

Microsoft prints these exactly. So does your answer.

Every name checked against Microsoft Learn — latest check 15 September 2026 · 444 names · sessions 1–40 of 40

IN THIS SHEET

PAGE

A · Plan and manage an Azure AI solution	25–30% of the exam · 138 names	1
B · Implement generative AI and agentic solutions	30–35% of the exam · 140 names	7
C · Implement computer vision solutions	10–15% of the exam · 57 names	13
D · Implement text analysis solutions	10–15% of the exam · 63 names	15
E · Implement information extraction solutions	10–15% of the exam · 43 names	18
Across the course · review and capstone	session 40 · 3 names	20

A · Plan and manage an Azure AI solution

25–30% of the exam

S1 · Choosing the model: LLMs, small language models, multimodal models, Foundry Tools

Microsoft Foundry	SERVICE	Microsoft's platform that brings agents, models and tools together under one management grouping.
Foundry Tools	SERVICE	Prebuilt, customisable APIs and models for tasks such as speech, translation, language and vision.
Azure Speech in Foundry Tools	SERVICE	The Foundry Tool for speech to text, text to speech and speech translation. · also S37
Azure Translator in Foundry Tools	SERVICE	The Foundry Tool for neural machine translation of text. · also S36
small language models	MODEL	Compact generative models for narrow tasks where cost and speed matter.
multimodal models	MODEL	Models that accept more than one type of input, such as text and images.
model catalog	PORTAL	The place in the Foundry portal to discover, compare and deploy models.
Foundry Models sold by Azure	CONCEPT	Catalog models that Microsoft sells and supports directly.
Foundry Models from partners and community	CONCEPT	Catalog models from other providers; some need an Azure Marketplace subscription.
Model leaderboard	PORTAL	The portal page ranking models on quality, safety, throughput and estimated cost.
Quality index	CONCEPT	A leaderboard score from zero to one summarising benchmark quality; higher is better.
Budgets	PORTAL	The Cost Management menu item where you create a spending budget and its alerts.
Forecasted alerts	SETTING	Budget alerts that fire when projected spending is likely to exceed a threshold.

S2 · Choosing the Foundry service, and the retrieval and indexing method

Foundry Agent Service	SERVICE	The managed platform for building, deploying and scaling AI agents. · also S39
Foundry IQ	SERVICE	Knowledge bases for agents, built on Azure AI Search and agentic retrieval, with citations and permissions. · also S39
knowledge base	CONCEPT	In Foundry IQ, a set of knowledge sources plus retrieval settings that several agents can share.

A · Plan and manage an Azure AI solution · continued

S2 · Choosing the Foundry service, and the retrieval and indexing method

Agentic retrieval	FEATURE	An LLM-assisted retrieval pipeline that plans queries across sources and returns structured responses for agents. · also S20, S39
Hybrid search	FEATURE	One request running full-text and vector queries together, merged into one ranked result set. · also S20
Reciprocal Rank Fusion (RRF)	CONCEPT	The method hybrid search uses to merge the keyword and vector result lists.
Semantic ranker	FEATURE	A secondary ranking pass that re-orders results by meaning and can return captions and answers. · also S20, S39
Vector search	FEATURE	Retrieval by similarity between numeric embeddings of the query and the content.
Integrated vectorization	FEATURE	Chunking and embedding inside the indexer pipeline, plus vectorising the query at query time.
AzureOpenAIEmbedding skill	FEATURE	The indexer skill that calls an Azure OpenAI embedding model during indexing.
Azure Content Understanding in Foundry Tools analyzer	SERVICE	The Foundry Tool that turns documents, images, audio and video into Markdown or schema-shaped output. · also S39
Microsoft.CognitiveServices/accounts/projects	CONCEPT	In Content Understanding, the component that defines extraction settings, field schema and model deployments.
Foundry Account Owner	ROLE	The Azure resource type of a Foundry project, a child of the Foundry resource.
		A built-in role that can create Foundry resources and projects and manage account-level resources.

S3 · Designing the Azure infrastructure, and choosing deployment options

Microsoft.CognitiveServices/accounts	CONCEPT	The Azure resource type of a Foundry resource, with kind AIServices.
standard agent setup	SETTING	Agent setup where you bring your own Azure resources for all customer data, including files, conversations and vector stores. · also S20
Azure Cosmos DB for NoSQL	SERVICE	In Standard setup, stores messages, conversation history and agent metadata.
Foundry User	ROLE	Least-privilege built-in role for developers building and testing in a Foundry project; formerly Azure AI User. · also S5, S10, S20, S39
Foundry Project Manager	ROLE	Built-in role to manage a Foundry project and conditionally assign Foundry User to others.
Selected networks	SETTING	Network access option that admits only specific IP addresses and virtual networks.
Connected resources	PORTAL	The Project details tab listing the Azure resources connected to a Foundry project. · also S10
Serverless API	FEATURE	The preferred deployment option, billed on token usage or provisioned throughput units. · also S20, S39
Managed compute	FEATURE	Deployment option on dedicated accelerators, billed hourly per accelerator SKU. · also S20
Global Standard	TIER	Pay-per-token deployment type processed in any Azure region; the recommended starting point. · also S20, S39
Data Zone Standard	TIER	Pay-per-token deployment type processed only within a US, EU or Asia Pacific data zone. · also S20, S39

A · Plan and manage an Azure AI solution · continued

S3 · Designing the Azure infrastructure, and choosing deployment options

Regional Provisioned	TIER	Reserved-throughput deployment type (SKU ProvisionedManaged) processed within the Azure geography.
Global Batch	TIER	Asynchronous deployment type with separate quota and a 24-hour target turnaround.
az cognitiveservices account list-models	CLI	Lists the models, versions and SKUs a Foundry resource can deploy.

S4 · Hands-on: configuring model and agent deployments

Custom settings	PORTAL	The Deploy option that shows every deployment setting instead of the defaults.
az cognitiveservices account deployment create	CLI	Azure CLI command that creates a model deployment on a Foundry resource.
Cognitive Services Contributor	ROLE	Role Learn lists for creating and managing deployments with the CLI or Bicep.
Version update policy	SETTING	The deployment setting that controls automatic model version upgrades.
OnceNewDefaultVersionAvailable	SETTING	Upgrade policy: move to the new default version within two weeks of it being designated.
NoAutoUpgrade	SETTING	Upgrade policy: never upgrade; the deployment stops working at retirement.
Tokens-per-Minute (TPM)	CONCEPT	The unit in which quota is assigned and a deployment's rate limit is set. · also S18, S20
Requests-Per-Minute (RPM)	CONCEPT	A request rate limit set in proportion to a deployment's TPM. · also S18
Microsoft.DefaultV2	SETTING	The guardrail assigned to model deployments by default. · also S39
PromptAgentDefinition	CLASS	The Python class describing a prompt agent: its model and instructions, and optionally tools.
create_version	METHOD	Creates an agent, or a new immutable version of an existing agent.
delete_version	METHOD	Deletes one version of an agent, given agent_name and agent_version.
version_selector	SETTING	Chooses which agent version the stable agent endpoint serves: latest, or a pinned version.

