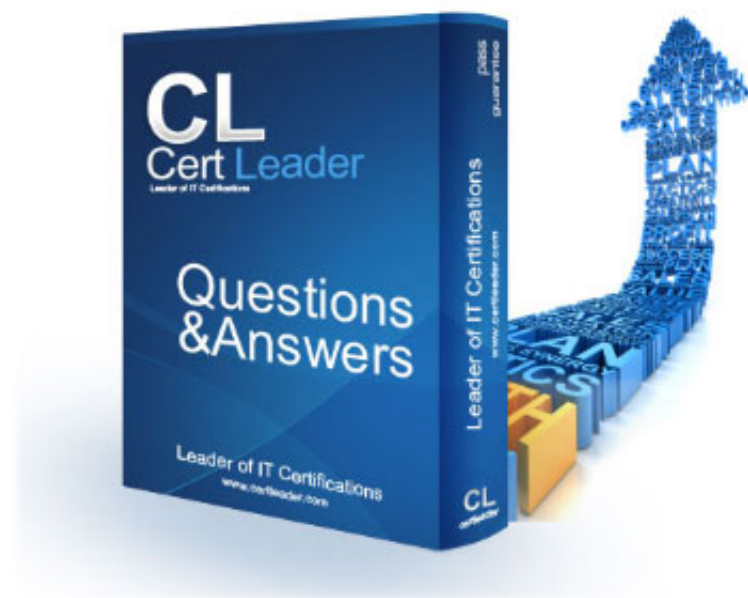


AI-103 Dumps

Developing AI Apps and Agents on Azure

<https://www.certleader.com/AI-103-dumps.html>



NEW QUESTION 1

- (Topic 1)

You need to recommend an invoice review solution that resolves the issue reported by the finance department.
What should you include in the recommendation?

- A. Azure Content Understanding in Foundry Tools
- B. chat completions
- C. Azure Document Intelligence in Foundry Tools
- D. Image Analysis

Answer: A

NEW QUESTION 2

HOTSPOT - (Topic 1)

You need to ensure that the marketing department can generate videos by using the model deployed to Project2. now 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

```
import time

from openai import OpenAI

client = OpenAI(
    base_url="https://Contoso.openai.azure.com/openai/v1/",
    api_key=...,
)

video = client.videos.
    model=deployment_name_
    prompt="A video of ou
)

while video.status not in ["completed", "failed", "cancelled"]:
    time.sleep(20)
    video = client.videos.
print(video.status)
```

create
download_content
list
retrieve

(video.id)

create
download_content
list
retrieve

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area

```
import time

from openai import OpenAI

client = OpenAI(
    base_url="https://Contoso.openai.azure.com/openai/v1/",
    api_key=...,
)

video = client.videos.create(
    model=deployment_name,
    prompt="A video of our product",
)

while video.status not in ["completed", "failed", "cancelled"]:
    time.sleep(20)
    video = client.videos.retrieve(video.id)
    print(video.status)
```



NEW QUESTION 3

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=  ,
    

AzureKeyCredential()
        DefaultAzureCredential()
        None


)
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=  ,
    

AzureKeyCredential()
        DefaultAzureCredential()
        None


)

myAgent = "Agent1"

agent = project_client.agents.  (agent_name=myAgent)

