

AI - 102

Azure AI Engineer Associate

Topics:



Assessed on this exam

- Plan and manage an Azure AI solution
- Implement generative AI solutions
- Implement an agentic solution
- Implement computer vision solutions
- Implement natural language processing solutions
- Implement knowledge mining and information extraction solutions

Languages

= Python, C → programming language by Microsoft

Time

= 90 minutes

Questions = 51

Overview

As a Microsoft Azure AI engineer, you build, manage, and deploy AI solutions that leverage Azure AI.

Your responsibilities include participating in all phases of AI solutions development, including:

- ✓ Requirements definition and design
- ✓ Development
- ✓ Deployment
- ✓ Integration
- ✓ Maintenance
- ✓ Performance tuning
- ✓ Monitoring

You work with solution architects to translate their vision. You also work with data scientists, data engineers, Internet of Things (IoT) specialists, infrastructure administrators, and other software developers to:

- Build complete and secure end-to-end AI solutions.
- Integrate AI capabilities in other applications and solutions.

Skills at a glance

✓ Plan and manage an Azure AI solution (20–25%)

✓ Implement generative AI solutions (15–20%)

- Implement an agentic solution (5–10%)

→ Less

✓ Implement computer vision solutions (10–15%)

✓ Implement natural language processing solutions (15–20%)

✓ Implement knowledge mining and information extraction solutions (15–20%)

AI Search

Doc Intelligence

20 - 25 %

PLAN

&

MANAGE

AZURE AI

SOLUTION

1. Select the appropriate Azure AI Foundry services

- Select the appropriate service for a generative AI solution
- Select the appropriate service for a computer vision solution
- Select the appropriate service for a natural language processing solution
- Select the appropriate service for a speech solution
- Select the appropriate service for an information extraction solution
- Select the appropriate service for a knowledge mining solution

Document
Intelligence

AI Search

2. Plan, create and deploy an Azure AI Foundry service

- Plan for a solution that meets Responsible AI principles
- Create an Azure AI resource
- Choose the appropriate AI models for your solution
- Deploy AI models using the appropriate deployment options
- Install and utilize the appropriate SDKs and APIs
- Determine a default endpoint for a service
- Integrate Azure AI Foundry Services into a continuous integration and continuous delivery (CI/CD) pipeline
- Plan and implement a container deployment

3. Manage, monitor, and secure an Azure AI Foundry Service

- ✓ • Monitor an Azure AI resource
- ✓ • Manage costs for Azure AI Foundry Services
- ✓ • Manage and protect account keys
- ✓ • Manage authentication for an Azure AI Foundry Service resource

4. Implement AI solutions responsibly

- Implement content moderation solutions
- Configure responsible AI insights, including content safety
- Implement responsible AI, including content filters and blocklists
- Prevent harmful behavior, including prompt shields and harm detection
- Design a responsible AI governance framework

1. Select the appropriate Azure AI Foundry services

1. AI Vision Solution

OCR - Optical character Recognition

Image / Video Analysis

Face - Liveness

The screenshot shows the Azure AI Foundry services interface with the 'Featured' tab selected. The page is organized into two rows of service cards. The first row includes: 'Search photos with image retrieval' (description: Retrieve specific moments within your photo album. For example, you can search for: a wedding you attended last summer, your pet, or your favorite city.), 'Add dense captions to images' (description: Generate human-readable captions for all important objects detected in your image.), 'Add captions to images' (description: Generate a human-readable sentence that describes the content of an image.), 'Detect common objects in images' (description: Recognize the location of objects of interest in an image and assign them a label.), 'Extract common tags from images' (description: Use an AI model to automatically assign one or more labels to an image.), and 'Create smart-cropped images' (description: Create cropped image thumbnails based on the key areas of a larger image.). The second row includes: 'Detect faces in an image' (description: Detect the location of one or more human faces in images, along with attributes such as gender and age.), 'Liveness detection' (description: Ensures the authenticity of users by verifying that the face in front of the camera), 'Extract text from images' (description: Extract printed and handwritten style text from images and documents for supported languages.), 'Portrait processing' (description: Enhance portrait images with features like face beautification, hair color, and more.), and 'Photo ID matching' (description: Match a photo ID to a person's face).

Imp.

1. OCR supports over 150 languages
2. Image Analysis  → A man in frame
↳ can also be used for content moderation
3. Face liveness
↳ whether face behind camera is
4. Video Analysis
↳ Analyzing videos.
e.g., count of objects in video

2. NLP Solution - AI language

Azure AI Language Mostly Text

- Language Detection
- Sentiment Analysis
- Imp. → Q & A chatbot
- NER, Key Phrase Extraction
- Text Classification

3. Azure AI Speech → mostly Audio

- Speech to Text STT
- Text to Speech TTS
- Speech Translation
- Speech Recognition

4. Generative AI Solution

Types of AI

Type	Purpose	Example
Conversational AI <u>GPT</u>	Communicate with humans via text or speech	ChatGPT, Alexa
Generative / Creation AI <u>DALL-E</u>	Create new content (text, image, video, code)	DALL-E, Copilot
Agentic AI <u>Autonomous</u>	Acts autonomously to achieve goals using tools & reasoning	AutoGPT, Azure AI Agents
Predictive AI <u>ML</u>	Analyzes data to predict outcomes	Forecasting, fraud detection
Perceptual AI <u>vision, speech</u>	Understands visuals, speech, and environment	Computer Vision, Speech Recognition
Autonomous AI <u>R.L</u>	Operates independently in real-world settings	Self-driving cars, drones

Topics to Focus:

- Generative
- Conversational }
- Agentic }

AI

RAG

Data Grounding

5. Document Intelligence Solution

"It detects tables, key-value pairs, and text from PDFs, invoices, or forms, and converts them into structured formats like spreadsheets or databases."

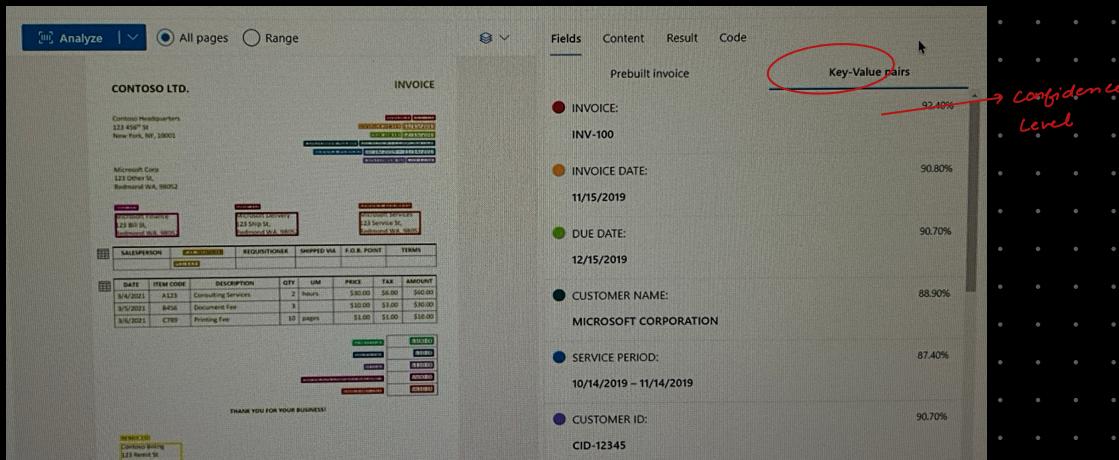
Azure AI Document Intelligence (formerly Form Recognizer) is a cloud service that automatically extracts structured data—like text, tables, and key-value pairs—from documents (invoices, receipts, forms, PDFs).

Key points:

- Prebuilt and custom models for various document types.
- Supports handwritten and printed text.
- Integrates with other Azure services for automation and AI workflows.
- Offers APIs, SDKs, and a web-based studio for building and testing models.

- Invoice
- Receipt
- Contract
- Tax Forms

In short: it turns documents into usable data automatically.



6. Azure AI Search → Knowledge mining.

Data is search through Indexes + Embeddings

Data Sources

1. Azure Storage

- Blob Storage – PDFs, images, text files.
- Data Lake Storage (Gen1 & Gen2) – Structured and unstructured datasets.
- Azure Files – Shared documents and files.

2. Databases

- Azure SQL Database – Relational tables.
- SQL Server – On-premises or in Azure.
- Cosmos DB – NoSQL JSON documents.
- PostgreSQL / MySQL – Via connectors.
- Azure Table Storage – Semi-structured data.

3. Microsoft 365 & Collaboration

- SharePoint Online / On-Premises – Documents and lists.
- OneDrive / Teams – Files and messages.
- Exchange Online – Emails.

4. External Sources

- Web pages / HTTP endpoints – Via custom indexers.
- Custom data sources – Any source that can be exposed through API or connectors.

