solution: You configure protected material detection. Does this meet the goal?

- A. Yes
- B. No

Answer: B

NEW QUESTION 37

DRAG DROP - (Topic 2)
You have a Microsoft Foundry project that contains a customer support agent grounded in internal documentation.
After a recent update, users report the following issues:
• Some answers are unsupported by retrieved documents.
• A small number of responses are flagged for policy violations. You need to evaluate each issue.
Which observability signals should you use for each issue? To answer, drag the appropriate observability signals to the correct issues. Each observability signal may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.
NOTE: Each correct selection is worth one point.

Observability signals

⋮ Groundedness evaluation metrics

⋮ Latency breakdown traces

⋮ Risk and safety metrics

⋮ Token usage analytics

Answer Area

Unsupported responses:

Observability signal

Policy violations:

Observability signal

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Observability signals

⋮

Groundedness evaluation metrics

⋮

⋮

Latency breakdown traces

⋮

⋮

Risk and safety metrics

⋮

⋮

Token usage analytics

⋮

Answer Area

Unsupported responses

⋮

Groundedness evaluation metrics

Policy violations:

⋮

Risk and safety metrics

NEW QUESTION 41

HOTSPOT - (Topic 2)

You have a Microsoft Foundry project that contains two agents named PolicyWriter and RiskReviewer. PolicyWriter generates draft updates for customer policies, and RiskReviewer reviews the drafts. In the visual builder, you need to create a workflow that meets the following requirements:

- Finalizes low-risk updates without manual intervention
- Ensures predictable execution across the agents

Answer Area

Orchestration pattern:

The sequential template that passes outputs node-by-node

The group chat template to dynamically route control between the agents

The human-in-the-loop template that pauses execution of the workflow for input

Approval checkpoints:

Add a Basic chat node.

Add a condition statement.

Add an Ask a question node.

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

Answer Area

Orchestration pattern:

The sequential template that passes outputs node-by-node

The group chat template to dynamically route control between the agents

The human-in-the-loop template that pauses execution of the workflow for input

Approval checkpoints:

Add a Basic chat node.

Add a condition statement.

Add an Ask a question node.

NEW QUESTION 42

.....

Thank You for Trying Our Product

We offer two products:

1st - We have Practice Tests Software with Actual Exam Questions

2nd - Questons and Answers in PDF Format

AI-103 Practice Exam Features:

- * AI-103 Questions and Answers Updated Frequently
- * AI-103 Practice Questions Verified by Expert Senior Certified Staff
- * AI-103 Most Realistic Questions that Guarantee you a Pass on Your FirstTry
- * AI-103 Practice Test Questions in Multiple Choice Formats and Updatesfor 1 Year

100% Actual & Verified — Instant Download, Please Click
[Order The AI-103 Practice Test Here](#)