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


create_version
    get
    get_version


```

NEW QUESTION 4

- (Topic 2)

You are building an app that will use Azure AI to monitor workspaces for safety. You need to recommend a service that meets the following requirements:

? Generates alerts when employees enter high-risk areas

? Monitors video feeds in real time

? Minimizes development effort

What should you recommend?

- A. Azure Vision in Foundry Tools Image Analysis
- B. Azure AI Video Indexer
- C. Azure Vision in Foundry Tools Spatial Analysis
- D. Object detection in Azure Custom Vision

Answer: C

NEW QUESTION 5

- (Topic 2)

You have a Microsoft Foundry project that contains a prompt agent used by a customer support web app.

The agent is invoked from a Python service that does NOT run in the Foundry portal.

You need to implement end-to-end tracing to capture latency breakdowns and exceptions across agent runs.

Which two components can you use? Each correct answer presents a complete solution. NOTE: Each correct selection is worth one point.

- A. a Log Analytics workspace
- B. OpenTelemetry
- C. Application Insights
- D. the Azure Monitor Agent
- E. Microsoft Sentinel

Answer: BC

NEW QUESTION 6

- (Topic 2)

You have a Microsoft Foundry project that contains an agent.

The knowledge source for the agent is a set of scanned PDF troubleshooting guides stored in Azure Blob Storage. The guide pages contain two-column layouts and tables.

You use Azure Content Understanding in Foundry Tools to process the PDFs.

You plan to ingest the processed content into an index for Retrieval Augmented Generation (RAG) and store extracted fields for downstream automation. Stakeholders must be able to verify where each extracted field value came from in the original PDF and route low-reliability extractions for manual review. You need to ensure that the Content Understanding document analyzer output includes a per-field confidence score and source grounding locations within the source document. What should you do?

- A. Enable estimateFieldSourceAndConfidence.
- B. Configure the analyzer to use generative extraction for all fields.
- C. Set enableSegment to true.
- D. Provide labeled samples.

Answer: A

NEW QUESTION 7

HOTSPOT - (Topic 2)

You have a Microsoft Foundry project that contains an agent named PaymentAgent.

PaymentAgent includes a function tool that issues customer refunds by using an external API.

You are creating a workflow in YAML.

You need to ensure that the workflow pauses for human approval and continues with the refund step only after approval is granted.

How should you complete the workflow definition? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

steps:

- id: propose_refund

type: agent

agent: PaymentAgent

- id: approval

type:

ask_question
basic_chat
data_transformation

- id: execute_refund

type: agent

agent: PaymentAgent

condition:

approval == "approved"
propose_refund.output != null
true

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

steps:

- id: propose_refund

type: agent

agent: PaymentAgent

- id: approval

type:

▼

ask_question

basic_chat

data_transformation

- id: execute_refund

type: agent

agent: PaymentAgent

condition:

▼

approval == "approved"

propose_refund.output != null

true

NEW QUESTION 8

DRAG DROP - (Topic 2)

You have an app that uses Azure AI Custom Vision and a custom trained classifier to identify products in images. You need to add new products to the classifier.

The solution must meet the following requirements:

- Minimize how long it takes to add the products.
- Minimize development effort.

Which five 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

⋮

From the Azure Machine Learning studio, open the workspace.

⋮

From the Custom Vision portal, open the project.

⋮

Label the sample images.

⋮

Upload sample images of the new products.

⋮

Publish the model.

⋮

From Vision Studio, open the project.

⋮

Retrain the model.

Answer Area

A. Mastered

B. Not Mastered

Answer: A

Explanation:

Actions

From the Azure Machine Learning studio, open the workspace.

From the Custom Vision portal, open the project.

Label the sample images.

Upload sample images of the new products.

Publish the model.

From Vision Studio, open the project.

Retrain the model.

Answer Area

From Vision Studio, open the project.

Upload sample images of the new products.

Label the sample images.

Retrain the model.

Publish the model.

NEW QUESTION 9

- (Topic 2)

You have an Azure Speech in Foundry Tools resource that hosts a custom speech to text model deployed to a custom endpoint. An agent uses the endpoint to perform real-time speech recognition.
You are approaching the expiration date of the custom speech to text model. What is the expected behavior when the model expires?