5. AI-Processed Data

- Outputs from Azure Document Intelligence (tables, text, key-value pairs).
 - Embeddings or semantic data stored in supported databases.
- 💡 In short: Azure AI Search can access Azure storage, databases, Microsoft 365, web endpoints, and AI-processed data.

Imp. Even 3ps like { AWS, Adobe, Meta, GCP Cloud, Jira, Oracle etc }

2. Plan, create and deploy an Azure AI Foundry service

1. Responsible AI Principles

Responsible AI ensures AI systems are ethical, transparent, fair & accountable

Maximizing Benefit Minimizing Harm

1 Fairness

AI should treat everyone equally and avoid bias.

🧠 Example: A hiring model should not favor any gender or region.

2 Reliability & Safety

AI must perform consistently, safely, and accurately.

⚡ Example: Medical AI should give dependable results for all patients.

3 Privacy & Security

AI must protect personal data and ensure confidentiality.

🔒 Example: Use data encryption, anonymization, and secure model storage.

4 Inclusiveness

AI should empower and include all users, regardless of background or ability.

🌐 Example: Speech models that recognize diverse accents.

5 Transparency

AI systems should be understandable and explainable.

📊 Example: Using SHAP or LIME to explain model predictions.

6 Accountability

Humans are ultimately responsible for AI decisions.

⚖️ Example: Having clear audit trails and ethical oversight for deployed systems.

In Short — Responsible AI Principles Table

Principle	Meaning	Example / Practice
Fairness	Ensure unbiased and equal treatment	Remove bias from datasets, test fairness metrics
Reliability & Safety	System performs accurately and safely	Test under diverse scenarios, monitor drift
Privacy & Security	Protect user data and confidentiality	Encrypt, anonymize, comply with GDPR
Inclusiveness	Design for everyone, including minorities	Support diverse accents, accessible UIs
Transparency	Explain how AI makes decisions	Use SHAP/LIME, model documentation
Accountability	Humans remain responsible for outcomes	Human review, ethical governance e.g., if Self Driving Car Crashes

Someone should be accountable, it can't be none

2. Create Azure AI Resource

Create Azure AI services

Basics Network Identity Tags Review + create

Get access to Vision, Language, Search, and Speech Azure AI services with a single API key. Quickly connect services together to achieve more insights into your content and easily integrate with other services like Azure Search.

Project Details

Subscription * Azure subscription 1

Resource group * (New) ai102

Instance Details

Region * East US

Previous Next Review + create

End point is a web address of azure resource

ai102rs Overview Essentials JSON View

Resource group (move) : ai102

Status : Active

Location : East US

Subscription (move) : Azure subscription 1

Subscription ID : f153f722-aa61-4505-bdcb-9dfb5b3e4294

Tags (edit) : Add tags

API Kind : CognitiveServices

Pricing tier : Standard

Endpoint : https://ai102rs.cognitiveservices.azure.com/

Manage keys : Click here to manage keys

Autoscale : Disabled

Get Started Decision Language Speech Vision Document Intelligence Metrics Advisor Containers

Build intelligent apps using a comprehensive family of AI services and cognitive APIs

The Azure AI services multi-service resource combines various services from Decision, Language, Speech, Vision, and Applied AI into a single key and endpoint to enable you to easily build solutions that can see, hear, speak, understand, and make decisions. Follow the cards below to learn the basics, read documentation, and join the community.

Q. what is the difference b/w Azure AI service and individual service?

In Short:

Question

Need one AI capability (e.g., Vision only)?

Answer

Choose Individual Service

Need multiple AI features (Vision + Language + Speech)?

Choose Azure AI Services

Want simple, isolated billing?

Individual Service

Want unified management & scalability?

Azure AI Services

3.. Azure CI / CD - Deployment

💡 In Short

Term	Meaning in ML	Azure Example
CI (Continuous Integration)	Automate training, testing, and registering models	Azure DevOps or GitHub Action pipeline
CD (Continuous Delivery)	Automate model deployment	Deploy to Azure ML Managed Endpoint
CT (Continuous Training)	Automate retraining when new data arrives	Trigger ML pipeline from Data Lake update

🚀 Why CI/CD Matters in ML

- ✓ Reduces manual work
- ✓ Ensures consistent, reliable model deployment
- ✓ Tracks every version of data, code, and model
- ✓ Makes ML scalable and production-ready

🧠 In Simple Words:

CI/CD in ML = "Auto-train, auto-test, auto-deploy your models — reliably, repeatedly, and safely."

4.. Azure AI Container - Docker

🧠 Azure AI Container (Docker Deployment) — In Short

- It lets you **run Azure AI services or ML models locally** (on-premises or edge) using **Docker containers**, instead of calling cloud APIs.
- You **pull the container image** (e.g., for Computer Vision or Text Analytics) from Microsoft's registry and run it via Docker.
- The container exposes a **local REST API endpoint** (like `localhost:5000`) for your apps to call.
- You still use your **Azure key** for authentication and billing.
- Ideal for **data privacy, low-latency, offline, or hybrid cloud scenarios**.

Example:

bash

Copy code

```
docker pull mcr.microsoft.com/azure-cognitive-services/vision/computervision:latest  
docker run -p 5000:5000 -e ApiKey=<your_key> mcr.microsoft.com/azure-cognitive-services/v
```

✓ In short:

"Azure AI Containers let you bring cloud AI models to your local or private environment — securely, offline, and with full control."

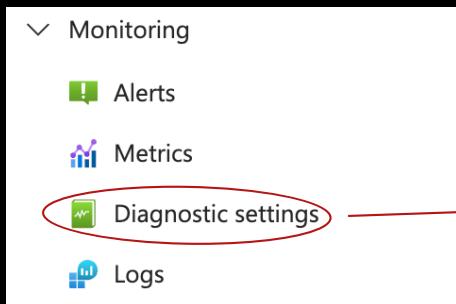
*Imp: It like running Azure AI models locally
i.e., let's say in Remote Area*

3. Manage, monitor, and secure an Azure AI Foundry Service

1. Monitoring

track performance, usage of your cloud resources.

Go to Resource — Turn on monitoring



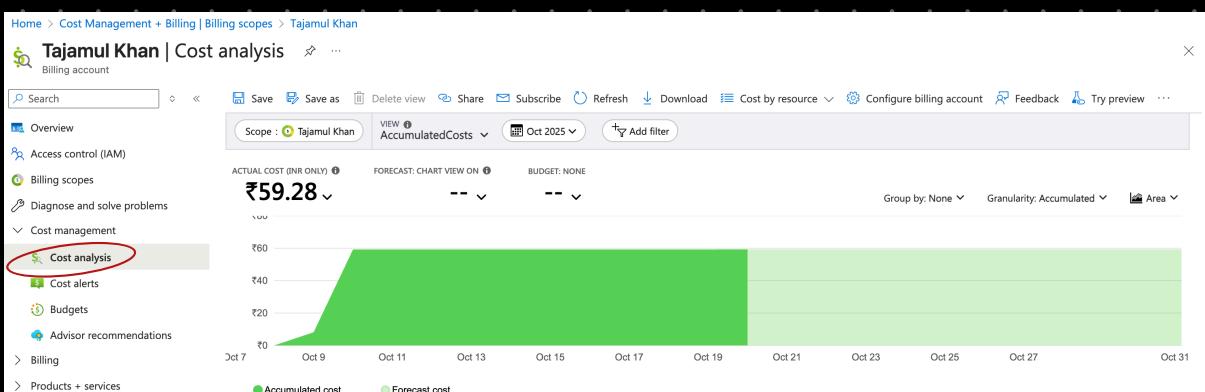
Turn On

KQL is used.

2. Manage Costs

a. Use pricing calculator by Azure

b. Use cost management tool



3. Azure Keys

To access we need

- Endpoint URL
- API Key

Azure Key Vault is like a password manager
key, certificates manager

Imp:
So, SSL certificates are managed by
Azure Key Vault

Recommended: Store keys in Azure Key Vault

Give
IAM → Access os
Key Vault Admin

Q: How to add Secret?

```
python
from azure.identity import DefaultAzureCredential
from azure.keyvault.secrets import SecretClient

VAULT_URL = "https://<your-vault-name>.vault.azure.net/"
SECRET_NAME = "my-secret"

client = SecretClient(vault_url=VAULT_URL, credential=DefaultAzureCredential())
secret = client.get_secret(SECRET_NAME)

print("Secret value:", secret.value)
```

4. Manage Private Communications

we can use

virtual networks; give access to specific IP's

private endpoints

4. Implement AI solutions responsibly

1. Text Moderation

- Hate
- Sexual
- Violence
- Self-harm

2. Image Moderation

we need to create Content Studio;

Get started with Content Safety Studio

Safeguard your text content with built-in features

Leverage our abilities to identify harmful text content across over 100 languages, and address concerns related to jailbreaking, hallucinations, and copyright infringements.



Moderate text content

Run moderation tests on text contents. Assess the test results with detected severities. Experiment with different threshold levels.