S11 · Choosing memory, tool and knowledge services for agents

Web search	TOOL	Built-in tool that grounds answers in real-time public web results with inline citations.
File Search	TOOL	Built-in tool that searches files you upload into a vector store.
Code Interpreter	TOOL	Built-in tool that writes and runs Python code in a sandboxed environment.
Function calling	TOOL	Custom functions the agent can ask for; your app runs them and returns the result.
OpenAPI tool	TOOL	Connects an agent to an external HTTP API described by an OpenAPI 3.0 or 3.1 specification.
Model Context Protocol (MCP)	TOOL	Connects an agent to tools hosted on a remote MCP server endpoint.
Toolbox	FEATURE	A curated, versioned set of tools exposed through one MCP-compatible endpoint for many agents.

A · Plan and manage an Azure AI solution · continued

S11 · Choosing memory, tool and knowledge services for agents

Grounding with Bing Search	TOOL	Web grounding tool that needs your own Grounding with Bing Search resource and a project connection.
memory store	CONCEPT	The managed store (preview) that holds an agent's long-term memories across sessions.
User profile memory	CONCEPT	Memory type for durable preferences and personal context, such as language or accessibility needs.
Chat summary memory	CONCEPT	Memory type holding distilled summaries of earlier conversation topics.
Procedural memory	CONCEPT	Memory type holding reusable routines inferred from earlier interactions.
knowledge sources	CONCEPT	The connections to indexed or remote content that make up a Foundry IQ knowledge base.

S18 · Quotas, scaling, rate limits and cost footprints

Request quota	PORTAL	The button on Manage > Quota for asking for more standard quota.
retry-after-ms	SETTING	The 429 response header giving the recommended wait, in milliseconds, before retrying.
max_retries	PARAMETER	The openai client setting for built-in retries with exponential backoff; the default is two.
RateLimitError	CLASS	The openai package exception raised for a 429 response.
az cognitiveservices usage list	CLI	Shows quota usages for Cognitive Services, including Foundry, in a region.
Provisioned throughput units (PTUs)	TIER	The unit of reserved model processing capacity for provisioned deployments.
Spillover	FEATURE	Sends overflow from a full provisioned deployment to a standard deployment in the same resource.
x-ms-spillover-deployment	SETTING	The request header that controls spillover for an individual request.
Azure Reservations	FEATURE	One-month or one-year discounts on the PTU meter; they don't guarantee capacity.
Cost analysis	PORTAL	The Cost Management view of spend, groupable by meter and filterable by tag.
Cost Management Reader	ROLE	A role that lets you view costs and usage data.

S19 · Monitoring models and data: drift, safety events, grounding, index health, relevance

Azure OpenAI Requests	CONCEPT	The metric counting calls to the API, splittable by deployment and status code.
Processed Inference Tokens	CONCEPT	The metric for prompt plus generated tokens; REST name TokenTransaction.
Blocked Volume	CONCEPT	The metric counting calls rejected by the content filter.
Harmful Volume Detected	CONCEPT	The metric counting calls the content filter detected as harmful, blocked or annotated.
Diagnostic settings	PORTAL	Where you route a resource's logs and metrics to a destination such as Log Analytics. · also S24
Continuous evaluation	FEATURE	Quality and safety evaluation of production traffic at a sampled rate.
Scheduled evaluation	FEATURE	Evaluation on a schedule against test datasets, to detect system drift.

A · Plan and manage an Azure AI solution · continued

S19 · Monitoring models and data: drift, safety events, grounding, index health, relevance

ContinuousEvaluationRuleAction	CLASS	The rule action that runs an evaluation on agent responses, with an hourly cap.
max_hourly_runs	PARAMETER	Caps how many continuous evaluation runs happen per hour; extra runs are skipped.
get_indexer_status	METHOD	Returns an indexer's current status and execution history from Python. · also S25
Docs Succeeded	PORTAL	The count of documents indexed in a run, shown in the indexer's execution details.
Throttled search queries percentage	CONCEPT	The AI Search metric for the share of queries that were throttled.
@search.rerankerScore	CONCEPT	The semantic ranker's relevance score, from 0.00 to 4.00.
Max failed items	SETTING	The error count below which an indexer run still reports success.

S21 · Security: managed identity, private networking, keyless credentials, role policies

System-assigned managed identity	CONCEPT	An identity switched on inside one Azure resource and deleted with it.
User-assigned managed identity	CONCEPT	A standalone identity resource that several resources can share; you delete it yourself.
az identity create	CLI	Azure CLI command that creates a user-assigned managed identity.
ManagedIdentityCredential	CLASS	The azure-identity credential Learn suggests in production on Azure instead of the full default chain.
https://ai.azure.com/.default	SETTING	The Microsoft Entra ID token scope for calls to Foundry.
disableLocalAuth	SETTING	Resource property that turns off key-based authentication when set to true.
az role assignment create	CLI	Assigns a role to a principal at a scope.
Foundry Owner	ROLE	Highly privileged role: full management plus building in projects; formerly Azure AI Owner.
Cognitive Services User	ROLE	Role for calling AI Services resources directly; Learn's RBAC page says not to use it for Foundry project work. · also S28, S33, S36
Private endpoint connections	PORTAL	The Networking tab where private endpoints to the Foundry resource are managed.
privatelink.services.ai.azure.com	SETTING	One of the private DNS zones a Foundry private endpoint needs.
Microsoft.App/environments	SETTING	The delegation required on the agent subnet for virtual network injection.
Allow Only Approved Outbound	SETTING	Managed virtual network mode that permits only approved outbound traffic, through a managed firewall.

S22 · Foundry projects in CI/CD pipelines

OpenID Connect	CONCEPT	Token exchange that lets GitHub Actions sign in to Azure without a stored secret.
azure/login@v2	TOOL	The GitHub Action that signs a workflow in to Azure.
azure/cli@v2	TOOL	The GitHub Action that runs Azure CLI scripts after sign-in.
id-token: write	SETTING	Workflow permission required to fetch the OIDC token.
Federated credentials	PORTAL	The Settings page of a user-assigned managed identity where you add trust for GitHub.