- A. Speech recognition requests will fall back to the most recent base model for the same locale.
- B. Speech recognition requests will continue to use the expired custom model until the model is removed manually.
- C. Speech recognition requests will return a 4xx error until a new custom model is deployed.
- D. The custom model will be deleted automatically when the model expires.

Answer: A

NEW QUESTION 10

- (Topic 2)

You have a Microsoft Foundry project that contains a high-traffic agent. After a recent update, operational costs increase significantly. Monitoring confirms that the volume of user traffic to the agent remains unchanged.
You suspect that changes to the request or response characteristics are causing the increase.
You need to identify whether the additional costs are driven by the model input size, the model output size, or expanded tool usage.
Which observability capability should you use?

- A. evaluation metrics
- B. token usage
- C. latency
- D. run success rate

Answer: B

NEW QUESTION 10

DRAG DROP - (Topic 2)

You have a Microsoft Foundry project that processes procurement documents submitted by suppliers.
You need to implement two pipelines by using Azure Content Understanding in Foundry Tools. The solution must meet the following requirements:

- Include a pipeline named Pipeline1 that supports cost-effective, high-volume processing of standalone PDF invoices.
- Include a pipeline named Pipeline2 that supports cross-document validation by using multi-step reasoning and reference data.

How should you configure each pipeline? To answer, drag the appropriate configurations to the correct pipelines. Each configuration 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.

Configurations

Multi-file task in pro mode

Multi-file task in standard mode

Single-file task in pro mode

Single-file task in standard mode

Answer Area

Pipeline1: Configuration

Pipeline2: Configuration

A. Mastered

B. Not Mastered

Answer: A

Explanation:

Configurations

Multi-file task in pro mode

Multi-file task in standard mode

Single-file task in pro mode

Single-file task in standard mode

Answer Area

Pipeline1

Single-file task in standard mode

Pipeline2

Multi-file task in pro mode

NEW QUESTION 13

- (Topic 2)

You have a Microsoft Foundry project that generates product marketing images from text prompts. After publishing several images, the legal team at your company identifies a competitor's logo on a sign in the background of an image. You need to remove only the logo, while preserving the rest of the image. What should you do?

- A. Increase the prompt guidance strength.
- B. Modify the original prompt to exclude brand names.
- C. Apply a mask-based inpainting edit to the part of the image that contains the logo.
- D. Rerun the prompt by using a different random seed.

Answer: C

NEW QUESTION 16

HOTSPOT - (Topic 2)

You have a Microsoft Foundry project that contains a customer support agent built by using the Foundry Agent Service. The agent uploads user-provided screenshots to Azure Storage through a ticketing tool and receives a blob URL for additional reasoning. You need to use image moderation during agent runs and prevent harmful content from being returned during runs. Azure AI Content Safety must access the images by using the blob URL. The solution must follow the principle of least privilege. What should you configure for Content Safety? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.

Guardrails:

Select Tool call and set Action to Block.

Select User input and Output and set Action to Annotate.

Select User input and Tool response and set Action to Annotate.

Select User input, Output, Tool response, and Tool call and set Action to Block.

Storage access:

Storage account access keys

A user-assigned identity that is assigned the Storage Queue Data Contributor role

A system-assigned managed identity that is assigned the Storage Blob Data Reader role

A system-assigned managed identity that is assigned the Storage Blob Data Contributor role

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Guardrails:

Select Tool call and set Action to Block.

Select User input and Output and set Action to Annotate.

Select User input and Tool response and set Action to Annotate.

Select User input, Output, Tool response, and Tool call and set Action to Block.

Storage access:

Storage account access keys

A user-assigned identity that is assigned the Storage Queue Data Contributor role

A system-assigned managed identity that is assigned the Storage Blob Data Reader role

A system-assigned managed identity that is assigned the Storage Blob Data Contributor role

NEW QUESTION 17

- (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 18

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 23

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

The Leader of IT Certification