[Try it out](#)



Groundedness detection

Region not supported

Groundedness detection detects ungroundedness generated by the large language models (LLMs).

[Try it out](#)



Protected material detection for text

Use protected material detection to detect and protect third-party text material in LLM output.

[Try it out](#)



Protected material detection for code

Region not supported

PREVIEW

Run tests on code generated by LLM and identify whether the code already exists in GitHub repo.

[Try it out](#)



Prompt Shields

Prompt Shields provides a unified API that addresses Jailbreak attacks and Indirect attacks.

[Try it out](#)

Azure AI | Content Safety Studio

Content Safety Studio > Moderate text content

2. Test

Chopping tomatoes and cutting them into cubes or wedges are great ways to practice your knife skills.

Configure filters Use blocklist View code

Set the Severity thresholds for each category. Content with a severity level less than the threshold will be allowed. Learn more about categories and threshold

Category	Threshold level
Violence	Medium
Self-harm	Medium
Sexual	Medium
Hate	Medium

101/10000 characters

[Run test](#) [Select an Azure AI resource to run test. Go to select](#)

Content Studio

can also be used later to get API and moderate our content.

SUMMARY

Services;

- AI Search
- Document Intelligence
- AI Vision
 - OCR
 - Image Analysis
 - Video Analysis
 - Face
- AI Language
 - Text classification
 - Q&A chatbot
 - Sentiment Analysis
- AI Speech
 - Text To Speech
 - Speech To Text
 - Language Detection
- Generative AI
 - Text - GPT
 - Voice - whisper
 - Image - Dall e
 - Embedding - Ada



Responsible AI

F - Fairness
R - Reliable
P - Privacy
I - Inclusiveness
T - Transparency
A - Accountability



Moderation

- Content Studio
- Image moderation
- Text moderation

10 - 15 %

AI

VISION

SOLUTIONS

Analyze images

- Select visual features to meet image processing requirements
- Detect objects in images and generate image tags
- Include image analysis features in an image processing request
- Interpret image processing responses
- Extract text from images using Azure AI Vision
- Convert handwritten text using Azure AI Vision

Implement custom vision models

- Choose between image classification and object detection models
- Label images
- Train a custom image model, including image classification and object detection
- Evaluate custom vision model metrics
- Publish a custom vision model
- Consume a custom vision model
- Build a custom vision model code first

Analyze videos

- Use Azure AI Video Indexer to extract insights from a video or live stream
- Use Azure AI Vision Spatial Analysis to detect presence and movement of people in video

1. Analyze Images - AI vision Studio

The screenshot shows the Azure AI Vision Studio interface. At the top, there's a navigation bar with 'Azure AI | Vision Studio' and various icons. Below it, a sub-header says 'Vision Studio > Add captions to images'. A note says 'Use one of your own files or choose from a sample below.' On the left, there's a dashed box for file upload with options: 'Drag and drop a file here', 'Browse for a file', or 'Take a photo'. To the right are several sample images: two cows, a person with a surfboard, a building facade, a baseball player, a skateboarder, and some food. Below these is a section titled 'Sample image 4' showing a baseball player in action. To the right of the image are tabs for 'Detected attributes' (which is selected) and 'JSON'. Under 'Detected attributes', the text 'A baseball player holding a bat' is displayed, with 'Next steps' highlighted by a red circle. Below this are links: 'Get started', 'Try out the SDK' (which is bolded), 'Review pricing', and 'Responsible use of AI'. Further down are sections for 'Select a Vision resource' (with dropdowns for 'Subscription', 'Azure Resources', 'Region', and 'Vision endpoint') and 'Follow the quickstart' (with a link to 'Open the quickstart' and 'View the sample code on GitHub').

Services

This screenshot shows the 'Featured' section of the AI services page. It includes links to 'Optical character recognition', 'Face', and 'Image analysis'. Below are twelve service cards arranged in two rows of six:

- Search photos with image retrieval**: Detect specific moments within your photo album. [Try it out](#)
- Add dense captions to images**: Generate human-readable captions for all important objects detected in your image. [Try it out](#)
- Add captions to images**: Generate a human-readable sentence that describes the content of an image. [Try it out](#)
- Detect common objects in images**: Recognize the location of objects of interest in an image and assign them a label. [Try it out](#)
- Extract common tags from images**: Use an AI model to automatically assign one or more labels to an image. [Try it out](#)
- Create smart-cropped images**: Create cropped image thumbnails based on the key areas of a larger image. [Try it out](#)
- Detect faces in an image**: Detect the location of one or more human faces in images, along with attributes such as age and gender. [Try it out](#)
- Liveness detection**: Ensure the authenticity of users by verifying that the face in front of the camera is real. [Try it out](#)
- Extract text from images**: Extract printed and handwritten text from images and documents for supported languages. [Try it out](#)
- Portrait processing**: Enhance portrait images with features like face beautification, hair color, and more. [Try it out](#)
- Photo ID matching**: Match a photo ID to a person's face. [Try it out](#)

Imp: You will be charged for each service

Code; use Endpoint
API key

AI vision models = Image classify
object detect

Image classify = what is in the image
(add single/multi labels)

Object detect = Location of image

2. *Customize Models*; where we train our own

Image classifier
Object Detector

Those are available in Azure Custom Vision

Visual Intelligence Made Easy

Easily customize your own state-of-the-art computer vision models that fit perfectly with your unique use case. Just bring a few examples of labeled images and let Custom Vision do the hard work.

SIGN IN

Upload Images

Train

Evaluate

Let's test; Custom AI Vision

CustomizeSamples

Training Images Predictions Quick Test

Upload Images

Train labels
e.g.,
10 Images = 1
10 Images = 2

The longer we train model, the more accuracy.

Let's predict

View Endpoint Quick Test

Quick Test

This is a test image

Tag	Probability
Hemlock	99.9%
Japanese Cherry	0%

How to use the Prediction API

If you have an image URL:

```
https://ai102ci-prediction.cognitiveservices.azure.com/customvision/v3.0/Prediction
Set Prediction-Key Header to : x0NngcSg4j3NYXjneyFQoSZkugRom8iX3WYW3IFxjybI2... QJ99BJACYeBjFXJ3w3AA...
Set Content-Type Header to : application/json
Set Body to : {"Url": "https://example.com/image.png"}
```

If you have an image file:

```
https://ai102ci-prediction.cognitiveservices.azure.com/customvision/v3.0/Prediction
Set Prediction-Key Header to : x0NngcSg4j3NYXjneyFQoSZkugRom8iX3WYW3IFxjybI2... 2TiYyVhJQQJ99BJACYeBjFXJ3w3AA...
Set Content-Type Header to : application/octet-stream
Set Body to : <image file>
```

Got it!

3. Video Analysis

- AI video Indexer
- AI Spatial Analysis

💡 Video Analysis in Azure (Azure AI Video Indexer)

- meta data

Azure AI Video Indexer is a service that automatically analyzes videos to extract insights, metadata, and searchable information using AI.

It combines vision, speech, and language models to understand content inside videos.

🧠 What It Does

e.g., keywords used etc.

It can:

- Detect faces, emotions, brands, objects, and spoken words
- Extract transcripts, translations, and speaker identification
- Create summaries, scene segmentation, and searchable indexes

📍 What is Spatial Analysis?

- live Data.

Spatial Analysis (part of Azure Video Analytics) uses computer vision to analyze live video feeds (CCTV, cameras) for movement and behavior in a physical space.

It can:

e.g., count of people in
the video walking

- Count people
- Detect social distancing or occupancy
- Track motion paths
- Trigger alerts for rule violations

👉 In short:

Spatial Analysis = Real-time understanding of how people or objects move in space.

🛠 Summary Table

Feature	Purpose
Video Indexer	Extracts insights from recorded videos
Indexer	Processes and indexes video metadata
Spatial Analysis	Analyzes live video streams for movement, counting, and zones

Check code on Microsoft

AI Vision

repository.

15 - 20 %

AI

LANGUAGE

SOLUTIONS

Implement natural language processing solutions (15–20%)

Analyze and translate text

- Extract key phrases and entities
- Determine sentiment of text
- Detect the language used in text
- Detect personally identifiable information (PII) in text
- Translate text and documents by using the Azure AI Translator service

Process and translate speech

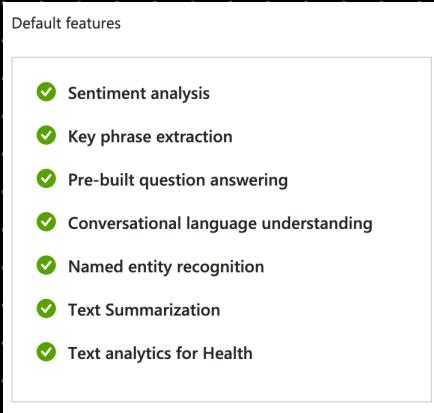
- Integrate generative AI speaking capabilities in an application
- Implement text-to-speech and speech-to-text using Azure AI Speech
- Improve text-to-speech by using Speech Synthesis Markup Language (SSML)
- Implement custom speech solutions with Azure AI Speech
- Implement intent and keyword recognition with Azure AI Speech
- Translate speech-to-speech and speech-to-text by using the Azure AI Speech service

Implement custom language models

- Create intents, entities, and add utterances
- Train, evaluate, deploy, and test a language understanding model
- Optimize, backup, and recover language understanding model
- Consume a language model from a client application
- Create a custom question answering project
- Add question-and-answer pairs and import sources for question answering
- Train, test, and publish a knowledge base
- Create a multi-turn conversation
- Add alternate phrasing and chit-chat to a knowledge base
- Export a knowledge base
- Create a multi-language question answering solution
- Implement custom translation, including training, improving, and publishing a custom model

1. AI Foundry - Language Services

Features



Azure AI | Language Studio

Try CLU and CQA in Azure AI Foundry—fine-tune your models faster with new features, agent integration, and a unified experience with all of Azure AI. See [CLU](#) and [CQA](#) for more information on configuring your resources in Azure AI Foundry.

