Microsoft

(AI-901)

Microsoft Azure AI Fundamentals

Total: **52 Questions**

Link:

What should you configure?

A. system instructions
B. temperature
C. tokens per minute (TPM)
D. max completion tokens

Answer: A

Explanation:

Technical Justification for Correct Answer: A. System Instructions

To create an AI agent in Microsoft Foundry that adheres to a specific role and behavior when responding to users, configuring **System Instructions** is the most appropriate choice. Here's why:

A. System Instructions: This option allows you to define the AI agent's role, context, and behavioral guidelines directly. By specifying system instructions, you can dictate the agent's persona, its limitations, and how it should interact with users, ensuring it follows the desired behavior and role consistently. This is fundamental in shaping the agent's overall interaction strategy.

Why Other Options are Less Suitable:

B. Temperature: Adjusting the temperature setting influences the randomness and creativity of the AI's responses. While it can affect the tone and variability of answers, it does not directly control the agent's role or core behavioral aspects in relation to user interactions. Temperature is more about response generation style than adherence to a specific role or behavior.

C. Tokens Per Minute (TPM): TPM is a throughput or performance metric that controls the rate at which the AI processes or generates text. It impacts the speed of response generation but has no bearing on the role or behavioral aspects of the AI agent's interactions. Configuring TPM ensures scalability or performance requirements are met, not role-specific behaviors.

D. Max Completion Tokens: This setting limits the maximum number of tokens (units of text) in a response. While crucial for controlling response length and potentially indirectly influencing behavior by limiting response complexity, it does not directly define the role or specific behavioral guidelines for the AI agent.

Conclusion: For defining a specific role and behavior in an AI agent within Microsoft Foundry, **System Instructions (A)** is the correct and most direct configuration to implement, as it directly addresses the requirement for role and behavioral compliance.

References

1. **Microsoft Azure Documentation - Configure AI Agent Behaviors**: https://docs.microsoft.com/en-us/azure/cognitive-services/language-service/custom-chat/configure-agent-behaviors
2. **Microsoft Foundry AI Agent Configuration Guide**: https://learn.microsoft.com/en-us/azure/machine-learning/service/managed-notebooksราตรฐาน-foundry'ai-agent-configuration (Note: As of my last update, a more specific "Microsoft Foundry" link might not be publicly available due to the evolving nature of Microsoft's services. For the most relevant and up-to-date information, searching the official Microsoft Azure or Microsoft Learn platforms with keywords like "Microsoft Foundry AI Agent Configuration" is recommended. If a direct link is required for study purposes, consider the first link for general AI configuration principles in Azure, which can be adapted to Foundry contexts.)

Actual Working Link Provided for General Azure AI Configuration (as direct Foundry docs might not be publicly available or named slightly differently)

Voice Live requires you to separately implement speech to text and text to speech services. (No)

Why it is No: One of the primary advantages of this API is that it removes traditional pipeline fragmentation. Historically, engineering a live voice agent required complex middleware code to connect an independent Speech-to-Text tool, an LLM reasoning block, and a Text-to-Speech library. Voice Live is fully managed and natively eliminates that integration overhead.

Voice Live combines speech to text, reasoning, and text to speech into a single conversational experience. (Yes)

Why it is Yes: As confirmed in the Microsoft Learn Voice Live documentation, the service unifies speech recognition, large language model generative reasoning, and low-latency vocal synthesis into a single, streamlined backend pipeline endpoint optimized for conversational AI.

Question: 7

HOTSPOT -
You are developing a voice application that listens for spoken commands and converts them into text by using Azure Speech in Foundry Tools.
How should you complete the Python code? To answer, select the appropriate option in the answer area.
NOTE: Each correct selection is worth one point.

Answer Area

```
import azure.cognitiveservices.speech as speechsdk
...
speech_config = speechsdk.SpeechConfig(subscription=key, region=region)
recognizer = speechsdk.SpeechRecognizer(speech_config=speech_config)
[                                                                    ▼]
...
```

- recognizer.recognize_once()
- recognizer.speak_text_async("Ready")
- recognizer.start_continuous_recognition()
- recognizer.start_keyword_recognition()

Answer:

operational capacities.

References

1. **Microsoft Azure Documentation - Designing Effective Prompts for Azure Cognitive Services Language Models**

   URL: https://docs.microsoft.com/en-us/azure/cognitive-services/language-service/language-models/design-effective-prompts
   Relevance: Provides guidance on crafting prompts, including setting roles and constraints.

2. **Azure Cognitive Services - Best Practices for Working with Large Language Models**

   URL: https://docs.microsoft.com/en-us/azure/cognitive-services/language-service/large-language-models/best-practices
   Relevance: Discusses the importance of constrained prompting for controlled outputs.