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

NEW QUESTION 25

- (Topic 2)

You have a customer support agent built by using the Microsoft Foundry Agent Service. The agent calls an Azure OpenAI model deployment. During load testing, calls intermittently fail and return an HTTP 429 rate limit exceeded error. You need to handle throttling to reduce call failures and improve reliability under load. The solution must remain within the service and model limits. What should you do?

A. Implement a retry policy that uses exponential backoff and jitter.

B. Create a new thread and retry the calls immediately.

C. Reduce the number of registered tools.

D. Split uploaded content into smaller files.

Answer: A

NEW QUESTION 28

- (Topic 2)

You have a Microsoft Foundry project that contains an agent. The agent uses two tools to perform the following actions:

? Use Azure AI Search to retrieve answers from a private product documentation index.

? Use the web search tool to retrieve public information on the internet.

You need to ensure that for a specific run, the agent deterministically retrieves information only from the internet. To what should you set tool_choice?

A. "required"

B. {"type": "azure_ai_search"}

C. "auto"

D. {"type": "bing_grounding"}

Answer: D

NEW QUESTION 31

- (Topic 2)

You are creating an agent workflow in a Microsoft Foundry project to support natural voice interactions. The agent must receive continuous audio input, convert the input into text for reasoning, and then return spoken responses to a user. The workflow must meet the following requirements:

- . Support turn-taking dynamics, where the agent begins to generate the speech output before the user finishes speaking.
- . Operate with low latency to maintain a conversational experience.

You need to enable both speech to text and text to speech in a real-time agent interaction. What should you do?

A. Use an embeddings model to encode the audio, and then decode the audio into text and speech.

B. Use batch transcription to convert the audio input and return text responses from the agent.

C. Use speech translation to convert the audio into another language and return the translated text.

D. Use real-time speech to text for incoming audio and text to speech for agent responses.

Answer: D

NEW QUESTION 36

HOTSPOT - (Topic 2)

Your company is piloting a customer support agent in a Microsoft Foundry project name Project1. Project1 is connected to an existing Application Insights resource, and the company's support team reviews runs in the Traces tab. The Foundry Agent Service is configured to perform the following actions:

- Retrieve the Application Insights connection string by calling project_client.telemetry.get_application_insights_connection_string().
- Call configure_azure_monitor(connection_string=...) to enable telemetry.

A separate LangChain service configured to use OpenTelemetry and has the following configurations:

- Uses AzureAIOpenTelemetryTracer(connection_string=..., enable_content_recording=False)
- Passes the tracer by using config={"callbacks":[azure_tracer]}

Company policy has the following requirements:

- Telemetry from LangChain and OpenTelemetry must be distinguishable within the same Application Insights resource.
- Secrets and credentials must NOT be stored in prompts, tool arguments, or span attributes.

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

NOTE: Each correct selection is worth one point.

Answer Area

Statements	Yes	No
The LangChain service will appear in Traces without configuring a tracer.	☐	☐
Setting different OTEL_SERVICE_NAME values separates the services in Application Insights.	☐	☐
When using enable_content_recording=False, prompts and tool data will be captured in the telemetry.	☐	☐

A. Mastered

B. Not Mastered

Answer: A

Explanation:

Answer Area		
Statements	Yes	No
The LangChain service will appear in Traces without configuring a tracer.	☐	☒
Setting different OTEL_SERVICE_NAME values separates the services in Application Insights.	☒	☐
When using enable_content_recording=False, prompts and tool data will be captured in the telemetry.	☐	☒

NEW QUESTION 37

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 38

- (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 image moderation to block unsafe content before processing the images.
Does this meet the goal?

- A. Yes
- B. No

Answer: B

NEW QUESTION 39

- (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 43

HOTSPOT - (Topic 2)

You have a Python application named App1 that integrates with a Microsoft Foundry project named Project1.

You need to ensure that App1 meets the following requirements:

- Authenticates by using a Microsoft Entra managed identity
- Sends prompts to a deployed model by using the Azure OpenAI Responses API

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.

