

AI-3008: Extract insights from visual data on Azure

Course Description

This 1-day course focuses on building intelligent applications that can see, interpret, and reason over images and documents using different multimodal models and agent-based tools. Learners explore how visual and document inputs can be combined with language models to enable structured extraction, analysis, and decision-making workflows. The course emphasizes practical patterns for extracting information, orchestrating tools, and grounding model responses in visual data.

Course Duration

1 day

Course Outline

1 - Develop a vision-enabled generative AI application

- Use a vision-capable model in the Microsoft Foundry portal
- Develop a vision-based chat app
- Module assessment

2 - Generate images with AI

- What are image-generation models?
- Explore image-generation models in Microsoft Foundry portal
- Create a client application that uses an image generation model
- Module assessment

3 - Generate videos with Microsoft Foundry

- Deploy a video generating model
- Generate video from a prompt
- Generate video in Python
- Module assessment

4 - Analyze images with Content Understanding

- What is Content Understanding?
- Analyze images with Content Understanding
- Module assessment

5 - Create a multimodal analysis solution with Azure Content Understanding

- What is Azure Content Understanding?
- Create a Content Understanding analyzer
- Use the Content Understanding API
- Module assessment

6 - Create an Azure Content Understanding client application

- Prepare to use the AI Content Understanding API
- Create a Content Understanding analyzer
- Analyze content
- Module assessment

7 - Extract data with Azure Document Intelligence

- What is Azure Document Intelligence?
- Use the Document Intelligence Studio
- Use prebuilt models

- Train and use custom models
- Module assessment

8 - Create a knowledge mining solution with Azure AI Search

- What is Azure AI Search?
- Extract data with an indexer
- Enrich extracted data with AI skills
- Search an index
- Persist extracted information in a knowledge store
- Module assessment

Who Should Attend

This course is designed for developers, AI engineers, and technical professionals who want to build applications that work with images and documents using multimodal, agent-driven approaches. It's best suited for learners with basic programming experience and a general understanding of cloud or AI concepts.