A · Plan and manage an Azure AI solution · continued

S22 · Foundry projects in CI/CD pipelines

GitHub Actions deploying Azure resources	PORTAL	The federated credential scenario to choose for GitHub workflows.
az deployment group create	CLI	Deploys a Bicep file to a resource group.
Microsoft.CognitiveServices/accounts/deployments	CONCEPT	The resource type of a model deployment, a child of the Foundry resource.
raiPolicyName	PARAMETER	The deployment property that assigns a guardrail.
microsoft/ai-agent-evals	TOOL	Preview GitHub Action that evaluates Foundry agents inside a workflow.
AIAgentEvaluation@2	TOOL	Preview Azure Pipelines task that evaluates Foundry agents.
azd pipeline config	CLI	Azure Developer CLI command that generates a CI/CD pipeline and its federated credentials.
result_counts	PARAMETER	The run field with total, passed, failed and errored counts, used for a gate.

S23 · Responsible AI: safety filters, guardrails, risk detection, moderation, evaluators

Create Guardrail	PORTAL	The Build > Guardrails button that opens the guardrail wizard. · also S40
Annotate and block	SETTING	Control action that records a detected risk and stops the content; works for models and agents. · also S40
Tool response	SETTING	Agent-only intervention point that scans content a tool returns to the agent.
User prompt attacks	FEATURE	Risk covering jailbreak-style attempts in the user's own prompt.
Prompt Shields	FEATURE	Detection of adversarial input: user prompt attacks and document (indirect) attacks. · also S33
Spotlighting	FEATURE	Preview defence that encodes document content so the model trusts it less than direct prompts. · also S33
Protected material for text	FEATURE	Risk that detects known third-party text such as lyrics or articles in output.
Groundedness detection	FEATURE	Content Safety feature that checks whether responses are grounded in supplied sources.
Custom blocklist	FEATURE	Your own list of exact or regex terms to filter.
Azure AI Content Safety	SERVICE	Standalone service that detects harmful user- and AI-generated content through its own APIs. · also S33
BadRequestError	CLASS	The OpenAI Python SDK error raised when a guardrail blocks input (code content_filter).
AI Red Teaming Agent	TOOL	Foundry tool that runs PyRIT-based attacks against your model or app and scores them.
Attack Success Rate (ASR)	CONCEPT	Percentage of red-teaming attacks that succeeded; the key red-teaming metric.

S24 · Auditing and governing agents: trace logs, provenance, approvals, tool-access controls

Microsoft Foundry Control Plane	PORTAL	Unified interface for visibility, governance and control of agents, models and tools across Foundry.
guardrail policies	FEATURE	Minimum guardrail controls mandated across a subscription or resource group, enforced with Azure Policy.
Resource Policy Contributor	ROLE	Role that, like Owner, can create or edit guardrail policies.
Microsoft Entra Agent ID	SERVICE	Entra ID's framework for identities built for AI agents, used by Foundry.

A · Plan and manage an Azure AI solution · continued

S24 · Auditing and governing agents: trace logs, provenance, approvals, tool-access controls

agent identity blueprint	CONCEPT	The Entra object from which agent identities are created; the project managed identity authenticates it.
AI gateway	FEATURE	Preview API Management route for MCP traffic that enforces authentication, rate limits, IP rules and audit logging.
Microsoft Purview	SERVICE	Data security and compliance service that captures Foundry prompts and responses in the unified audit log.

B · Implement generative AI and agentic solutions

30–35% of the exam

S5 · First code: connect an app to a Foundry project and call deployed models

azure-ai-projects	PACKAGE	The Python package for the Foundry SDK, which works through the project endpoint.
azure-identity	PACKAGE	The Python package that provides DefaultAzureCredential and other Microsoft Entra ID credentials.
AIProjectClient	CLASS	The Foundry SDK client, created with the project endpoint and a credential.
DefaultAzureCredential	CLASS	A credential that tries sign-in methods in turn, such as your az login session or a managed identity.
get_openai_client	METHOD	Returns the project's OpenAI-compatible client for Responses, agents and evaluations.
FOUNDRY_PROJECT_ENDPOINT	SETTING	The environment variable Learn's samples use to hold the project endpoint.
Foundry SDK	CONCEPT	Microsoft's thin-client SDK exposing all Foundry project APIs through one project endpoint.
get_bearer_token_provider	METHOD	Turns a credential into a token provider the OpenAI client can use instead of a key.
chat.completions.create	METHOD	The Chat Completions call, taking a messages list, for models without Responses API support.
previous_response_id	PARAMETER	Chains a Responses call to an earlier response to continue the exchange. · also S20
input_image	PARAMETER	The content part type for sending an image, by URL or base64 data URL, in the Responses API.
400 Model not supported	CONCEPT	The error a deployment returns when its model does not support the Responses API.

S6 · Tuning generation: prompt engineering and model parameters

system message	CONCEPT	Instructions and context given to the model to guide its responses; sent first in chat APIs.
few-shot learning	CONCEPT	Including example input and output pairs in the prompt to show the model what you want.
recency bias	CONCEPT	The tendency for the end of a prompt to influence output more than the beginning.
Chain of thought prompting	CONCEPT	Asking a non-reasoning model to proceed step by step and show the steps; not for reasoning models. · also S16
grounding data	CONCEPT	Data you give the model to draw its answer from, for reliable, up-to-date responses.
temperature	PARAMETER	Sampling randomness, from 0 to 2; lower is more focused. Not supported by reasoning models.

B · Implement generative AI and agentic solutions · continued

S6 · Tuning generation: prompt engineering and model parameters

<code>top_p</code>	PARAMETER	Nucleus sampling, an alternative to temperature; change one or the other, not both.
<code>max_output_tokens</code>	PARAMETER	Responses API cap on generated tokens, including reasoning tokens.
<code>max_completion_tokens</code>	PARAMETER	Chat Completions cap on generated tokens, including reasoning tokens; replaces the deprecated <code>max_tokens</code> .
<code>stop</code>	PARAMETER	Chat Completions: up to four sequences at which generation stops.
<code>frequency_penalty</code>	PARAMETER	Chat Completions, -2.0 to 2.0; positive values reduce verbatim repetition.
<code>presence_penalty</code>	PARAMETER	Chat Completions, -2.0 to 2.0; positive values push the model towards new topics.
<code>reasoning_effort</code>	PARAMETER	Chat Completions setting for how much a reasoning model reasons; Responses uses <code>reasoning={"effort": ...}</code> .