```
from azure.identity import DefaultAzureCredential

from azure.ai.projects import AIProjectClient

credential =  ()
    AzureKeyCredential
    ClientSecretCredential
    DefaultAzureCredential

project_client = AIProjectClient(
    endpoint="https://contosoai.services.ai.azure.com/api/projects/project1",
    credential=credential,
)

with project_client.get_openai_client() as openai_client:
    response = openai_client.responses.  (
        compact
        create
        retrieve

    model="trail-guide-chat",
    input="Create a 3-day hiking itinerary near Seattle.",
)

print(response.output_text)
```

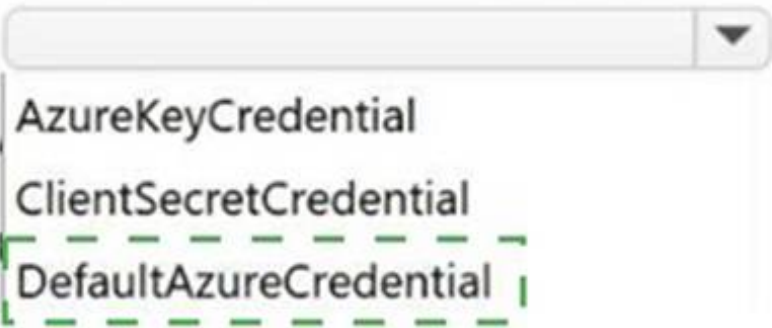
- A. Mastered
- B. Not Mastered

Answer: A

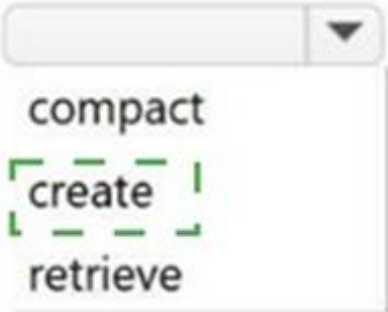
Explanation:

```
from azure.identity import DefaultAzureCredential

from azure.ai.projects import AIProjectClient

credential =  ()

project_client = AIProjectClient(
    endpoint="https://contosoai.services.ai.azure.com/api/projects/project1",
    credential=credential,
)

with project_client.get_openai_client() as openai_client:
    response = openai_client.responses. (

        model="trail-guide-chat",
        input="Create a 3-day hiking itinerary near Seattle.",
    )