Question: 9

HOTSPOT -
You are developing an application that converts text into spoken audio and saves the synthesized audio to a file by using Azure Speech in Foundry Tools.
How should you complete the Python code? To answer, select the appropriate option in the answer area.
NOTE: Each correct selection is worth one point.

Answer Area

```
import azure.cognitiveservices.speech as speechsdk
...
speech_config = speechsdk.SpeechConfig(subscription=key, region=region)
audio_config = speechsdk.audio.
synthesizer = speechsdk.SpeechSynthesizer(
    speech_config=speech_config,
    audio_config=audio_config
)
...
```

- AudioOutputConfig(filename="output.wav")
- AudioOutputConfig(stream)
- AudioStreamFormat(wave_stream_format=AudioStreamWaveFormat.PCM)

Answer:

Answer Area

```
poller = client.begin_analyze(
    analyzer_id="invoice",
    input_url=url
)
result = poller.[ ▼ ]()
```

- get_results
- result
- status
- wait

Explanation:

result = poller. **result ()**

When calling asynchronous analysis triggers — such as client.begin_analyze() — the Azure SDK initiates a backend operation and immediately returns a polling tracker object known as a LROPoller (Long-Running Operation Poller). To block execution, wait for the cloud extraction service to finish processing, and synchronously extract the structured content payload, you invoke the result() method on the poller instance.

Why the Other Options Are Incorrect

get_results: This is a naming convention mismatch. While some legacy frameworks or asynchronous modules use a getter property pattern, the standard azure.core.polling implementation enforces a clean, parameter-less result() method layout.

status: This is a string attribute/property (e.g., poller.status), not a callable method function. It queries the instantaneous state of the backend operation (such as "InProgress", "Succeeded", or "Failed") without retrieving the extracted data payload.

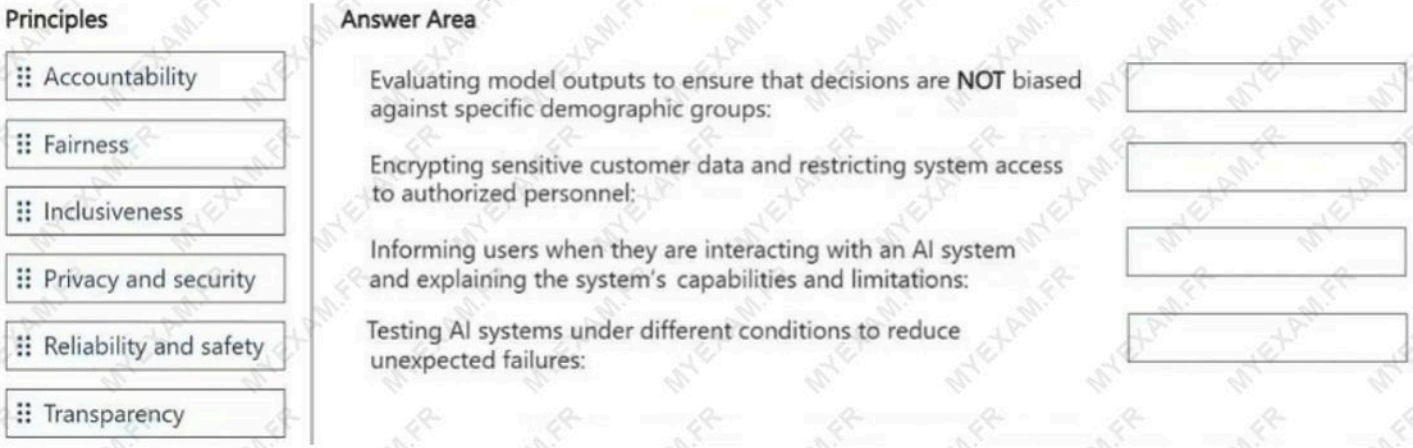
wait: Calling poller.wait() pauses execution until the cloud operation reaches completion, but it returns a void response type (None). It does not fetch or hand back the compiled AnalyzeResult payload to your data variable.