S7 · Retrieval-augmented generation (RAG) in an application

Azure OpenAI On Your Data	FEATURE	Managed grounding of GPT-4o models on your data; deprecated and approaching retirement.
<code>text-embedding-3-small</code>	MODEL	An embedding model sold by Azure, producing 1,536-dimension vectors.
<code>embeddings.create</code>	METHOD	The OpenAI SDK call that turns input text into an embedding vector.
Text Split skill	FEATURE	The built-in indexer skill that chunks text by pages or sentences.
<code>azure-search-documents</code>	PACKAGE	The Python package for Azure AI Search index management and queries. · also S26
<code>SearchIndexClient</code>	CLASS	Creates, updates and deletes search indexes.
<code>SearchClient</code>	CLASS	Uploads documents to one index and queries it.
<code>VectorizedQuery</code>	CLASS	A vector query built from an embedding you computed; pass it in <code>vector_queries</code> . · also S26
<code>vector_search_dimensions</code>	PARAMETER	The number of dimensions a vector field holds; must match the embedding model.
Search Index Data Contributor	ROLE	Data-plane role to load, change and query documents in an index. · also S13
Search Service Contributor	ROLE	Role to create and manage indexes and other objects on a search service. · also S25, S27
Dedicated pricing model	TIER	Provisioned search capacity, including Basic, billed at a fixed hourly rate while the service exists.
<code>az search service delete</code>	CLI	Deletes a search service and its data; the only way to stop Dedicated billing.

S8 · Workflows, tool-augmented flows and multistep reasoning

<code>tool_choice</code>	PARAMETER	Controls whether the model may call tools: auto (default), none, or a named function to force. · also S12, S24
<code>function_call_output</code>	PARAMETER	Responses API input item that returns your function's result, matched by <code>call_id</code> . · also S12
<code>call_id</code>	PARAMETER	The identifier on a <code>function_call</code> item that your <code>function_call_output</code> must repeat. · also S12

B · Implement generative AI and agentic solutions · continued

S8 · Workflows, tool-augmented flows and multistep reasoning

<code>parallel_tool_calls</code>	PARAMETER	Whether to enable parallel function calling during tool use; true by default in Chat Completions.
<code>Sequential orchestration</code>	CONCEPT	Agents or steps run in a fixed linear order, each using the previous output.
<code>prompt chaining</code>	CONCEPT	Another name for sequential orchestration: one call's output becomes the next call's input.
<code>Magentic orchestration</code>	CONCEPT	A manager agent builds and tracks a task ledger for open-ended problems.
<code>Handoff orchestration</code>	CONCEPT	Agents decide whether to handle a task or transfer it to a better-suited agent.
<code>Human in the loop</code>	FEATURE	Workflow pattern that asks the user a question and waits for input before continuing. · also S15, S16
<code>Microsoft Agent Framework</code>	TOOL	Open-source SDK for agents and code-first multi-agent workflows; Learn's path for new workflows. · also S15
<code>agent-framework</code>	PACKAGE	The pip package for Microsoft Agent Framework in Python. · also S15

S9 · Evaluating models and apps: fabrications, relevance, quality, safety

<code>azure-ai-evaluation</code>	PACKAGE	The Azure AI Evaluation SDK package for local evaluation; documented under the Foundry (classic) portal.
<code>GroundednessEvaluator</code>	CLASS	Classic SDK class that scores how well a response is grounded in context.
<code>RelevanceEvaluator</code>	CLASS	Classic SDK class that scores how well a response addresses the query.
<code>builtin.groundedness</code>	SETTING	Current evaluator identifier for Groundedness in cloud evaluation. · also S20
<code>builtin.relevance</code>	SETTING	Current evaluator identifier for Relevance in cloud evaluation.
<code>builtin.violence</code>	SETTING	Risk and safety evaluator identifier; uses hosted models and a 0 to 7 severity scale. · also S23
<code>Groundedness Pro</code>	FEATURE	Preview evaluator that checks strict consistency with context using Azure AI Content Safety.
<code>Response Completeness</code>	FEATURE	Preview evaluator for the recall side: does the response miss information in the ground truth?
<code>TestingCriterionAzureAIEvaluator</code>	CLASS	Foundry SDK class that adds a built-in evaluator, with its data mapping, to an evaluation. · also S40
<code>testing_criteria</code>	PARAMETER	The list of evaluators an evaluation applies to each row.
<code>data_mapping</code>	PARAMETER	Maps evaluator inputs to dataset fields using templates such as <code>{{item.response}}</code> .
<code>evals.create</code>	METHOD	Creates an evaluation definition through the project's OpenAI client.
<code>report_url</code>	PARAMETER	Run property linking to the evaluation results in the Foundry portal. · also S40
<code>AI red teaming agent</code>	TOOL	Simulates attacks with Microsoft's PyRIT framework to find safety and security weaknesses before deployment.

S10 · Integrating generative workflows with the Foundry SDKs and connectors

<code>Add connection</code>	PORTAL	The button on the Connected resources tab that starts a new project connection.
-----------------------------	--------	---

B · Implement generative AI and agentic solutions · continued

S10 · Integrating generative workflows with the Foundry SDKs and connectors

Project client	CONCEPT	The Foundry SDK client for Foundry-native operations such as listing connections and enabling tracing.
OpenAI-compatible client	CONCEPT	The client from <code>get_openai_client()</code> for Responses, agents, evaluations and fine-tuning.
connections.get	METHOD	Returns one project connection by name, including its id and target.
SEARCH_CONNECTION_NAME	SETTING	The name Learn's Azure AI Search tool sample uses for the search connection's name.
Custom key	CONCEPT	Connection type that stores keys and related properties securely for a custom target.
Foundry Tools Catalog	PORTAL	Where you browse over 1,000 pre-built connectors to SaaS, data and business systems.
Connector Namespace	SERVICE	A managed service per Foundry account that hosts connector runtimes and managed MCP servers.
MCPTool	CLASS	Python class for an MCP server tool on an agent; takes <code>project_connection_id</code> for credentials. · also S14, S27
project_connection_id	PARAMETER	Tool setting that references credentials stored in a project connection instead of embedding them. · also S27

S12 · Defining an agent: role, goals, conversation tracking, tool schemas

FunctionTool	CLASS	The Foundry SDK class that defines a function tool: name, description, parameters schema, strict.
strict	PARAMETER	FunctionTool setting that enforces stricter validation of the parameters schema.
additionalProperties	PARAMETER	JSON schema key; set to false so the model cannot send fields you did not define. · also S34
conversations.delete	METHOD	Deletes a stored conversation and the items it holds.
agent_reference	PARAMETER	The <code>extra_body</code> entry that tells a responses call which agent to run. · also S40
conversations.items.create	METHOD	Adds items, such as a user message, to an existing conversation.
Message items	CONCEPT	Conversation items holding user or assistant messages.
Tool call items	CONCEPT	Conversation items recording the tool invocations the agent attempted.
Conversations	CONCEPT	The current replacement for classic threads; they hold streams of items, not only messages.