    print(response.output_text)
```

NEW QUESTION 44

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

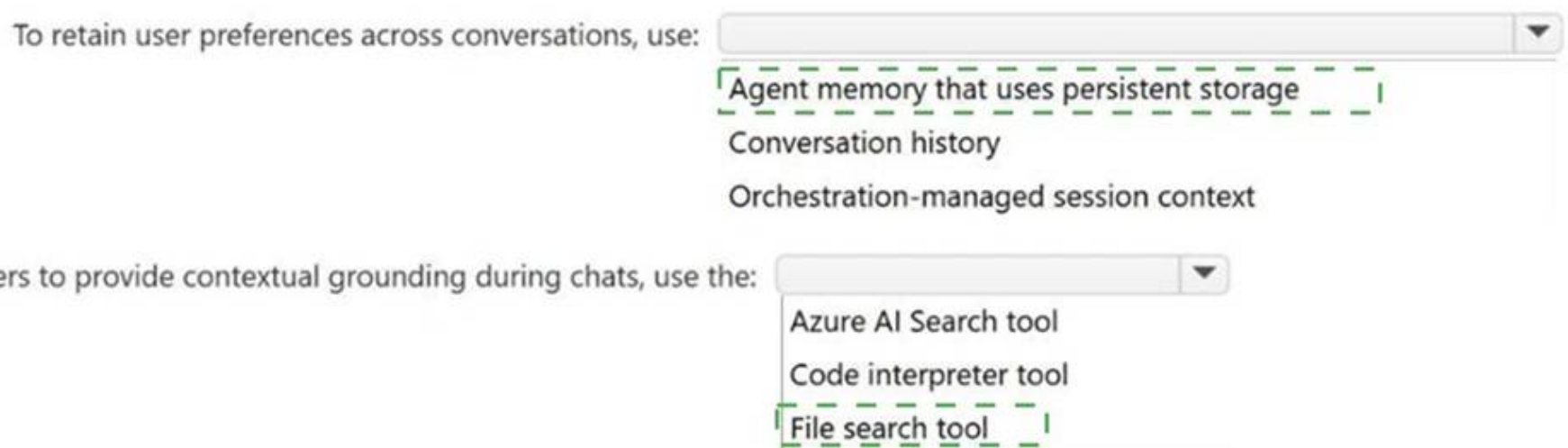
Code interpreter tool

File search tool

B. Not Mastered

Answer: A

Explanation:



NEW QUESTION 47

- (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 a prompt shield for documents. Does this meet the goal?

A. Yes

B. No

Answer: B

NEW QUESTION 51

- (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 53

- (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 Microsoft Foundry project that contains an agent. The agent generates summaries from retrieved policy documents.

Users report that some responses omit required regulatory clauses, even when the clauses are present in the retrieved content.

You need to improve response completeness.

Solution: You run an evaluation flow that scores responses for completeness and blocks responses that fall below a defined threshold.

Does this meet the goal?

A. Yes

B. No

Answer: B

NEW QUESTION 56

- (Topic 2)

You have a Microsoft Foundry project that contains a customer support agent built on a deployed chat model. The agent responses are validated by using an automated testing system that compares generated answers to stored expected outputs. Identical prompts must return consistent responses to prevent automated test failures. You need to reduce response variability, without modifying the prompt or reducing factual accuracy.

A. Increase the max_tokens parameter.

B. Remove stop sequences from the requests.

- C. Increase the temperature parameter.
- D. Decrease the temperature parameter.

Answer: D

NEW QUESTION 59

- (Topic 2)

You have a Microsoft Foundry project that contains an agent. The agent uses Azure Speech in Foundry Tools.

You fine-tune a baseline speech to text model for the en-us locale and publish the model. The agent calls the Speech to text REST API and returns an error message indicating that the project ID is invalid.

You need to set the project property to the correct ID. To what should you set the project property?

- A. the custom speech endpoint URL
- B. the project URL
- C. the project ID
- D. the custom speech project ID

Answer: D

NEW QUESTION 60

- (Topic 2)

You are building a web app named App1 that generates responses by using a model deployed to a Microsoft Foundry project named Project1.

Before sending the prompts to the model, App1 must retrieve documents by using Azure AI Search.

You need to integrate Project1 and App1. The solution must meet the following requirements:

- Multiple client applications must use the same search configuration.
- A security policy must prevent key-based authentication.
- Administrative effort must be minimized. What should you do?

- A. Enable a managed identity for each application and call Azure AI Search directly.
- B. Create a custom HTTP connection in Foundry and manually configure Azure AI Search endpoints per application.
- C. Call Azure AI Search directly from each application by using Microsoft Entra authentication.
- D. Configure an Azure AI Search connection in Project1 and reference the connection in each application.

Answer: D

NEW QUESTION 61

- (Topic 2)

You plan to configure an evaluation in Microsoft Foundry for a Retrieval Augmented Generation (RAG) chat app. You need to provide scores for groundedness, relevance, and harmful-content categories. Which two evaluation categories can you use? Each correct answer presents part of the solution.

- A. risk and safety metrics
- B. AI quality (NLP) metrics
- C. AI quality (AI assisted) metrics
- D. fluency evaluator
- E. similarity evaluators

Answer: AC

NEW QUESTION 63

- (Topic 2)

You have a Microsoft Foundry project. You need to deploy a model from the model catalog to support real-time inference. The solution must meet the following requirements:

? Use key-based authentication

? Support real-time REST API access

? Not consume the vCPU quota of the virtual machines in the Azure subscription

Which type of deployment should you use?

- A. serverless API
- B. self-hosted container
- C. standard
- D. batch

Answer: A

NEW QUESTION 64

- (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 68

HOTSPOT - (Topic 2)

You have a configurable guardrail in a Microsoft Foundry project.

You have a Python application. A text blocklist named ConfidentialTerms already contains terms that must be detected in user messages. The application stores each incoming user message in a variable named input_text.

You plan to moderate input_text before the text is sent to a language model. You need to analyze input_text by using the existing blocklist.

Answer Area