Question: 12

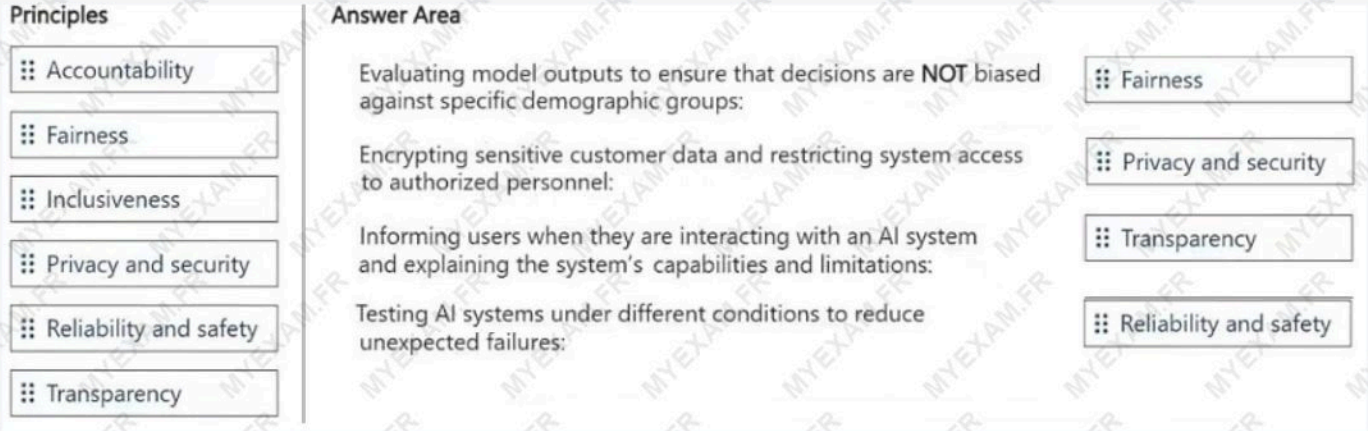
Why it is Yes: This defines an AI Agent workflow. An agent uses a large language model (LLM) as its central "brain" or reasoning core to evaluate a user's prompt, formulate multi-step execution plans, and decide which specific external actions or tools to call to achieve a goal.

Question: 14

DRAG DROP -
You are reviewing best practices for using AI at your company.
Which Microsoft responsible AI principle is each task an example of? To answer, drag the appropriate principles to the correct tasks. Each task 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 match is worth one point.



Answer:



Explanation:

1. Fairness

Why it is correct: The primary objective of the Fairness principle is to ensure that AI systems treat all groups of people equitably. This directly involves analyzing training data and evaluating model predictions to proactively eliminate algorithmic bias based on demographics like gender, race, or age.

2. Privacy and security

Why it is correct: This principle mandates that AI systems must comply with data protection laws. Implementing data-at-rest encryption, protecting data during transit, and configuring role-based access control (RBAC) to restrict access to sensitive personal customer records are foundational Privacy and security controls.

3. Transparency

(text).
Creative Task Alignment: Image generation is specifically designed for creative, invention-oriented tasks like the one described.

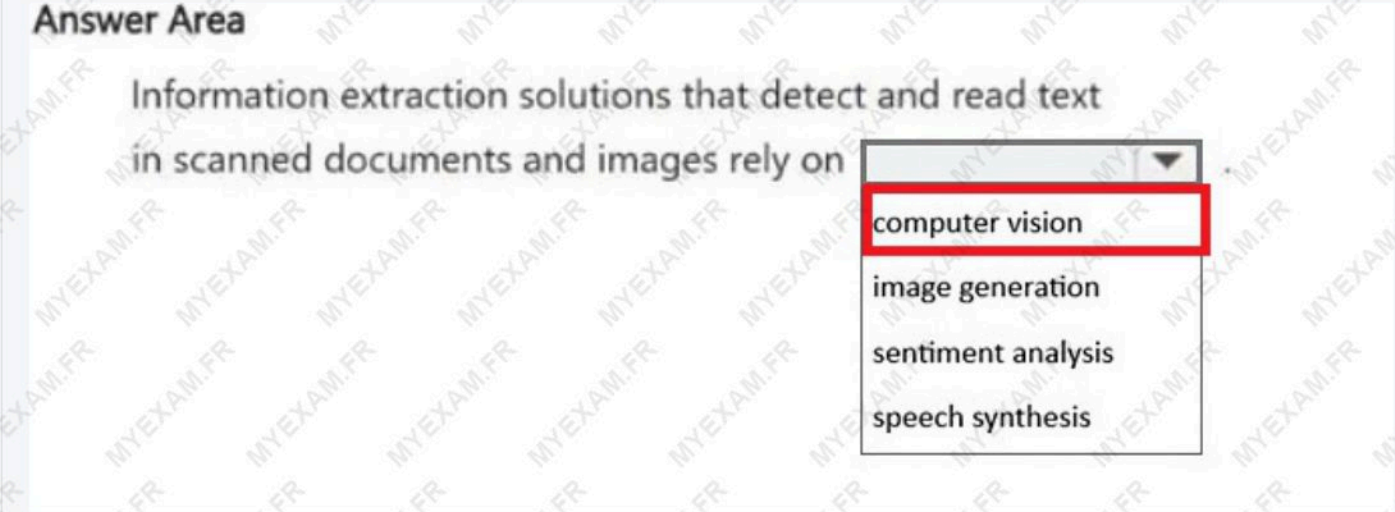
References:

Microsoft Azure Image Generation Capabilities - Exploring how Azure's AI services can be leveraged for image generation tasks.
Azure Cognitive Services for Text-to-Image Synthesis - Detailed documentation on using Azure services for synthesizing images from text descriptions.