S13 · Agents with retrieval, function calling and conversation memory

FileSearchTool	CLASS	Attaches File Search to an agent, pointing it at a vector store. · also S40
vector_store_ids	PARAMETER	The FileSearchTool parameter listing the vector store the agent searches.
vector_stores.create	METHOD	Creates a vector store through the project's OpenAI client.
upload_and_poll	METHOD	Uploads a file into a vector store and returns once processing has finished. · also S40
vector_stores.delete	METHOD	Deletes a vector store when you no longer need it.
AzureAISearchTool	CLASS	Attaches an existing Azure AI Search index to an agent through a project connection. · also S27

B · Implement generative AI and agentic solutions · continued

S13 · Agents with retrieval, function calling and conversation memory

AzureAISearchQueryType	CLASS	Chooses the query type: simple, vector, semantic or one of the two hybrids.
MemorySearchPreviewTool	CLASS	Attaches a memory store to a prompt agent so it reads and writes long-term memory.
MemoryStoreDefaultDefinition	CLASS	Defines a memory store: its chat model, embedding model and options.
update_delay	PARAMETER	Seconds of inactivity before memories are written; the default is 300.
Short-term memory	CONCEPT	Context for the current session, such as a conversation's items.
Long-term memory	CONCEPT	Distilled knowledge an agent keeps across sessions in a memory store.

S14 · Agent tools: APIs, knowledge stores, search, Content Understanding, custom functions

OpenApiTool	CLASS	Adds an OpenAPI-described HTTP API to an agent as a tool.
OpenApiFunctionDefinition	CLASS	Holds an OpenAPI tool's name, specification, description and authentication.
OpenApiAnonymousAuthDetails	CLASS	Authentication setting for public APIs that need no credential.
operationId	PARAMETER	Required name for each operation in the spec; letters, hyphens and underscores only.
server_label	PARAMETER	Your identifying label for an MCP server in the tool definition.
require_approval	PARAMETER	Whether MCP tool calls need approval: always (default), never, or per tool. · also S16, S20, S24
mcp_approval_request	CONCEPT	The output item type raised when an MCP call awaits approval. · also S16, S24
McpApprovalResponse	CLASS	Your app's answer to an approval request: approve and approval_request_id. · also S24
allowed_tools	PARAMETER	An allow-list of the MCP server's tools the agent may call. · also S24
knowledge_base_retrieve	TOOL	The knowledge base MCP tool an agent is allowed to call to query Foundry IQ. · also S27
RemoteTool	SETTING	The project connection category used to reach a knowledge base's MCP endpoint.
Azure Content Understanding skill	FEATURE	Azure AI Search skill that uses Content Understanding analyzers to extract and chunk content at indexing time. · also S25

S15 · Multi-agent orchestration, and orchestrating models, flows and rules engines

Connected Agents	TOOL	Classic agent-to-agent tool; not available in the new Foundry Agent Service, where Learn recommends the A2A tool.
A2APreviewTool	CLASS	Adds an Agent2Agent (A2A) endpoint, through a project connection, as a tool for an agent.
Foundry Agent Consumer	ROLE	Role the calling identity needs, at least, on a target Foundry project for A2A calls. · also S21
Group chat	CONCEPT	Workflow pattern that passes control between agents based on context or rules.
Power Fx	FEATURE	The Excel-like, low-code formula language used for expressions and conditions in Foundry workflows.

B · Implement generative AI and agentic solutions · continued

S15 · Multi-agent orchestration, and orchestrating models, flows and rules engines

Azure Logic Apps	SERVICE	Low-code workflow service that can call Foundry agents as steps alongside deterministic actions.
FoundryChatClient	CLASS	Agent Framework client that uses a model deployed in a Foundry project from your own code.
SequentialBuilder	CLASS	Builds a pipeline in which agents process the task in turn.
Executor	CLASS	Agent Framework base class for a custom code step inside a workflow.
Handoff	CONCEPT	Orchestration pattern in which agents transfer control to each other based on context.
Magentic	CONCEPT	Orchestration pattern in which a manager agent dynamically coordinates specialised agents.

S16 · Autonomy with safeguards and approvals; reflection and self-critique loops

mcp_approval_response	CONCEPT	The input item type you send back to approve or reject the call.
approval_request_id	PARAMETER	Links your approval response to the request item it answers.
Ask a question	PORTAL	The workflow node that pauses and waits for the user's input or approval.
approval_mode	PARAMETER	Agent Framework's tool setting; always_require makes a function wait for human approval.
user_input_requests	CONCEPT	Where an Agent Framework run lists the tool calls waiting for approval.
maker-checker loop	CONCEPT	One agent proposes, another checks against criteria and sends it back until approved or capped.
iteration cap	CONCEPT	The limit that stops a refinement loop, with a defined fallback when it is reached.
reasoning summary	FEATURE	The supported way to see a summary of a reasoning model's chain of thought.
output_tokens_details.reasoning_tokens	PARAMETER	Where a Responses API response reports how many reasoning tokens were used.

S17 · Observability: tracing, token analytics, safety signals, latency; agent error analysis

AIProjectInstrumentor	CLASS	Turns on GenAI tracing for Foundry SDK and OpenAI Responses calls when you call instrument(). · also S20, S40
azure-monitor-opentelemetry	PACKAGE	The package that exports OpenTelemetry traces to Azure Monitor Application Insights.
configure_azure_monitor	METHOD	Sends your app's spans to the Application Insights resource named by a connection string. · also S40
get_application_insights_connection_string	METHOD	Reads the connection string of the project's connected Application Insights resource. · also S40
AZURE_EXPERIMENTAL_ENABLE_GENAI_TRACING	SETTING	Must be true before instrument() is called, or no GenAI spans are produced.
OTEL_INSTRUMENTATION_GENAI_CAPTURE_MESSAGE_CONTENT	SETTING	Records message content and tool arguments in spans. Off by default; development use only. · also S24
trace_function	METHOD	A decorator that creates a span for each call to your own function.
Log Analytics Reader	ROLE	The role needed on the connected Application Insights resource to query traces. · also S40

B · Implement generative AI and agentic solutions · continued

S17 · Observability: tracing, token analytics, safety signals, latency; agent error analysis

<code>execute_tool</code>	CONCEPT	The span name that records a tool execution inside an agent trace.
Agent Monitoring Dashboard	PORTAL	The agent's Monitor tab: token usage, latency, run success rate, evaluations and red teaming.
Time to Response	CONCEPT	The Azure Monitor latency metric for streaming requests: time to the first response chunk.
Task Adherence	CONCEPT	A system evaluator: did the agent follow the rules and constraints in its instructions?
Tool Call Accuracy	CONCEPT	A process evaluator: did the agent make the right tool calls with correct parameters?
Cluster analysis	FEATURE	Groups evaluation samples with similar problems so you can spot recurring errors (preview).