```
...
blocklist_name = "ConfidentialTerms"
client = ContentSafetyClient(endpoint, credential)
request =
    text=input_
    .. ..
    (
        AddOrUpdateTextBlocklistItemsOptions
        AnalyzeImageOptions
        AnalyzeTextOptions
        TextBlocklist
    )
```

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area

```
...
blocklist_name = "ConfidentialTerms"
client = ContentSafetyClient(endpoint, credential)
request =
    text=input_
    .. ..
    (
        AddOrUpdateTextBlocklistItemsOptions
        AnalyzeImageOptions
        AnalyzeTextOptions
        TextBlocklist
    )
```

NEW QUESTION 70

- (Topic 2)

You are building a speech processing solution in Microsoft Foundry for a customer support platform.

The platform will transcribe live phone calls, so that supervisors at your company can view call transcripts and detect issues while the calls are in progress. The call audio will arrive as a continuous stream from the telephony system.

You need to ensure that the call transcripts appear within only a few seconds of the audio stream.

What should you do?

- A. Run a batch transcription job on recorded audio files.
- B. Use real-time speech to text to process streaming audio input.
- C. Use speech translation to generate the transcripts into multiple languages.
- D. Use text to speech by using a custom neural voice.

Answer: B

NEW QUESTION 75

- (Topic 2)

In Microsoft Foundry, you use the Chat playground with the GPT-35 Turbo model. You have a prompt that contains the following code.

```
function F(n)
{
    var f = [0, 1];
    for (var i = 2; i < n; i++) f[i] = f[i-1] + f[i-2];
    return f;
}
```

You need the model to create an explanation of the code. The solution must minimize costs. What should you do?

- A. Add function F (explanation) to the prompt.
- B. Change the model to GPT-4-32k.
- C. Add// what does function F do? to the prompt.
- D. Set the temperature parameter to 1.

Answer: C

NEW QUESTION 80

- (Topic 2)

You have a Microsoft Foundry project that ingests scanned PDF invoices stored in Azure Blob Storage. Each invoice contains printed line items and has a table-based layout.

Extracted results are stored as structured JSON and used as grounding data for an agent in a Retrieval Augmented Generation (RAG) solution.

You need to create a single analyzer that meets the following requirements:

- Extracts the invoice number, invoice date, vendor name, and total amount across varying templates
- Returns confidence scores so that results with confidence below 0.80 can be routed for supervisor review

What should you use?

- A. the Azure Content Understanding in Foundry Tools prebuilt-layout analyzer
- B. a Foundry agent that has groundedness guardrails enabled to extract invoice fields and confidence scores
- C. a custom Azure Content Understanding in Foundry Tools analyzer that defines the required fields as the extracted fields and the returned confidence scores for routing
- D. the Azure Content Understanding in Foundry Tools prebuilt-documentSearch analyzer and search.score from the Azure AI Search results for routing

Answer: C

NEW QUESTION 82

- (Topic 2)

You have an Azure subscription that contains an Azure Language in Foundry Tools service resource. You need to identify the URL of the REST interface for the Language service. Which blade should you use in the Azure portal?

- A. Networking
- B. Keys and Endpoint
- C. Identity
- D. Properties

Answer: B

NEW QUESTION 86

- (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 91

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 93

- (Topic 2)

You need to measure the public perception of your brand on social media by using natural language processing. Which Azure service should you use?

- A. Azure Document Intelligence in Foundry Tools
- B. Content Safety in Foundry Control Plane
- C. Azure Language in Foundry Tools
- D. Azure Vision in Foundry Tools

Answer: C

NEW QUESTION 96

- (Topic 2)

You have a Microsoft Foundry project that contains a customer support agent. The agent calls an internal knowledge API tool before generating responses.

Users report the following issues:

- Some requests take more than 15 seconds to complete.
- Some responses are incorrect, even when the knowledge API returns the expected data. You need to inspect individual agent runs to view the ordered sequence of large language model (LLM) calls, tool invocations, and timing information. Which observability capability should you use?

- A. token usage
- B. safety metrics
- C. tracing
- D. monitoring

Answer: C

NEW QUESTION 98

HOTSPOT - (Topic 2)

You have a Microsoft Foundry project that contains a deployed chat model.

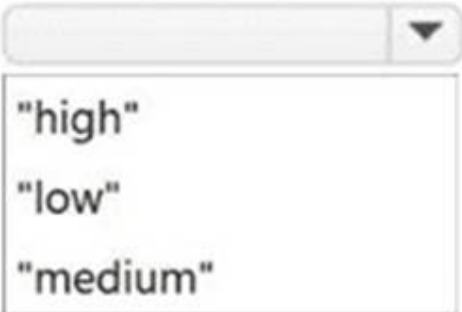
You have a Python service that sends API requests to the model. The service is integrated with an automated validation system that compares generated outputs against approved response patterns.

Stakeholders report that small wording differences are causing validation mismatches.

You need to update the request parameters to improve output stability. The solution must maximize reasoning quality.

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.