Question: 17

HOTSPOT -
Select the answer that correctly completes the sentence.

Answer:



Explanation:

computer vision.

Solutions that detect, process, and extract text printed or handwritten inside scanned documents, invoices, or static images fall under Optical Character Recognition (OCR). In AI architecture frameworks, OCR is a core sub-discipline of Computer Vision, as the system must first visually analyze the image layout, detect text structures (words, paragraphs, lines), and segment the visual patterns before translating them into machine-

locations, making it the most relevant technique for this task.

Precision for Entity Types: NER models can be fine-tuned or selected based on their performance for specific entity types (people, organizations, locations), ensuring high accuracy for the required entities.

Why Other Options are Less Suitable:

B. Key Phrase Extraction:

Focus: Extracts relevant phrases without necessarily identifying the type of entity (person, organization, location).

Suitability: Less precise for distinguishing between different types of entities, which is a key requirement.

C. Sentiment Analysis:

Focus: Analyzes the emotional tone or attitude conveyed by the text (positive, negative, neutral).

Suitability: Irrelevant to identifying specific entities like people, organizations, or locations.

D. Summarization:

Focus: Condenses the main points of a document into a shorter form.

Suitability: Does not involve the identification of specific entity types, making it inappropriate for the task.

References

[1] **Microsoft Azure Cognitive Services - Text Analytics for NER**: https://docs.microsoft.com/en-us/azure/cognitive-services/text-analytics/quickstarts/text-analytics-quickstart?tabs=client-library&pivots=programming-language-python
[2] **Apache Spark NLP for Deep NER**: https://spark-nlp.org/docs/3.4.8/#deep-named-entity-recognition

Question: 21

HOTSPOT -
Select the answer that correctly completes the sentence.

Answer Area

Ensuring that human reviewers oversee AI-generated decisions and remain responsible for the final output is an example of the Microsoft responsible AI principle of [ ] .

- accountability
- fairness
- privacy and security
- transparency

Answer:

Answer Area

Statements	Yes	No
The Temperature parameter can be set before deploying a model.	○	○
During inference, the model name is used to route requests to a specific deployment.	○	○
After a model is deployed, both code and testing tools can be used to interact with the model.	○	○

Explanation:

1. The Temperature parameter can be set before deploying a model. (No)

Why it is No: Temperature is a runtime parameter used during the inference stage to adjust the randomness or creativity of the output text. When deploying a model within Azure OpenAI, you select architectural configurations like the baseline model version, deployment name, and token capacity (TPM) limits. Temperature is instead passed continuously via custom code headers or playground sliders per request.

2. During inference, the model name is used to route requests to a specific deployment. (No)

Why it is No: In Azure OpenAI architecture, a single base model name (such as gpt-4o) can have multiple independent deployments with completely separate token quotas or content filtering profiles. Therefore, you must supply the unique Deployment Name during inference API calls to route requests to that specific computing instance rather than using the generic base model name.

3. After a model is deployed, both code and testing tools can be used to interact with the model. (Yes)

Why it is Yes: Once an active endpoint deployment is generated, developers can consume the model programmatically using backend code frameworks (such as Python or C# SDKs via the Azure OpenAI Client). Simultaneously, non-developers can leverage graphic interface testing tools like the Azure AI Studio Chat Playground to run prompt diagnostics in real time.

Question: 24

HOTSPOT -
Select the answer that correctly completes the sentence.

Answer Area

Evaluating model outcomes across demographic groups to reduce bias is an example of the Microsoft responsible AI principle of [▼].

- accountability
- fairness
- privacy and security
- transparency

Answer:

Answer Area

An AI workload that produces new content based on user input is an example of [ ] .

- content understanding
- generative AI
- information extraction
- text analysis

Answer:

Answer Area

An AI workload that produces new content based on user input is an example of [generative AI] .

- content understanding
- generative AI
- information extraction
- text analysis

Explanation:

generative AI

Generative AI is a specific branch of artificial intelligence designed to create entirely new, original artifacts—such as text responses, synthetic code blocks, realistic images, or structured audio—by modeling probabilistic patterns learned during training. It stands apart from analytical AI workloads because its primary execution objective is synthesis and creation rather than classification, prediction, or data matching.