C · Implement computer vision solutions

10–15% of the exam

S29 · Content Understanding for images and video: pipelines, segments, objects and regions

<code>prebuilt-image</code>	MODEL	The base analyzer a custom image analyzer inherits from.
<code>prebuilt-video</code>	MODEL	The base analyzer a custom video analyzer inherits from.
<code>prebuilt-imageSearch</code>	MODEL	The RAG analyzer that generates a one-paragraph description and insights for an image.
<code>enableSegment</code>	SETTING	Config switch that turns on video segmentation by <code>contentCategories</code> .
<code>contentCategories</code>	SETTING	Natural-language category descriptions used to segment or classify content, optionally routed to other analyzers.
<code>cameraShotTimesMs</code>	CONCEPT	Video output listing the timestamps, in milliseconds, where camera shots change.
<code>keyFrameTimesMs</code>	CONCEPT	Video output listing the timestamps of the key frames extracted for analysis.
<code>pro mode</code>	FEATURE	Preview mode (2025-05-01-preview) adding multi-step reasoning, multiple inputs and reference data, for documents only.
<code>agentic mode</code>	FEATURE	Preview document-analyzer workflow (<code>config.workflow</code> set to <code>agentic</code>) that reasons over a document for multistep answers.
<code>AudioVisualContent</code>	CLASS	The Python result class for audio and video content, with timing and, for video, frame details.
<code>Object detection</code>	FEATURE	Image Analysis 4.0 feature returning bounding box coordinates for detected objects; retiring in 2028.

S30 · Multimodal understanding: visual context, captions, visual Q&A, alt-text

Vision-enabled chat models	MODEL	Large multimodal models that analyse images and answer questions about them in text.
GPT-5 series	MODEL	One of the model families Learn lists as vision-enabled.
<code>input_text</code>	PARAMETER	The Responses API content part type for text sent alongside images.
<code>detail</code>	PARAMETER	The <code>input_image</code> property that sets the detail level: low, high or auto.
<code>responses.parse</code>	METHOD	Responses call that parses the model's output into a Pydantic model you supply. · also S34

C · Implement computer vision solutions · continued

S30 · Multimodal understanding: visual context, captions, visual Q&A, alt-text

<code>text_format</code>	PARAMETER	The responses.parse argument that takes the Pydantic model defining the output schema. · also S34
<code>output_parsed</code>	PARAMETER	The response property holding the parsed structured result.
Contextual specificity	CONCEPT	Image prompting technique: add the scenario's context so the output fits its purpose.
Alternative text (alt text)	CONCEPT	The Microsoft Style Guide's guidance on text replacements for images.
Web Content Accessibility Guidelines (WCAG) 2.0	CONCEPT	The W3C accessibility standard the Microsoft Style Guide cites.
Azure Vision in Foundry Tools	SERVICE	The Foundry Tool whose Image Analysis 4.0 captioning is deprecated and retires in September 2028.

S31 · Image generation and editing: reference media, inpainting, masks, controls

Image Edit API	FEATURE	The endpoint that modifies existing images from a prompt, with optional mask and up to 16 input images.
<code>gpt-image-2</code>	MODEL	A generally available GPT-image model for high-resolution output and improved editing, with flexible sizes.
<code>gpt-image-1</code>	MODEL	A limited-access GPT-image model for realistic, instruction-following image generation and editing.
<code>gpt-image-1-mini</code>	MODEL	A limited-access, cost-efficient GPT-image model for fast or bulk generation; no input_fidelity support.
<code>images.generate</code>	METHOD	The OpenAI SDK call that generates images from a text prompt.
<code>size</code>	PARAMETER	The output dimensions, such as 1024x1024; allowed values depend on the model.
<code>quality</code>	PARAMETER	Image quality level: low, medium or high (more options on the 2.5 models).
<code>output_format</code>	PARAMETER	The returned file type: PNG by default, or JPEG.
<code>output_compression</code>	PARAMETER	A compression level from 0 to 100 for the generated image.
<code>background</code>	PARAMETER	Set to transparent, with PNG output, for a transparent background.
<code>mask</code>	PARAMETER	A PNG the same size as the input; fully transparent pixels mark the area to edit.
<code>input_fidelity</code>	PARAMETER	Controls how closely an edit matches the style and features, especially faces, of the input images.
<code>partial_images</code>	PARAMETER	With streaming, the number of intermediate images (1 to 3) returned before the final one.
Images playground	PORTAL	The Foundry portal page for trying image-generation prompts before writing code.

S32 · Video generation and editing generated video

Sora 2	MODEL	OpenAI's video model on Foundry (preview): text, image or generated video in; video with audio out.
<code>sora-2</code>	MODEL	The model ID you deploy for Sora 2 video generation.
<code>videos.create</code>	METHOD	Starts a video render job and returns at once with an ID and status.
<code>videos.retrieve</code>	METHOD	Gets the latest state of a video job; used to poll until it finishes.
<code>videos.download_content</code>	METHOD	Fetches the finished MP4 for a completed video job.

C · Implement computer vision solutions · continued

S32 · Video generation and editing generated video

<code>videos.remix</code>	METHOD	Edits a completed video by ID with a new prompt, keeping its structure, motion and framing.
<code>create_and_poll</code>	METHOD	The async helper that creates a video job and waits for it to finish.
<code>input_reference</code>	PARAMETER	One reference image anchoring the first frame; it must match the requested size exactly.
<code>seconds</code>	PARAMETER	Video duration: 4, 8 or 12 seconds in the Sora 2 parameter table.
<code>variant</code>	PARAMETER	Which content to download for a video; video fetches the MP4.
<code>in_progress</code>	CONCEPT	The job status while the video is being generated, after queued.

S33 · Responsible AI for images: unsafe content, injected text in images, visual policy

Analyze image API	FEATURE	Scans an image for sexual content, violence, hate and self-harm with severity levels.
ContentSafetyClient	CLASS	The Python client for Content Safety analysis calls.
AnalyzeImageOptions	CLASS	The request object that wraps the image data for an image analysis call.
<code>analyze_image</code>	METHOD	Sends the image request and returns a severity for each harm category.
Analyze multimodal API	FEATURE	Preview API that analyses an image and its text together for harm.
<code>enableOcr</code>	PARAMETER	Multimodal API setting that makes the service read text in the image and analyse it with the image.
<code>documentsAnalysis</code>	PARAMETER	The Prompt Shields response field holding attackDetected for each document you sent.
Indirect attacks	CONCEPT	The guardrail risk for instructions hidden in third-party content such as documents or tool output. · also S39
Content Credentials	FEATURE	C2PA-based, cryptographically signed provenance metadata that Foundry attaches to supported generated content.
Custom categories (rapid) API	FEATURE	Preview API for defining new harmful-content patterns and scanning text and images for them.