Language Studio

your scenario.

AI Language Studio

Star Featured Extract information Classify text Understand questions and conversational language Summarize text Translate text

Use Natural Language Understanding (NLU) to detect the language or classify the sentiment of any piece of text you have. You can also classify your text documents by customizing a classification model over your dataset. [Learn more about custom text classification.](#)

Analyze sentiment and mine opinions
Detect positive, negative and neutral sentiment in text. Get more insights by mining opinions.
[Try it out](#)

Detect language
Evaluate text and detect a wide range of languages and variant dialects.
[Try it out](#)

Custom text classification
Train a classification model to classify text using your own data.
[Open Custom text classification](#)

Product Review (short)
I bought a size S and it fit perfectly. I found the zipper a...

Service Review (long)
Long wait... BUT FOR GOOD REASON. Some awesome Italian food ...

Customer Complaint Email
Hello, My name is Mateo Gomez and I visited Contoso...

Social Media Post
I can't describe my experience of this game in two words: "TEC..."

Employee survey feedback
The cafeteria's food is getting worse by the day. \$15 for a pl...

Run

Examine the results

Result JSON

Analyzed sentiment

Document sentiment

Mixed Confidence: 50.00%

50.00% 1.00% 49.00%
POSITIVE NEUTRAL NEGATIVE

Sentiment Analysis

Real World Use Cases ;

1. Extract Key Phrases
2. Extract NER - DII
3. Sentiment Analysis
4. Language Detection - Over 100 languages

2. Azure AI Speech - Speech - Whisper

Imp: Responsible AI in voice Generation;

e.g., Deepika becomes voice of Llama.
(with consent).

Text to speech - Speech studio.

1.

Speech to text

Quickly and accurately transcribe in more than 100 languages and dialects. Enhance the accuracy of your transcriptions by creating a custom speech model that can handle domain-specific terminology, background noise, and accents. [Learn more about speech to text](#)

The interface includes sections for Real-time speech to text, Whisper Model in Azure OpenAI Service, Batch speech to text, Custom Speech, Pronunciation Assessment with speech to text, and Speech Translation.

Real-time speech to text: Quickly test live transcription capabilities on your own audio without writing any code. Try out Real-time speech to text.

Whisper Model in Azure OpenAI Service: Quickly test live transcription capabilities on your own audio utilizing your Azure OpenAI resource and use prompts to improve the quality of the transcripts. Try out Whisper Model in Azure OpenAI Service.

Batch speech to text: Quickly test batch transcription capabilities to transcribe a large amount of audio in storage and receive results asynchronously using Azure Speech models or OpenAI Whisper model. Try out Batch speech to text.

Custom Speech: Add your own data and adapt to specific speaking styles, vocabulary, and more with a customized speech to text model. Start a Custom Speech project.

Pronunciation Assessment with speech to text: Get instant feedback on pronunciation accuracy and fluency by reading a script aloud. Try out Pronunciation Assessment.

Speech Translation: Translate speech into other languages of your choice with low latency. Try out Speech Translation.

2.

Text to speech

Build apps and services that speak naturally with more than 150 voices across 500 languages and dialects. Create a customized voice to differentiate your brand and use various speaking styles to bring a sense of emotion to your spoken content. [Learn more about text to speech](#)

The interface includes sections for Voice Gallery, Professional voice fine-tuning, Personal Voice, Audio Content Creation, Text to speech Avatar, and Custom Avatar.

Voice Gallery: Browse expressive voices with humanlike speech to find the perfect speaker for your project. Explore the Voice Gallery.

Professional voice fine-tuning: Use your own audio recordings to create a distinct, one-of-a-kind voice for your text-to-speech apps. Start a Professional voice fine-tuning project.

Personal Voice: Create an AI voice easily from a human voice sample, providing your users with a personalized voice experience across 100 languages. Try out Personal Voice. A red arrow points from this section to the handwritten note "Synthetic voice".

Audio Content Creation: Craft nuanced speech by adjusting the speaking style, pacing, and pronunciation of your spoken content. Start an Audio Content Creation project.

Text to speech Avatar: Bring text to life with natural-sounding voices and photorealistic talking avatars, creating a more engaging and delightful communication experience. Try out Text-to-speech Avatar.

Custom Avatar: Create a unique and realistic avatar or digital twin using your own video data. Ensure brand consistency while producing captivating videos or giving personality to your chatbots. Start a Custom Avatar project.

Synthetic voice
Neural voice
more Realistic

3. Speech Synthesis Markup Language - SSML

is an XML based markup language that allows to fine tune our text to speech attributes such as;

- pitch
- pronunciation
- Speaking Rate
- Volume.

Prosody - Element which controls 3 attributes

Prosody controls the tone, speed, and volume of a synthetic voice — just like how humans change their speaking rhythm and pitch to sound natural or expressive.

In SSML, the `<prosody>` tag lets you fine-tune these three key aspects:

- `rate` → how fast or slow the voice speaks
- `pitch` → how high or low the voice sounds
- `volume` → how loud or soft the voice speaks

xml

Copy code

```
<speak xmlns="http://www.w3.org/2001/10/synthesis" version="1.0" xml:lang="en-US">
<voice name="en-US-AriaNeural">
  <prosody rate="+15%" pitch="+4st" volume="loud">
    Hey there! I sound brighter and more energetic now.
  </prosody>
</voice>
</speak>
```

Viseme → Animation of mouth

4. Intent Recognition matching intent of user

e.g., setting up manual chatbot / Assistant

- Pattern Matching - Keywords

- CLU - conversational lang. understanding

↳ Speech to Text conversion

2. Text → AI language

More preferred → To understand intent

Drawback = CLU needs to be trained.

5. Keyword Recognition

e.g.: Hi Siri,
Hey Alexa.

Go to Speech Studio → Custom Keywords

The screenshot shows the Azure AI Speech Studio interface. On the left, there's a sidebar with 'Custom Keyword' selected, showing a model named 'Preppy Ai' (English (United States)). The main area is titled 'Create a custom keyword for your virtual assistant' and shows a table with one row: 'Name' (Preppy Ai), 'Model type' (Basic), 'Keyword' (HEY TJ), 'Created' (10/22/2025 9:36 PM), and 'Status' (Processing). A red circle highlights the 'HEY TJ' keyword entry.

5. Translate Language → Language studio

→ Text to Text (Transformer)

Feature	Synchronous	Asynchronous	Real-time
Speed	Instant	Delayed (batch)	Continuous
Input Size	Small text	Large text/files	Streaming data
Best For	Apps, UIs, chatbots	Document translation	Live meetings, captions
Storage Required	✗ No	✓ Yes (Blob Storage)	✗ No
Common Pairing	Text Translation API	Document Translation API	Speech Translation API

Resources

Language Studio (Document)

Translator → should have storage Blob
Storage Account Contributor Role

The screenshot shows the Azure AI Language Studio interface. On the left, there's a sidebar with 'Language Studio' selected. The main area is titled 'Create a translator' and shows a form for 'Initial Configuration'. The 'Resource type' dropdown is set to 'Translator'. At the bottom of the form, there's a link 'Create a new Translator resource in the Azure portal' which is circled in red. To the right of the form, there's a note: 'The above translator resource must have "Storage Blob Data" accounts. These may take a while.' A red arrow points from this note to the circled link. Another red arrow points from the note to the text 'Upload Docs and get the translation in PDF.'

Upload Docs and get the translation in PDF.

The screenshot shows the Microsoft Azure portal interface. On the left, under 'Azure role assignments', there is a note: 'Set up permission to Access Data in Storage Acc by Translator'. The 'Identity' tab is highlighted in red. On the right, the 'Add role assignment (Preview)' dialog is open, showing the 'Storage' scope, 'Azure subscription 1' subscription, 'ai102acc' resource, and 'Role' dropdown set to 'Select a role'. A search bar contains 'blo'. Below it, 'Storage Blob Data Contributor' is selected, while other roles like 'Storage Actions Blob Data Operator', 'Storage Blob Data Owner', 'Storage Blob Data Reader', and 'Storage Blob Delegator' are listed.