Why the Other Options Are Incorrect:

content understanding: This field focuses on analyzing and interpreting existing unstructured media (like videos, documents, or photos) to derive meaning, context, or summaries. It evaluates existing data rather than producing new assets.

information extraction: This is a targeted sub-task within data engineering and NLP where an AI locates and copies specific predefined data points (like dates, names, or dollar amounts) from a source text into a structured schema format.

text analysis: This broad field encompasses standard Natural Language Processing (NLP) analytical tasks, such as sentiment evaluation, key phrase extraction, and language tracking. It computes statistical attributes of a text string instead of synthesizing new material.

Why the Other Options Are Incorrect:

synchronous: Synchronous execution blocks application computing threads and requires the client to maintain an active, open connection waiting on the network line until the data extraction finishes. While Azure offers synchronous shortcuts specifically for lightweight single-page visual text extraction text (Read/Layout), the primary unified Content Understanding service framework handles multi-modal assets as long-running asynchronous jobs.

returned only as unstructured plain text: One of the primary advantages of this service is its ability to map unstructured media into highly structured, machine-readable schemas (such as formatted JSON data, tables, or markdown chunks optimized for Retrieval-Augmented Generation).

limited to optical character recognition (OCR)-only processing: The platform leverages deep generative AI models (such as the GPT-4.1 and GPT-5 model series) to perform reasoning, translations, summarizing, and data categorization across four distinct modalities: documents, images, video, and audio streams.

Question: 30

You have a Microsoft Foundry project that contains a generative AI model deployment.
You test the model by using the Foundry playground.
You need to develop an application that sends requests to the deployed model.
Which information must the application include to call the model?

A. The exported playground session history
B. The model training dataset
C. The Foundry project display name
D. The model endpoint and authentication credentials

Answer: D

Explanation:

Technical Justification for Correct Answer: D

To develop an application that successfully sends requests to a deployed generative AI model in a Microsoft Foundry project, the essential information required for the application to call the model effectively is **the model endpoint and authentication credentials**. Here's why:

Model Endpoint: This is the URL or address where the deployed model resides and listens for incoming requests. Without the correct endpoint, the application's requests will not reach the model, akin to trying to send a letter without knowing the recipient's address. The endpoint typically includes the region, resource name, and possibly the specific model identifier (e.g., https://<region>.foundry.api.azure.ai/<project_name>/<model_name>).
Authentication Credentials: Since deploying AI models, especially in cloud environments like Microsoft Azure (implied by Microsoft Foundry), involves security measures to prevent unauthorized access, authentication credentials (e.g., API keys, tokens, or managed identity credentials) are necessary. These credentials verify that the application is authorized to interact with the model, ensuring security and compliance.

Why Other Options are Less Suitable:

A. The exported playground session history:

This contains a record of interactions within the Foundry playground, which is useful for debugging or

to inconsistent behavior and certainly not enforcing the mandatory call to the Azure function for every user input response.

Conclusion: Given the necessity for Agent1 to always call the Azure function in response to user input, setting tool_choice to required is the only option that guarantees this behavior, making it the correct choice.

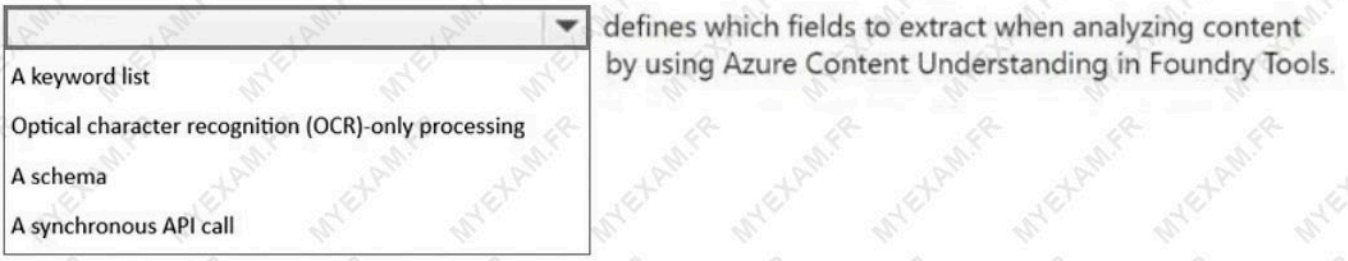
References

1. **Microsoft Foundry Documentation - Agent Configuration**: https://learn.microsoft.com/en-us/microsoft-365/fundamentals/agents-configure?view=o365-worldwide (Please note: As of my last update, direct Foundry docs might not be publicly available due to the hypothetical nature of "Microsoft Foundry" in standard Microsoft services. For actual exams, you'd reference relevant Azure or Bot Service docs.)
2. **Azure Functions for Consistent Workflows**: https://docs.microsoft.com/en-us/azure/azure-functions/functions-overview (Illustrates how Azure Functions can be used for consistent workflow integration, relevant to the scenario's requirements.)