D · Implement text analysis solutions

10–15% of the exam

S34 · Extracting entities, topics, summaries and structured JSON

Structured outputs	FEATURE	Makes a model's response follow the JSON Schema supplied with the call.
<code>json_schema</code>	SETTING	The format type that switches structured outputs on, in <code>text.format</code> or <code>response_format</code> .
<code>response_format</code>	PARAMETER	Where the schema goes for the Chat Completions API.
JSON mode	FEATURE	The older feature that guarantees valid JSON but not adherence to your schema.
Azure Language in Foundry Tools	SERVICE	The Foundry Tool for natural language processing: NER, PII, key phrases, summarization and more.
Legacy capabilities	CONCEPT	Language features supported for existing workloads, such as key phrases, summarization and sentiment.
<code>azure-ai-textanalytics</code>	PACKAGE	The Python package for Azure Language text analysis.

D · Implement text analysis solutions · continued

S34 · Extracting entities, topics, summaries and structured JSON

TextAnalyticsClient	CLASS	The Python client for Language text analysis, created with an endpoint and a credential.
recognize_entities	METHOD	Runs prebuilt named entity recognition and returns entities with categories, offsets and confidence.
extract_key_phrases	METHOD	Returns the main talking points of each document as key phrases.
begin_extract_summary	METHOD	Starts extractive summarization as a long-running operation; takes max_sentence_count.

S35 · Sentiment, tone, safety and sensitive content; domain-specific outputs

analyze_sentiment	METHOD	Returns document and sentence sentiment with confidence scores for each document.
show_opinion_mining	PARAMETER	Set to True on analyze_sentiment to return targets and assessments for each sentence.
mined_opinions	PARAMETER	The per-sentence list of opinions, each a target with its assessments.
Opinion mining	FEATURE	Aspect-based sentiment analysis: opinions about specific attributes mentioned in the text.
recognize_pii_entities	METHOD	Detects personal data in each document and returns the entities and a redacted copy.
redacted_text	PARAMETER	The copy of the document with detected personal data masked.
entitySynonyms	PARAMETER	REST setting that teaches PII detection your own words for a known entity type.
valueExclusionPolicy	PARAMETER	REST setting listing values that must not be detected or redacted.
analyze_text	METHOD	Content Safety call that returns a severity for each harm category in a text.
AnalyzeTextOptions	CLASS	The request object holding the text, and optionally categories and blocklists, for analyze_text.
outputType	PARAMETER	REST setting choosing FourSeverityLevels or EightSeverityLevels for text analysis.
blocklistNames	PARAMETER	REST setting naming your own blocklists of banned terms to check the text against.
Few-shot learning	CONCEPT	Giving the model worked examples inside the prompt to shape its answers.
Supervised fine-tuning	FEATURE	Training a base model on your labelled examples to teach it a task, style or format.

S36 · Translation: Azure Translator in Foundry Tools vs LLM translation flows

Text translation	FEATURE	Translator's real-time translation of strings; the 2026-06-06 version adds LLM choice.
Document translation	FEATURE	Translator for whole files, preserving structure and format, as batch or single-file jobs.
Custom Translator	FEATURE	Builds a customised NMT model from your parallel documents for domain terms and style. · also S39
Adaptive custom translation	FEATURE	Steers an LLM with your reference sentence pairs at request time; no model is trained.
deploymentName	PARAMETER	Per-target choice of engine: general (NMT, the default), an LLM deployment, or a custom model's category ID.

D · Implement text analysis solutions · continued

S36 · Translation: Azure Translator in Foundry Tools vs LLM translation flows

tone	PARAMETER	Per-target register for LLM translation: formal, informal or neutral.
Ocp-Apim-ResourceId	PARAMETER	Header carrying the resource ID when an Entra token is sent to the global Translator endpoint.
azure-ai-translation-text	PACKAGE	The Python client library for Translator text translation.
TextTranslationClient	CLASS	The Python client for text translation, transliteration and supported languages.
DocumentTranslationClient	CLASS	The Python client for asynchronous batch document translation.
begin_translation	METHOD	Starts a batch document translation from a source container to a target container and returns a poller.

S37 · Speech for agents: speech to text, text to speech, custom speech models

azure-cognitiveservices-speech	PACKAGE	The Python package for the Speech SDK, imported as speechsdk.
SpeechConfig	CLASS	Holds the Speech endpoint, credential, recognition language and voice settings.
SpeechRecognizer	CLASS	Turns audio from a microphone or file into text.
recognize_once_async	METHOD	Transcribes a single utterance of up to 30 seconds, or until silence.
SpeechSynthesizer	CLASS	Turns text into synthesised speech, to a speaker or a file.
speak_text_async	METHOD	Synthesises the given text with the configured voice.
token_credential	PARAMETER	SpeechConfig argument that takes a Microsoft Entra TokenCredential for keyless access.
Cognitive Services Speech User	ROLE	A role that lets an identity use a Speech resource with Microsoft Entra authentication.
Voice Live API	SERVICE	Managed, low-latency speech-to-speech for voice agents, combining recognition, generative AI and text to speech. · also S39
AgentSessionConfig	CLASS	Voice Live settings that connect a voice session to a named Foundry agent and project.
GPT Realtime API	SERVICE	Azure OpenAI speech-in, speech-out models over WebRTC, WebSocket or SIP.
custom speech	FEATURE	Trains and deploys a speech recognition model on your own data to improve accuracy.
word error rate	CONCEPT	The quantitative accuracy measure used to test a custom speech model.

S38 · Reasoning over audio, and speech translation

input_audio	PARAMETER	The content part that carries base64 audio data and its format in an audio chat request.
modalities	PARAMETER	Chooses the output types of an audio chat call: text, audio, or both.
gpt-4o-audio-preview	MODEL	An audio model for audio and text generation through chat completions.
AzureOpenAI	CLASS	The OpenAI package client used by the audio quickstart, with an endpoint and API version.
azure_ad_token_provider	PARAMETER	AzureOpenAI argument that takes a Microsoft Entra token provider instead of a key.