Q1 How to avoid Access Errors?

Go to storage Acc → Resource sharing (CORS)

The screenshot shows the 'Resource sharing (CORS)' settings for the 'ai102acc' storage account. The 'cors' tab is selected. It shows a table for 'Blob service' with columns: Allowed origins, Allowed methods, Allowed headers, Exposed headers, and Max age. Under 'Allowed origins', 'https://language.cognitive.azure.com' is listed. Under 'Allowed methods', all methods (DELETE, GET, HEAD, MERGE, POST, OPTIONS, PUT, PATCH) are checked. Under 'Exposed headers', 'Content-Type' is checked. Under 'Max age', '500' is specified.

Imp. For Translator job, user must have
blob storage contributor role to storage account.

The screenshot shows the 'Job history' section of the Language Studio. It lists a single job: 'c1188a76-c184-40ad-b4bd-b4dae24339d5 (1)' which 'Succeeded' on '10/23/2025, 03:19 PM'. The file 'sample_german_text.txt' is listed with status 'Succeeded', target language 'Hindi', and submitted date '10/23/2025, 03:19 ...'. A download icon is circled in red at the end of the row.

File is Ready in Downloadable Format.

Custom Translation ;

BLEU - Bi Lingual Evaluation Understudy

is a metric used to evaluate the quality of machine translated text.

Typical Interpretation:	
BLEU Score	Translation Quality
0.8 - 1.0	Excellent / Human-level
0.6 - 0.8	Good
0.4 - 0.6	Fair
0.2 - 0.4	Poor
0.0 - 0.2	Very Poor

Azure AI Speech - Speech Translation ;

Azure AI Speech Service;

Use Case Live Translation Service

↳ where voice from microphone
is directly translated

Speech to Text Translation

Translate to multiple languages

Simultaneously .

Imp. Use code from Azure Github

(Microsoft SDK).

6: Implement Custom Language Model.

Luis expired CLU new

In CLU, two of the most important concepts are

- **Utterance** → What the user says.
👉 Example: "Book a flight to Delhi."
- **Intent** → What the user wants to do.
👉 Example: BookFlight

In short:

Utterance = user's words

e.g., Yo Yo

Intent = user's goal

none

Entity extraction = identifying specific pieces of information in a user's utterance.

It helps pick out **details** like names, dates, places, etc.

Example:

Utterance: "Book a flight to **Delhi** on **Monday**."

- Intent: BookFlight
- Entities: Delhi (destination), Monday (date)

Activity pane

Changes are saved 4:48 PM

Add entity

- TypeOfFood (2) LEARNED
- Toppings (1) LEARNED
- Size (1) LEARNED
- Quantity (1) LEARNED

Manual Tagging for model learning

Imp: we can also create utterances using chatGPT and mark entities

Model performance Preview

Deployment = will get prediction link

Backup

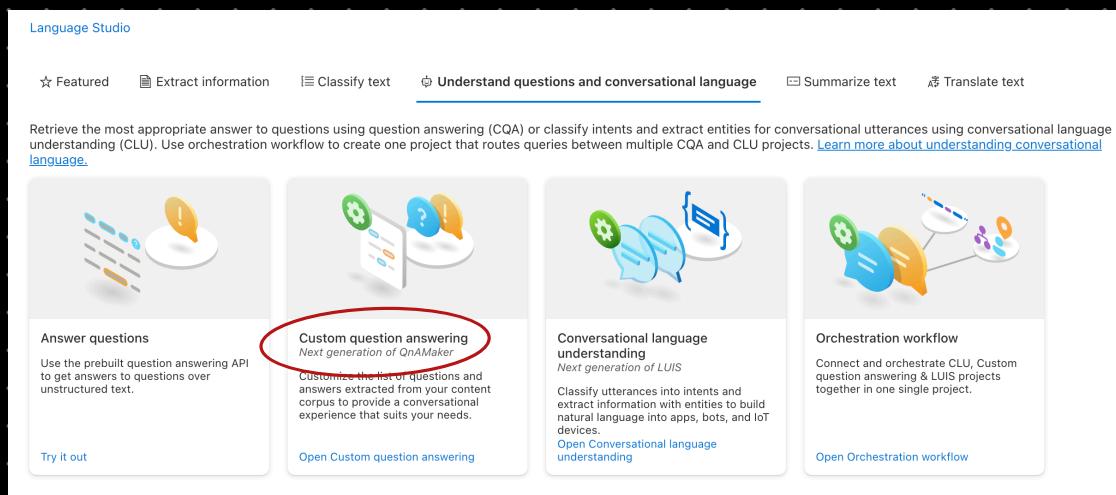
Create
RESTART API

&

Load
RESTART API

we can Test this model

7. Custom Q&A Bot ; uses Azure Search



Add Questions & Answers

we can add Q&A manually or with Q&A File.

Add a new question answer pair

Add a new question answer pair by entering a question, answer, and selecting the source.

Manual Q&A

Source *: Editorial

Question *: How many employees?

Answer *: 20

Add metadata (optional): + Metadata

Add follow up prompts (optional): + Follow up prompt

Done | **Cancel**

Manage sources

URLs (highlighted with a green circle)

Files (highlighted with a green circle)

Chatbot demo - Manage sources

How much is a pizza?
Our pizzas average \$16 per pie. We use only the finest and freshest ingredients.

Are you open late?
We're open until midnight on most weekdays, and until 2AM on Friday and Saturday night.

What if I want to work there?
Come in and speak to our manager Frank if you'd like to apply for a job.

Do you have tables for seating inside the restaurant?
Yes, we have a few tables for small and medium-sized parties up to 8 people. Please make a reservation in advance if you don't want to wait for available seats.

What is your opinion about pineapple on a pizza?
We do not allow such controversial subjects on this chatbot. This is your first warning.

Q11 How to use it as API ?

Get prediction URL

Use these sample requests as a starting point to call your model's endpoint

Prediction URL: <https://ifirstprojectcda.cognitiveservices.azure.com/language/query-knowledgebases?api-version=2021-10-01&deploymentName=test>

Sample request

```
curl -X POST https://ifirstprojectcda.cognitiveservices.azure.com/language/query-knowledgebases?api-version=2021-10-01&deploymentName=test -H "Ocp-Apim-Subscription-Key: 1JUQjP9BJA977RjwcvBrUzswtIewAAmlc8d831la9HSWzrR6Vg90ySvCjYebfRjX3w3AAaACQObJlw" -H "Content-Type: application/json" -d '{"topAnswersWithSpan":1,"question": "QUESTION_HERE","includeUnstructuredSources":true,"confidenceScoreThreshold":1,"your_score_threshold_here":1,"answerSpanRequest":1,"enable":false,"topAnswersWithSpan":1,"confidenceScoreThreshold":1,"your_score_threshold_here":1,"filter":1,"metadataFilter":1,"logicalOperation":1,"YOUR_LOGICAL_OPERATION_HERE":1,"metadata":1,"key":1,"YOUR_ADDITIONAL_PROP_KEY_HERE":1,"value":1,"YOUR_ADDITIONAL_PROP_VALUE_HERE":1}}'
```

Close

This can be in our programs as REST API!

Imp. Think of this Q&A as manual chatbot but one thing is confidence level is decreased it may come up with answers (less accurate)

Q. How to Deploy as Chatbot?

The screenshot shows the Azure AI portal with a modal dialog titled "Deploy this project?" asking if the user wants to deploy the project. Below the modal, a message says "Deploying your knowledge base will copy the knowledge base from the test index to the production index." To the right, the deployment status is shown as "Your knowledge base is now deployed. You can get your prediction URL or create a bot." It includes details like State: Deployed, Deployment Date: 10/23/2025, Location: eastus, Deployment Time: 10:04:05 PM, and Tier: Standard (S). A "Create a bot" button is also present.

Azure Bot Service Test

The screenshot shows the Azure Bot Service Test interface. On the left, there's a sidebar with options like Overview, Activity log, Access control (IAM), Tags, Settings (Bot profile, Configuration, Channels, Pricing), Test in Web Chat (selected), Encryption, Networking, Properties, Locks, Monitoring, and Conversational analytics. The main area shows a conversation in "Test" mode. The bot says "Hello and Welcome". The user asks "How much does pizza cost?", and the bot replies "I'm sorry! I don't know the answer." The user then types "Did you mean:" followed by "How much is 'A Pizza?'", "How much is a pizza?", and "None of the above". The input field at the bottom says "Type your message".