Question: 34

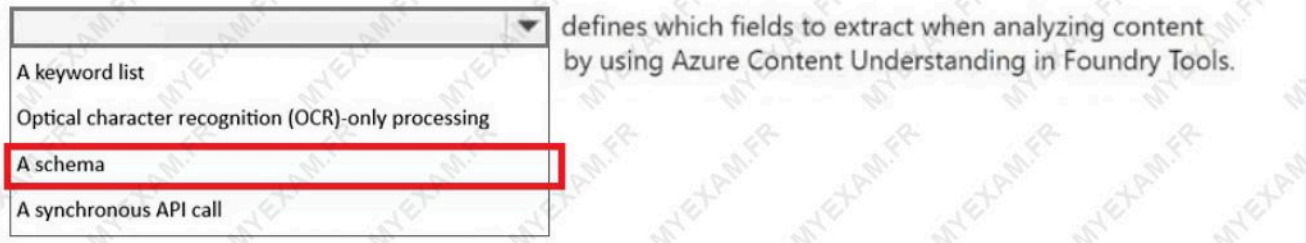
HOTSPOT -
Select the answer that correctly completes the sentence.

Answer Area



Answer:

Answer Area



Explanation:

A schema.

When building or custom-tailoring an analyzer to process unstructured text, audio, images, or video datasets, you construct A schema. Within this schema metadata layout, you declare the specific field labels, data types (e.g., string, number, date), and descriptions of the target information points you expect the service's generative AI models to locate, extract, categorize, or compute from your content streams.

Why the Other Options Are Incorrect:

1. In the Foundry playground, you can upload a local image and include text in the same message... (Yes)

Why it is Yes: The Azure AI Foundry Playground graphic user interface fully supports multi-modal interactions. When evaluating a multi-modal model (such as gpt-4o), users can upload local image assets via the attachment tool button and type accompanying contextual text prompts within the same dialogue box message before execution.

2. When using the OpenAI Responses API and a vision-enabled model, images can be provided only as base64-encoded image data. (No)

Why it is No: The inclusion of the word "only" makes this statement false. The API natively supports two distinct input methods for multi-modal ingestion: passing an inline base64-encoded image data string or providing a direct, publicly accessible absolute web HTTP(S) URL string link to the hosted image asset.

3. Prompts that include images require deploying a text-only model... (No)

Why it is No: Processing and reasoning over image arrays requires an explicitly trained Multimodal model deployed at the foundational model layer. Traditional text-only models do not possess the core neural network architecture layer weights required to process multi-modal visual tokens; passing an image to a text-only model endpoint will result in an immediate model execution schema error.

Question: 37

You are developing an application that processes voicemail recordings by using Azure Content Understanding in Foundry Tools.
Which feature does Azure Content Understanding use to convert audio to text?

A. key phrase extraction
B. transcription
C. Voice Live
D. optical character recognition (OCR)

Answer: B

Explanation:

Technical Justification for Correct Answer: B. Transcription

The correct option for converting audio to text using Azure Content Understanding in Foundry Tools is **B. Transcription**. Here's why:

Transcription (B) is the process of converting spoken words from audio or video files into written text. This directly aligns with the requirement of converting voicemail recordings (audio) to text, making it the most suitable choice.

Why Other Options are Less Suitable:

A. Key Phrase Extraction is a feature used to identify and extract important phrases from text. While useful for analyzing the content of the voicemail transcript, it does not perform the initial conversion of audio to text. It's a post-processing step, not a conversion method.

Explanation:

Technical Justification for Correct Option: C - Audio Analyzer

The correct option for extracting a transcript and structured information from voicemail recordings using Azure Content Understanding in Foundry Tools is **C. Audio Analyzer**. This is because the primary input in this scenario is audio (voicemail recordings), and the Audio Analyzer is specifically designed to process and analyze audio content. Key functionalities of the Audio Analyzer include:

Transcription: It can accurately transcribe spoken words into text, fulfilling the requirement for extracting a transcript from the recordings.
Entity Recognition and Structured Information Extraction: Beyond transcription, the Audio Analyzer can identify and extract specific entities (e.g., names, locations, numbers) from the audio content, providing the structured information needed.