D · Implement text analysis solutions · continued

S38 · Reasoning over audio, and speech translation

<code>gpt-audio</code>	MODEL	A newer audio model family for audio and text generation listed among Foundry Models.
<code>gpt-realtime-translate</code>	MODEL	A realtime audio model for multilingual translation with translated speech and text output.
LLM Speech	FEATURE	A Foundry speech API, enhanced by a large language model, for transcription and translation of audio files.
<code>enhancedMode</code>	PARAMETER	The LLM Speech setting that turns enhancement on and chooses the transcribe or translate task.
<code>SpeechTranslationConfig</code>	CLASS	Speech SDK configuration for translation: source language and target languages.
<code>TranslationRecognizer</code>	CLASS	Recognises speech and returns translations into the configured target languages.
<code>add_target_language</code>	METHOD	Adds an output language to a <code>SpeechTranslationConfig</code> .
<code>TranslatedSpeech</code>	SETTING	The result reason that signals a successful speech translation.
Live Interpreter	FEATURE	Real-time speech translation that detects the language and speaks the translation in a natural voice.

E · Implement information extraction solutions

10–15% of the exam

S25 · Ingesting and indexing documents, images, audio and video; the RAG ingestion flow with OCR

Document cracking	CONCEPT	The first indexer stage: opening files and extracting their content.
<code>imageAction</code>	PARAMETER	Indexer setting that extracts images from files so skills can read them.
<code>generateNormalizedImages</code>	SETTING	The <code>imageAction</code> value that produces normalized images for the OCR skill.
<code>Microsoft.Skills.Vision.OcrSkill</code>	FEATURE	The OCR skill type, which reads printed and handwritten text in images.
Document Layout skill	FEATURE	Skill that turns documents into layout-aware Markdown or text sections. · also S26
Import data	PORTAL	The search service Overview command that starts the wizard building a whole pipeline.
Storage Blob Data Reader	ROLE	Role the search service's identity needs to read blobs.
Cognitive Services OpenAI User	ROLE	Role a managed identity needs to send text to the embedding model.
<code>SearchIndexerClient</code>	CLASS	Python client for data sources, skillsets and indexers.
<code>prebuilt-audioSearch</code>	CONCEPT	Content Understanding RAG analyzer that transcribes and summarises conversations.
<code>prebuilt-videoSearch</code>	CONCEPT	Content Understanding RAG analyzer giving transcripts and descriptions per video segment. · also S29

S26 · Semantic, hybrid and vector search; enrichment with skills

semantic configuration	FEATURE	The index section that tells the semantic ranker which title, content and keyword fields to read.
<code>prioritizedFields</code>	SETTING	The part of a semantic configuration holding <code>titleField</code> , content fields and keyword fields in priority order.
Premium features	PORTAL	The Settings page of a search service where you choose the semantic ranker plan.

E · Implement information extraction solutions · continued

S26 · Semantic, hybrid and vector search; enrichment with skills

Hierarchical Navigable Small World (HNSW)	CONCEPT	The approximate nearest neighbour algorithm that suits most vector indexes.
Exhaustive K-Nearest Neighbors (KNN)	CONCEPT	The algorithm that scans every vector: precise, slower, useful as ground truth.
exhaustiveKnn	SETTING	The kind value for an exhaustive KNN algorithm configuration in <code>vectorSearch.algorithms</code> .
vectorSearchProfile	SETTING	The vector field property naming the profile that sets its algorithm, compression and vectorizer.
vectorizer	CONCEPT	Index component that turns query text or images into vectors at query time.
semantic_configuration_name	PARAMETER	The <code>search()</code> argument naming which semantic configuration the ranker uses.
Custom Web API skill	FEATURE	A custom skill that sends records to your own HTTPS endpoint during indexing.
Search Index Data Reader	ROLE	Built-in role that lets an identity query an index without a key.

S27 · Wiring retrieval into workflows and agents; multimodal extraction pipelines

AISearchIndexResource	CLASS	Names the index, its project connection and the query type for the Azure AI Search tool.
vector_semantic_hybrid	SETTING	The default query type of the Azure AI Search tool: vector plus keyword, semantically reranked.
knowledge_source	CONCEPT	A search service resource defining indexed or remote content that a knowledge base uses.
KnowledgeBaseRetrievalClient	CLASS	The Azure AI Search SDK client that calls a knowledge base's retrieve action from your own code.
Azure Document Intelligence in Foundry Tools	SERVICE	The Foundry Tool with read, layout, prebuilt and custom models for document extraction. · also S39
DocumentIntelligenceClient	CLASS	The Python client that analyses documents with prebuilt and custom Document Intelligence models.
begin_analyze_document	METHOD	Starts a document analysis with a given model ID and returns a poller for the result.
prebuilt-layout	MODEL	The model that returns text, tables, selection marks and document structure. · also S28
prebuilt-receipt	MODEL	The prebuilt model that extracts fields such as merchant name and transaction date from receipts.

S28 · Content Understanding for documents: grounded representations and analyzers

prebuilt-documentSearch	MODEL	The RAG analyzer that turns documents into Markdown with layout, figure descriptions, a summary and chunk-ready output.
prebuilt-invoice	MODEL	The prebuilt analyzer for invoices, utility bills, sales orders and purchase orders.
baseAnalyzerId	SETTING	The analyzer property naming the parent analyzer a custom analyzer inherits from.
fieldSchema	SETTING	The analyzer property that defines the structured fields to extract, generate or classify.
estimateFieldSourceAndConfidence	SETTING	Config option that returns a source location and confidence score for each field.
ContentUnderstandingClient	CLASS	The Python client for analysing content and creating and managing analyzers.

E · Implement information extraction solutions · continued

S28 · Content Understanding for documents: grounded representations and analyzers

<code>begin_create_analyzer</code>	METHOD	Creates a custom analyzer from a ContentAnalyzer definition as a long-running operation.
<code>ContentFieldSchema</code>	CLASS	The Python class holding a custom analyzer's named field definitions.
<code>GenerationMethod</code>	CLASS	The Python enum for a field's method: extract, generate or classify.
<code>azure-ai-contentunderstanding</code>	PACKAGE	The native Python SDK for Content Understanding, targeting the GA API.
Content Understanding Studio	PORTAL	The web experience for the full prebuilt catalogue and for building custom analyzers.
2025-11-01	SETTING	The generally available Content Understanding API version, recommended for production.

Across the course · review and capstone

session 40

S40 · Capstone review and exam technique

<code>Storage Blob Data Contributor</code>	ROLE	Role on the project's storage account that Learn lists for uploading files for file search.
<code>azure_ai_target_completions</code>	SETTING	Evaluation run data source type that sends each dataset item to a target such as an agent.
<code>builtin.coherence</code>	SETTING	The built-in AI-assisted evaluator that scores how coherent a response is.