Q. How to create multi-turn conversation?

More like adding follow up on Q & A

The screenshot shows the Azure AI Language Studio. On the left, there's a list of question-answer pairs under "Question answer pairs (4)". One pair is highlighted with a red circle, showing the question "Any Suggestion for Pizza ?" and the answer "Can u tell us more about your taste, Spicy Pizza". A "Follow up" button is visible next to the answer. A modal dialog titled "Follow-up prompt" is open, showing the text "Text displayed in the prompt to the user" with "Spicy" entered. Below it, "Answer: Spicy Pizza" is listed. A checkbox "Show only in contextual flow?" is checked. The "Add prompt" button is at the bottom. To the right, the "Test" pane shows a conversation where the user asks "Any Suggestion for Pizza ?" and the bot replies "Can u tell us more about your taste, Spicy Pizza". The "Follow up" button is circled in red. The "Tangy" option is also visible in the list.

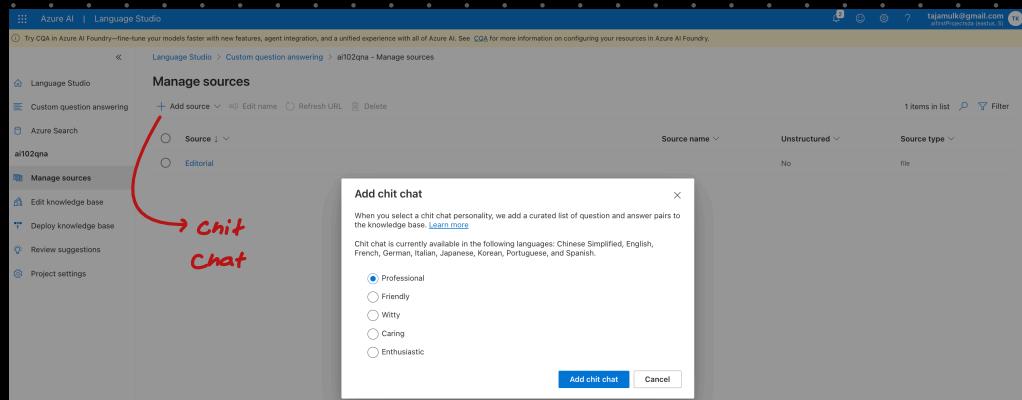
Imp: we can also add alternate Questions

Adding Chit Chat ;

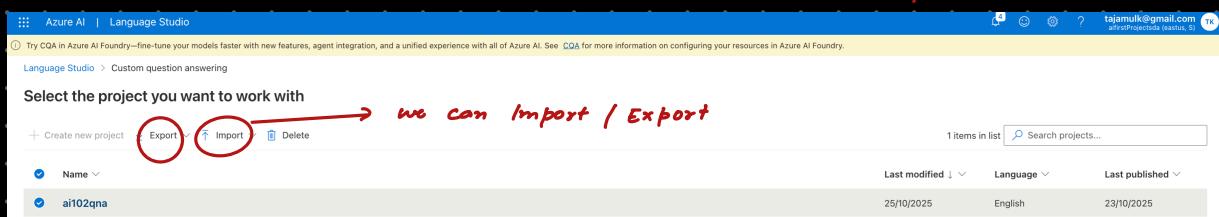
is more like adding basic chit chat capabilities to the Q&A chatbot.

e.g., Hi → Hey

How are you → I'm fine



Q, How to Reuse this Chatbot?



Multi Language BOT

Make multiple chatbots



Map Apps / users to different languages

15 - 20 %

KNOWLEDGE
MINING
&
INFORMATION
EXTRACTION

Implement knowledge mining and information extraction solutions (15–20%)

Implement an Azure AI Search solution

- Provision an Azure AI Search resource, create an index, and define a skillset
- Create data sources and indexers
- Implement custom skills and include them in a skillset
- Create and run an indexer
- Query an index, including syntax, sorting, filtering, and wildcards
- Manage Knowledge Store projections, including file, object, and table projections
- Implement semantic and vector store solutions

Implement an Azure AI Document Intelligence solution

- Provision a Document Intelligence resource
- Use prebuilt models to extract data from documents
- Implement a custom document intelligence model
- Train, test, and publish a custom document intelligence model
- Create a composed document intelligence model

Extract information with Azure AI Content Understanding

- Create an OCR pipeline to extract text from images and documents
- Summarize, classify, and detect attributes of documents
- Extract entities, tables, and images from documents
- Process and ingest documents, images, videos, and audio with Azure AI Content Understanding

1. **AZURE AI SEARCH** is a cloud based search service by Microsoft

- Supports Keyword BM25 - Exact & vector search HNSW - Approx

In short ↗

It helps you index, search, and analyze large amounts of structured and unstructured data (like PDFs, images, or documents) — enhanced with AI features such as OCR, NLP, and semantic search.

Think of it as:

→ Google-like search for your own data — powered by Azure AI.

Example use cases:

- Searching documents with meaning (not just keywords)
- Extracting info from resumes, reports, or PDFs
- Building enterprise knowledge bases

Capabilities :

- Chunking
- Vectorization
- Indexing
- Retrieving

Process

- Index Data
- Querying
- Results based on Semantic Rank

Imp.

With Azure AI Search, you're essentially building a custom search engine on top of your own data. Using the provided endpoint and API keys, your applications can interact with and query your data intelligently through Azure's AI-powered search capabilities.

Create AI Search Resource

The screenshot shows the Microsoft Azure portal with the search bar set to "Search resources, services, and docs (G+)" and the user "tajamulk@gmail.com" logged in. The main page displays a search service named "aifirstprojectsda-as3aiwewqygzane". The "Overview" tab is selected, showing basic details like the resource group "ai102", location "East US", and subscription "Azure_subscription_1". A red circle highlights the "+ Add index" button. The "Essentials" section lists properties such as URL, Pricing tier, Replicas, Partitions, Search units, and Compute type. The "JSON View" link is visible at the bottom right of the table.

Property	Value
Resource group	: ai102
Location	: East US
Subscription	: Azure_subscription_1
Subscription ID	: f153f722-aa61-4505-bdcb-9dfb5b3e4294
Status	: Running
Date created	: Oct 22, 2025, 11:50:49 AM
Url	: https://aifirstprojectsda-as3aiwewqygzane.search.windows.net
Pricing tier	: Basic
Replicas	: 1 (No SLA)
Partitions	: 1
Search units	: 1
Compute type	: Default

Data Sources ; Azure AI Search can index data from sources like

- Azure SQL
- Blob storage
- Data lake
- Cosmos DB
- Sharepoint

Supported formats include: by AI Search

- Structured data → SQL tables, CSV
- Semi-structured → JSON, XML
- Unstructured → PDFs, Word, Excel, PowerPoint, TXT, HTML, images (via OCR)

Imp: Json is the format used to push data via API. — Take and send data as JSON

Add Data Sources

The screenshot shows the Azure portal interface for managing a search service. The top navigation bar includes 'Microsoft Azure', 'Upgrade', 'Search resources, services, and docs (G+)', 'Copilot', and user information. Below the navigation is a search bar and a summary card for the search service. The main content area shows the service details: Resource group (ai102), Location (East US), Subscription (Azure subscription 1), Subscription ID, Status (Running), Date created (Oct 22, 2025, 11:50:49 AM). A red circle highlights the 'Import data' button in the top navigation bar.

The screenshot shows the 'Add data source' form in the Azure portal. The form fields are: Data Source (Azure Blob Storage), Name (azureblob-1761470865905-datasource), Subscription (Azure subscription 1), Storage account (sttajamulk98812340277953), Blob container (c4d40440-f509-4457-98bd-1b75ba8ec409-azureml-blobstore), and Blob folder (your/folder/here).

Azure AI Index Search Optimizer

adds unique keys (IDs) to data records to make them searchable and retrievable efficiently

Components

Field name	Type	Retrievable	Filterable	Sortable	Facetable	Searchable
id	String	<input type="checkbox"/>				

- **Retrievable** = what is shown (text)
- **Searchable** = what is searched (keywords)
- **Facetable** = ability to drill down

Q: Searchable vs Retrievable

- Searchable fields → are the ones Azure AI Search looks inside when you type a query.
Example: If "course_title" is searchable, typing "Azure" will return any record whose title contains "Azure."
- Retrievable fields → are the ones returned in the search results (i.e., what you see after the search).
Example: The API might return "course_title," "course_id," "description" — these are retrievable fields.

So yes ✓

When you search "Azure courses," Azure AI Search scans all *searchable fields* for that word, and then retrieves data (like course name, ID, etc.) from *retrievable fields* in the results.

3. Create Index

Manual Task takes lots of time so it is preferred to use Auto indexer

The screenshot shows the Azure portal interface for a search service named 'aifirstprojectsda-as3aiwewqygzane'. The 'Overview' tab is selected. A red circle highlights the 'Import data (new)' button in the top navigation bar, with a handwritten note 'use new' pointing to it. The main content area displays various service properties like location, subscription, and status.

The screenshot shows the 'Keyword search' wizard. The 'Optional Enrichments' section is highlighted with a red arrow and the handwritten note 'also klos sk'. The 'Apply AI enrichments' step is circled in red, with the handwritten note 'Skills' next to it. The wizard provides options for connecting data, applying AI enrichments, and previewing mappings.