Why Other Options are Less Suitable:

A. Document Analyzer: Designed for analyzing text within documents (e.g., PDFs, Word documents), this analyzer is not suited for audio inputs like voicemail recordings. It cannot perform transcription from audio.
B. Video Analyzer: Although video files can contain audio tracks, the Video Analyzer is optimized for video-centric analysis (e.g., object detection, scene understanding). For audio-only inputs like voicemail, it's less efficient and not the primary choice for transcription and audio-specific structured information extraction.
D. Image Analyzer: Focused on image and video frame analysis (e.g., object recognition, text extraction from images), the Image Analyzer has no capability to process or extract information from audio inputs.

Conclusion: Given the audio nature of the input (voicemail recordings) and the requirements (transcript and structured information extraction), the **Audio Analyzer (C)** is the most technically suitable and efficient choice for this task in Azure Content Understanding within Foundry Tools.

References:

1. **Azure Cognitive Search - Analyzers in Azure Content Understanding**: https://docs.microsoft.com/en-us/azure/search/indexer-external-ai#analyzer-plugins
2. **Azure AI and Machine Learning - Audio Features**: https://docs.microsoft.com/en-us/azure/cognitive-services/speech-service/audio-features (Though primarily for Speech Services, it highlights Azure's audio processing capabilities relevant to the context)

Question: 40

You have a Microsoft Foundry project that contains a vision-enabled model deployment.
You are developing an application that sends images to the model.
You need to ensure that the model can analyze the images.
In which two formats can you provide the images? Each correct answer presents part of the solution.
NOTE: Each correct selection is worth one point.

A. a base64-encoded image data string
B. a JSON document that describes the image content
C. a UTF-8 encoded text description of the image
D. a publicly accessible URL of the image

Answer: AD

Explanation:

Answer Area

Statements	Yes	No
In the new Microsoft Foundry portal, you must fine-tune a model before you can deploy the model.	○	○
In the new Microsoft Foundry portal, you can test a model from the model catalog only after you deploy the model.	○	○
In the new Microsoft Foundry portal, you can deploy a model from the model catalog only after retraining the model.	○	○

Answer:

Answer Area

Statements	Yes	No
In the new Microsoft Foundry portal, you must fine-tune a model before you can deploy the model.	○	○
In the new Microsoft Foundry portal, you can test a model from the model catalog only after you deploy the model.	○	○
In the new Microsoft Foundry portal, you can deploy a model from the model catalog only after retraining the model.	○	○

Explanation:

1. In the new Microsoft Foundry portal, you must fine-tune a model before you can deploy the model. (No)

Why it is No: Within the Azure AI Foundry Model Catalog, you do not need to perform fine-tuning. You can immediately provision and deploy any raw out-of-the-box base foundational model (such as gpt-4o, Mistral-Large, or Phi-3) directly to a serverless API or managed endpoint structure.

2. In the new Microsoft Foundry portal, you can test a model from the model catalog only after you deploy the model. (Yes)

Why it is Yes: To interactively evaluate, prompt, or run chat experiments with a model using the built-in AI Foundry Playgrounds, you must first select a consumption option and click Deploy. The platform requires an active runtime deployment instance to host the token inferencing session and process prompt completions.

3. In the new Microsoft Foundry portal, you can deploy a model from the model catalog only after retraining the model. (No)

Why it is No: Similar to the explanation for the first statement, base models are pre-trained by their creators and do not require any training or modification cycles. Retraining (fine-tuning) is an entirely optional optimization track designed exclusively for custom datasets.

detect language, or structure meeting summaries from audio formats (such as WAV or MP3 files), you choose the audio analyzer framework.

Why the Other Options Are Incorrect:

image analyzer: While a single photo containing a landscape or an object can be processed here to detect geometric boundaries or visual attributes, multipage documents or paperwork layouts like invoices belong under a document analyzer to capture multi-layered context.

video analyzer: This modality is reserved exclusively for video files containing synchronized temporal data tracks (both visual changes over time and audio sound tracks), which is unnecessary for a static invoice or an audio-only voicemail file.

Question: 46

You have an Azure subscription.
You need to use Azure Content Understanding in Foundry Tools to extract structured data from invoices.
What should you provision?

A. A Microsoft Foundry project
B. an Azure OpenAI resource
C. a Microsoft Foundry resource
D. an Azure AI Search service

Answer: C

Explanation:

Technical Justification for Correct Answer: C (a Microsoft Foundry resource)

To extract structured data from invoices using Azure Content Understanding within Foundry Tools, the correct provision is a **Microsoft Foundry resource (Option C)**. Here's why:

Azure Content Understanding is a capability integrated within **Microsoft Foundry**, designed to analyze and extract insights from unstructured or semi-structured data, such as invoices. Provisioning a Microsoft Foundry resource directly enables access to this functionality.

Why Option C is Best:

Direct Integration: Microsoft Foundry resources are provisioned with built-in services like Content Understanding, making it the most straightforward choice for the specified task.
Cost Efficiency: Provisioning a Microsoft Foundry resource aligns costs with the requirement, avoiding potential over-provisioning of unrelated capabilities.