```
message = client.messages.create(  
    model="deployment-name",  
    messages=[  
        {"role": "user", "content": "Summarize the release notes in 3 bullet points."}  
    ],  
    max_tokens=800,  
    temperature= ,  
    thinking={"type": "enabled"},  
    output_config={"effort":  }  
)
```

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

```
message = client.messages.create(  
    model="deployment-name",  
    messages=[  
        {"role": "user", "content": "Summarize the release notes in 3 bullet points."}  
    ],  
    max_tokens=800,  
    temperature=  


0



1



2

 ,  
    thinking={"type": "enabled"},  
    output_config={"effort":  


"high"



"low"



"medium"

 }  
)
```

NEW QUESTION 101

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 106

- (Topic 2)
You have a Microsoft Foundry project named Project1 that contains an agent. The agent uses an OpenAPI 3.0 specification to call an external weather service. The weather service requires a key to be passed in an HTTP header. The key value is stored as a connection in Project1. You need to ensure that the key value from the connection is included automatically whenever the OpenAPI tool is invoked. What should you configure in the OpenAPI specification?

- A. an Azure Key Vault connection
- B. a header parameter defined for each operation
- C. an API key security scheme
- D. a Bearer token security scheme

Answer: C

NEW QUESTION 107

HOTSPOT - (Topic 2)
You have a Microsoft Foundry project that contains an agent. The agent accepts user-uploaded screenshots and uses a multimodal chat model. Some screenshots contain potentially malicious embedded text. You need to prevent a prompt injection attack and ensure that third-party content is treated as lower trust. How should you configure prompt shields for document attacks? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.

Prompt shields action:

Disable the shield.

Set action to block.

Set action to annotate.

Additional mitigation:

Enable Spotlighting.

Create a custom blocklist.

Use optical character recognition (OCR) to extract the text from the images first.

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Prompt shields action:

▼

Disable the shield.

Set action to block.

Set action to annotate.

Additional mitigation:

▼

Enable Spotlighting.

Create a custom blocklist.

Use optical character recognition (OCR) to extract the text from the images first.

NEW QUESTION 109

.....

Thank You for Trying Our Product

* 100% Pass or Money Back

All our products come with a 90-day Money Back Guarantee.

* One year free update

You can enjoy free update one year. 24x7 online support.

* Trusted by Millions

We currently serve more than 30,000,000 customers.

* Shop Securely

All transactions are protected by VeriSign!

100% Pass Your AI-103 Exam with Our Prep Materials Via below:

<https://www.certleader.com/AI-103-dumps.html>