Home > Resource Manager | Recent resources > aifirstprojectsda-as3aiwewqygzane >

Keyword search ...

Auto Index Fields

We're building a better way to search. This wizard makes it easier to configure keyword search. [Share your feedback](#) or return to the classic view.

Switch to classic view

Connect to your data

Apply AI enrichments

Preview mappings

Advanced settings

Review and create

Preview index fields

Review the fields that will be included in your index. Index fields can obtain values from your data source and from skill outputs if you added enrichments in the previous step. [Learn more](#)

Add field Add subfield Configure field Reset Delete Hide auto-generated fields

Source column	Target index field name	Target index field type
content	content	Edm.String
metadata_storage_name	title	Edm.String
metadata_storage_path	id	Edm.String
content	keyPhrases	Collection(Edm.String)

Home > Resource Manager | Recent resources > aifirstprojectsda-as3aiwewqygzane > Keyword search ...

Semantic Ranking Synchronize

We're building a better way to search. This wizard makes it easier to configure keyword search. [Share your feedback](#) or return to the classic view.

Switch to classic view

Connect to your data

Apply AI enrichments

Preview mappings

Advanced settings

Review and create

Enable semantic ranker

Schedule indexing

Schedule how often the indexer runs to keep your index synchronized with your source data. [Learn more](#)

Schedule

Once

Once

5 minutes

10 minutes

30 minutes

Hourly

Daily

Custom

Home > Resource Manager | Recent resources > aifirstprojectsda-as3aiwewqygzane > Keyword search > search-1761475200015 ...

Testing AI Search

Took milliseconds

Save Discard Refresh Create demo app Edit JSON Delete Encryption

Documents Total storage Vector index quota usage Max storage

0 0 Bytes 0 Bytes 15 GB

Search explorer Fields CORS Scoring profiles Semantic configurations Vector profiles

Query options View

smart

Results

```

1 {
2   "@odata.context": "https://aifirstprojectsda-as3aiwewqygzane.search.windows.net/indexes('search-1761475200015')/s
3   "@odata.count": 1,
4   "@search.answers": [
5     {
6       "key": "aHR0cHM6Ly9zdHRhamFtdWx0Tg4MTIzNDAyNzc5NTMuYmxvYi5jb3JlLndpbmRvd3MubmV0L2M0ZDQwNDQwLWY1MDktNDQ1Ny05C
7       "text": "Smart Data Dashboard Overview The Smart Data Dashboard helps businesses visualize real-time KPIs ar
8       "highlights": "Smart Data Dashboard Overview The<em> Smart Data Dashboard </em>helps<em> businesses visualiz
9       "score": 0.889999856948853
10    }
11  ],
12  "value": [
13    {
14      "@search.score": 27.556522,
15      "@search.rerankerScore": 2.3929290771484375,
16      "@search.captions": [
17        {
18          "text": "Smart Data Dashboard Overview The Smart Data Dashboard helps businesses visualize real-time KPIs ar
19          "highlights": "Smart Data Dashboard Overview The<em> Smart Data Dashboard </em>helps<em> businesses visu

```

What is a Skillset in Azure AI Search?

A skillset is a collection of AI enrichment skills that process your raw data (like PDFs, images, or text) and add intelligent insights before indexing.

These skills can come from three main sources:

1. Built-in cognitive skills (from Azure Cognitive Services)
2. Custom skills (your own API or logic)
3. Azure OpenAI skills (semantic & generative enrichment)

1 Built-in Cognitive Skills

These are ready-to-use skills that come with Azure Cognitive Services.

Category	Examples of Built-in Skills	Purpose
Text Processing	Text extraction, language detection, key phrase extraction	Turn unstructured text into searchable data
Vision	OCR (optical character recognition), image analysis	Extract text and metadata from images or PDFs
Entity Recognition	Named entity recognition, PII detection	Identify people, places, dates, org names
Language	Sentiment analysis, translation	Analyze tone or convert to other languages
Knowledge Mining	Custom entity lookup, skill chaining	Combine skills for deeper enrichment

2 Custom Skills

These are **your own APIs or ML models** that Azure can call during enrichment.

For example:

- A REST API that classifies documents by topic
- A custom ML model that detects fraud patterns in text
- A logic app that runs business rules before indexing

💡 You define them in the skillset using an **endpoint URL**, and Azure passes your document data to them for processing.

API that indexer
calls to. e.g., Azure Fns
Azure web apps

3 Azure OpenAI Integration (Advanced Skill)

You can now integrate Azure OpenAI into your skillsets using **semantic and generative enrichment**:

- **Semantic search** → understands meaning, not just keywords
- **Generative skill** → uses GPT models (like GPT-4) to summarize, tag, or extract insights

Example: Summarize a long report into key points before indexing

In short (interview-ready one-liner):

A skillset in Azure AI Search combines **built-in AI skills**, **custom APIs**, and **Azure OpenAI models** to enrich raw data with insights like key phrases, entities, sentiment, and summaries — turning unstructured content into powerful, searchable knowledge.

Query on Index

The screenshot shows the Azure AI Search service interface. At the top, there's a navigation bar with 'Home > aifirstprojectsda-as3aiiewqygzane > Overview'. Below it, a toolbar has items like 'Search', 'Add index', 'Import data', 'Import data (new)', 'Search explorer' (which is circled in red), 'Upgrade', 'Refresh', 'Delete', and 'Move'. A red arrow points from the text 'Use this' to the 'Search explorer' button. The main area shows 'Essentials' details: Resource group (move) : ai102, Location (move) : East US, Subscription (move) : Azure subscription_1, Subscription ID : f153f722-aa61-4505-bdcb-9dfb5b3e4294, Url : https://aifirstprojectsda-as3aiiewqygzane.search.windows.net, Pricing tier : Basic, Replicas : 1 (No SLA), Partitions : 1. Below this is a 'Search explorer' tab, followed by 'Fields', 'CORS', 'Scoring profiles', 'Semantic configurations', and 'Vector profiles'. A search bar at the top of the 'Search explorer' section contains the word 'smart', which is also highlighted in the results list. The results list shows JSON output for a single document, with the score '0.889999856948853' underlined.

How to Integrate or Create a search App?

The screenshot shows the Azure AI Search service interface. At the top, there's a navigation bar with 'Home > aifirstprojectsda-as3aiiewqygzane > Indexes'. Below it, a toolbar has items like 'Search', 'Add index', 'Import data', 'Import data (new)', 'Search explorer' (which is circled in red), 'Upgrade', 'Refresh', 'Delete', and 'Move'. A red arrow points from the text 'Create demo app' to the 'Create demo app' button in the toolbar. The main area shows 'Essentials' details: Resource group (move) : ai102, Location (move) : East US, Subscription (move) : Azure subscription_1, Subscription ID : f153f722-aa61-4505-bdcb-9dfb5b3e4294, Status : Running, Date created : Oct 22, 2025, 11:50:49 AM UTC, Tags (edit) : Add tags. Below this is a 'Indexes' section (circled in red with a double-headed arrow). The 'Search explorer' tab is selected, showing a search bar with 'Search' and a results table with one item. The results table shows the same JSON output as the previous screenshot, with the score '1.6289719' underlined.

3. Azure Document Intelligence

Form
recognizer

Convert Documents into usable Data

e.g., Invoices → Data

Aim; extract text, key-value pairs
tables and structures from Docs.

Capabilities; extract handwritten &
printed both

Custom Models

- **Classification model** to determine what type of document it is
- **Custom neural models** - trained on a large collection of documents
- **Custom template models** - consistent visual template
- **Custom composed models** - use a collection of models

Imp. No locked PDF's

Try Document Intelligence

Document analysis
Extract text, tables, structure, key-value pairs, and named entities from documents.

OCR/Read: Extract printed and handwritten text from images and documents, and turn unsearchable PDFs into searchable PDFs.
Layout: Extract tables, check boxes, and text from forms and documents.
General documents: Extract key value pairs and structure like tables and selection marks from any form or document.
Prebuilt models:

- Invoices: Extract invoice ID, customer details, vendor details, ship to, bill to, total tax, subtotal, line items and more.
- Receipts: Extract time and date of the transaction, merchant information, amounts of taxes, totals and more.
- Identity documents: Extract name, expiration date, machine readable zone, and more from passports and ID cards.
- US health insurance cards: Extract insurer, member, prescription, group number and more information from US health insurance cards.

Custom models: Train custom models to classify documents and extract text, structure and fields from your forms or documents.

Custom models
Train custom models to classify documents and extract text, structure and fields from your forms or documents.

Custom extraction model: Label and build a custom model to extract a specific schema from your forms and documents.
Custom classification model: Build a custom classification model to split and classify documents.