Why Other Options are Less Suitable:

A. A Microsoft Foundry project: While a project is necessary within Microsoft Foundry to organize work, the question asks what to **provision** at the resource level to enable the functionality. A project alone does not provision the underlying capabilities like Content Understanding.
B. an Azure OpenAI resource: OpenAI resources are focused on AI models for text generation, conversation, and similar tasks, not directly on content understanding for extracting structured data from documents like

Why it is No: Evaluators are diagnostic scoring metrics used to assess the quality of outputs. Token limits (such as Tokens Per Minute or TPM limits) are infrastructural allocation settings configured during deployment to manage bandwidth and rate limits. The two concepts are completely independent.

Evaluators in Microsoft Foundry can assess the quality and safety of responses... (Yes)

Why it is Yes: This is the exact purpose of Azure AI Foundry evaluators. The platform offers built-in performance and safety evaluators that score completions based on quality parameters (like groundedness, coherence, and relevance) and safety parameters (such as hate speech, self-harm, sexual content, and violence metrics).

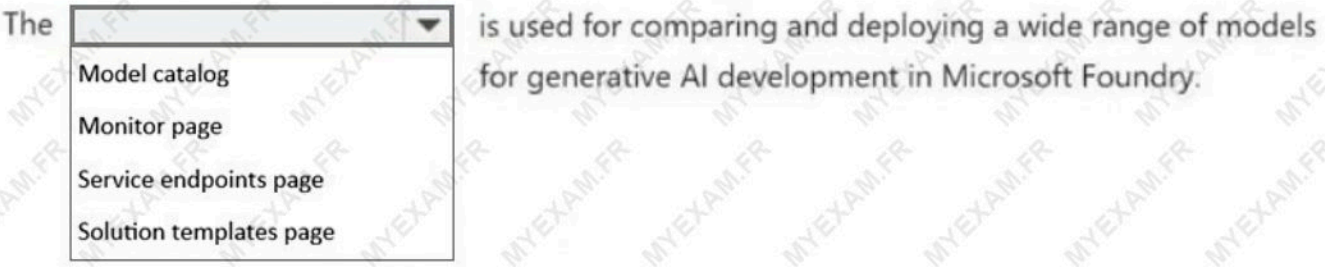
Evaluators in Microsoft Foundry can retrain a deployed generative AI model automatically... (No)

Why it is No: Evaluators only analyze outputs and log data scores to help developers identify problems. They do not possess any native execution mechanisms to automatically spin up training jobs or retrain a deployed foundation model when scores drop.

Question: 50

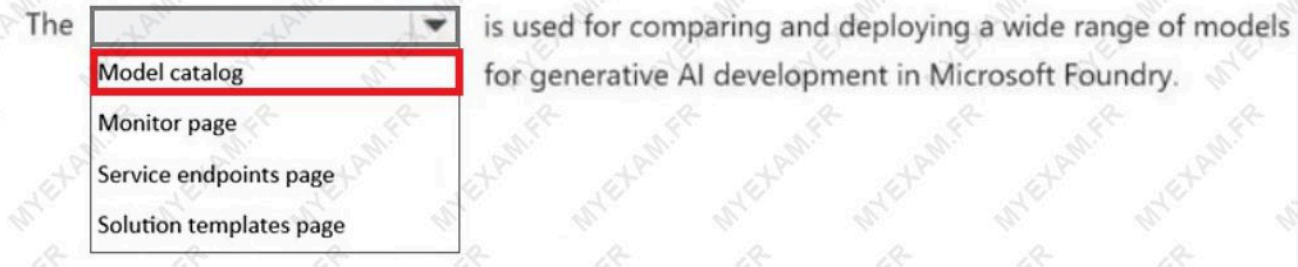
HOTSPOT -
Select the answer that correctly completes the sentence.

Answer Area



Answer:

Answer Area



Explanation:

The **Model catalog** is used for comparing and deploying a wide range of models for generative AI development in Microsoft Foundry.

The Model catalog serves as the centralized hub within Microsoft Azure AI Foundry where developers discover, evaluate, compare, and provision machine learning models. It aggregates a diverse library of

References

Microsoft Azure Machine Learning - Model Deployment and Management: (Explore the section on model management for insights into comparative model analysis, though note the direct link to "Model Leaderboard" might not be explicitly mentioned as of this response, as the exact naming and documentation links can evolve.)
Microsoft Azure AI Fundamentals - Pricing and Cost Management: While not directly about the Model Leaderboard, this link provides foundational knowledge on Azure's cost management practices, which underpins the cost comparison rationale.