Invoices

Document Intelligence Studio > Prebuilt

Prebuilt Invoices ▼

API version: 2024-11-30 (4.0 General Availability) Service resource: docintel1209

Fields Content Result Code

Prebuilt invoice

DocType: invoice

- AmountDue #1 95.00% USD 610
- BillingAddress #1 123 Bill St, Redmond WA, 98052 HouseNumber 123 Road Bill St PostalCode 98052 City Redmond

Drag & drop file here or Browse for files or Fetch from URL

Run analysis Query fields Analyze options

Receipts

Azure AI | Document Intelligence Studio

Document Intelligence Studio > Prebuilt

Prebuilt Receipts ▼

API version: 2024-11-30 (4.0 General Availability) Service resource: docintel1209

Fields Result Code

DocType: receipt.retailMeal

- CountryRegion #1 USA 99.50%
- Items (2) #1
 - Content 2 Surface Pro 6 \$1,998.00 Description Surface Pro 6 Quantity 99.40%
 - 2 TotalPrice USD 1998
- Content 3 Surface Pen \$299.97 Description Surface Pen Quantity 99.40%

Drag & drop file here or Browse for files or Fetch from URL

Run analysis Query fields Analyze options

Contracts

Document Intelligence Studio > Prebuilt

Prebuilt Contracts ▼

API version: 2024-11-30 (4.0 General Availability) Service resource: docintel1209

Fields Content Result Code

DocType: contract

- Title #1 WEB HOSTING AGREEMENT 99.90%
- EffectiveDate #1 2022-10-15
- Parties (2) #1
 - Region Washington
 - Clause
- Jurisdictions #1
 - Region Washington
 - Clause

This Agreement shall be governed by and construed in accordance with the internal laws of the State of Washington, without regard to any conflicts of law principles.

WEB HOSTING AGREEMENT

This web hosting agreement is entered into as of the 1st day of October 2022. Wherein Contoso Ltd and its business operations (hereinafter referred to as "Contoso") are the provider of web hosting services and the customer (hereinafter referred to as "Customer") is the subscriber of such services.

Contoso must provide a minimum term of service of 12 months. The Customer may cancel the service at any time with a 30-day notice period. Both parties shall provide the same 24 hour service at the breakaway location.

Both parties shall abide by all laws and regulations made applicable to the performance of this contract.

Contoso Corporation
By Angel Brown
Date: 10/01/2022
Address: 123 Main Street, Redmond, WA, 98052

Admiral Works Corp
By Aaron Smith
Date: 10/01/2022
Address: 456 Elm Street, Seattle, WA, 98101

10 - 15 %

AZURE

GEN

AI

SOLUTIONS

Turning Test whether content written is

- Human written or AI

Imp: Gen AI models can easily pass this Test

Chat GPT → Nov 2022

Imp: Open AI provides REST API's so people can use these LLM's

Models = GPT → Mostly text to text
Whisper → Mostly Audio
Embedding → Vector
Dall-E → Video / Image

Imp: GPT model's like ChatGPT don't have Chat built in

💬 So how does chat actually work?

To make a conversation feel like a chat:

1. You maintain a **thread** (a list of past messages).
2. Each message has:
 - `role` : system, user, or assistant
 - `content` : the text of that message
3. Every new query sends the **entire conversation history** (or a truncated version) back to the model.
4. The model predicts the next logical response — that's what you see as a "reply."

So the "memory" is **simulated context**, not internal storage.

⚙️ Example (Simplified)

json

Copy code

```
[  
  {"role": "system", "content": "You are a helpful data science tutor."},  
  {"role": "user", "content": "What is regression?"},  
  {"role": "assistant", "content": "Regression is a statistical method to predict continuous values based on one or more independent variables."},  
  {"role": "user", "content": "Explain it like I'm 10."}  
]
```

→ **Tokens**

- Azure OpenAI works with tokens not directly with text
- Tokens are words or chunks of characters
- "Hamburger" => "ham", "bur" and "ger"
- Limitation on the number of tokens you can use as input
- Image inputs also take tokens

Imp:
?

→ **Prompt Engineering**

is the process of engineering inputs
or commands to LLM models

→ **RAG**

Retrieval Augmentative Generation

→ combine personal + LLM models
Data

→ **Fine Tuning** → PEFT

is like creating an entirely customized
model

Imp:
? Our preference should always be
getting job done with

RAG + Prompts

because Fine Tuning is a tiring
Job,
takes time.
Costly.

Azure AI Foundry previously AI Studio

Imp.

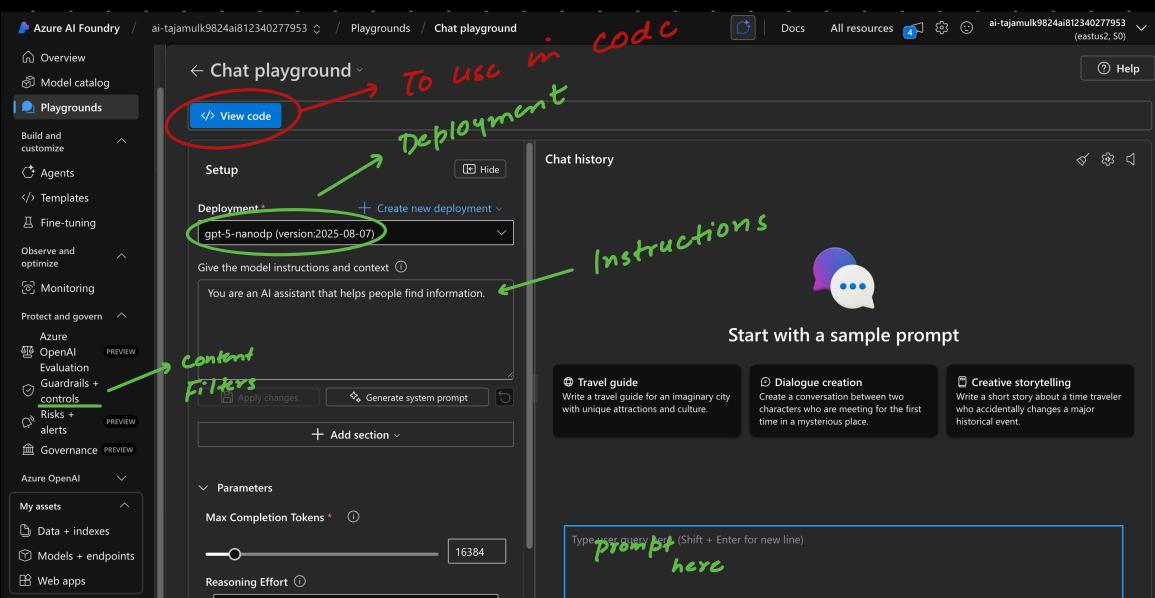
Q: Difference b/w Model & Deployment?

In Azure OpenAI, a **model** is the base AI system (like GPT-4), while a **deployment** is your **own configured instance** of that model — with a unique name, endpoint, and settings for your applications.

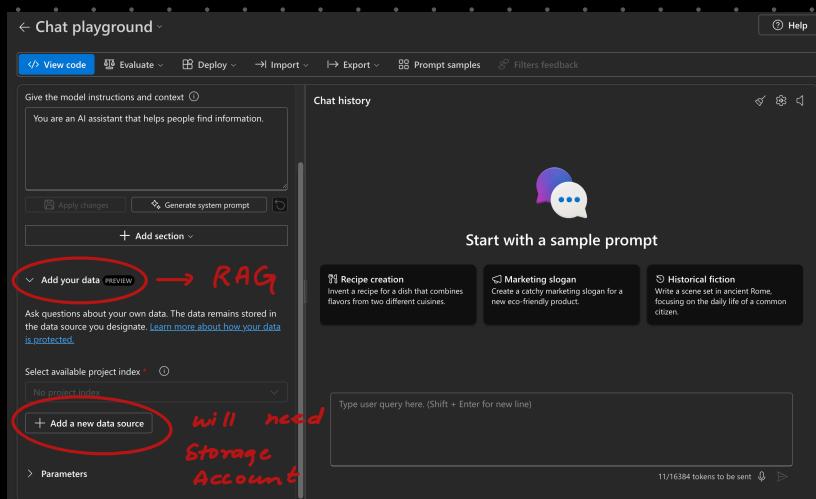
Model = Base model

Endpoint = Own Configured Instance
of the Base model

Model Deployment in Playground



Add Your Data → RAG.



Important Points

1. **Turing Test**; whether content created is AI Generated or Human written
2. **SSML**; Speech synthesis markup lang.
built on top of XML
3. **PROSODY**; Element which defines quality of language.
There are 3 attributes
 - volume
 - Rate
 - pitch
4. **BLEU SCORE**; Evaluation score of machine translations (0 - 1)
 $> 0.6 = \text{Good}$
5. **CLU** = Conversational Lang. Understanding
 - Two parts
 1. Utterance - Query
 2. Intent - Action
6. Skills of AI = AI Enrichments
7. **AI Index** = chunking
vectorization
Indexing
→ Semantic Ranking
8. **Viseme** = Lip shape when you speak
→ (AI speaks)