Microsoft

AI-900

Document Intelligence and Knowledge Mining

© Copyright Microsoft Corporation. All rights reserved.

AI-900 Agenda

1: AI Overview

2: Computer Vision

3: Natural Language Processing

Lab 4: Document Intelligence and Knowledge Mining ←

Lab 5: Generative AI

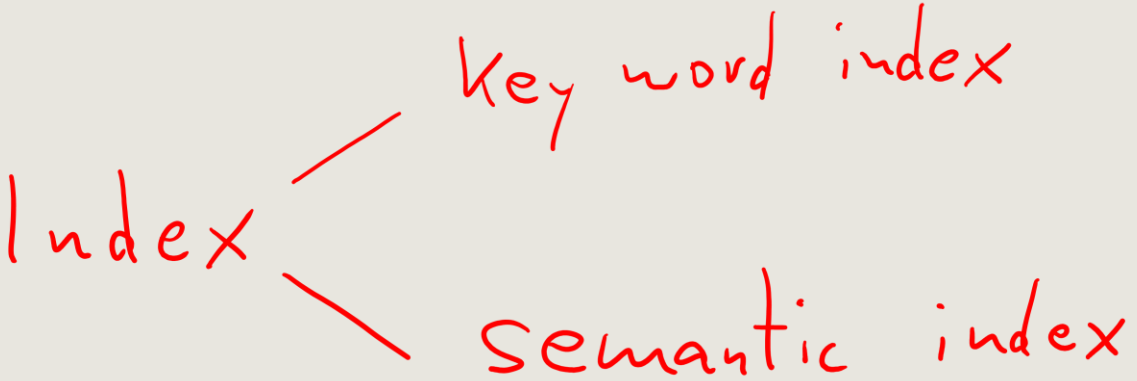
LLM
GPT
Copilots

RAG
Retrieval Augmented Gen.

© Copyright Microsoft Corporation. All rights reserved.

LP Agenda

- Azure AI Document Intelligence
- Azure AI Search

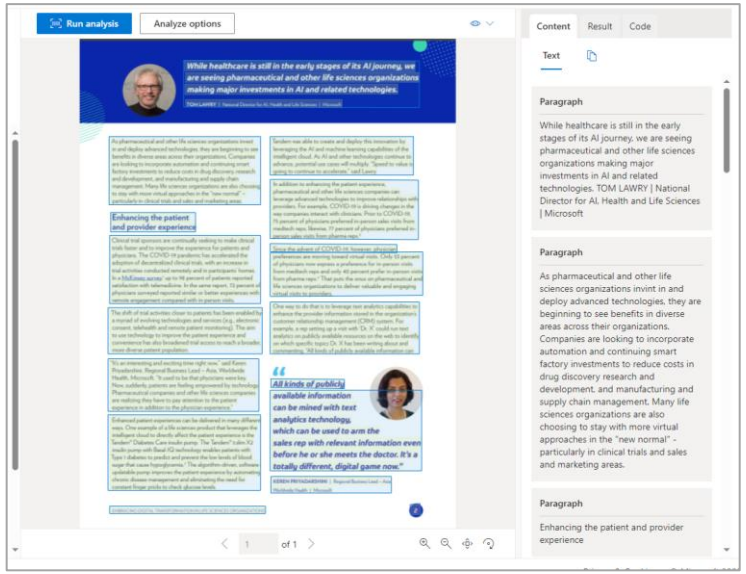


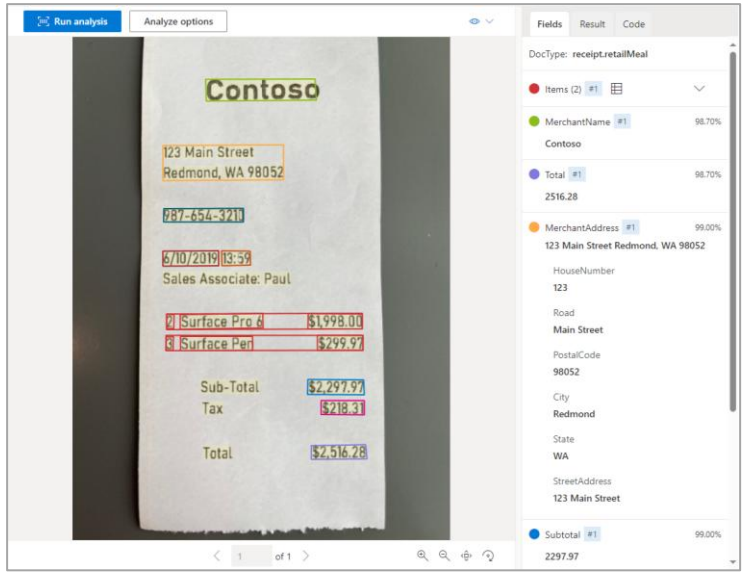
© Copyright Microsoft Corporation. All rights reserved.

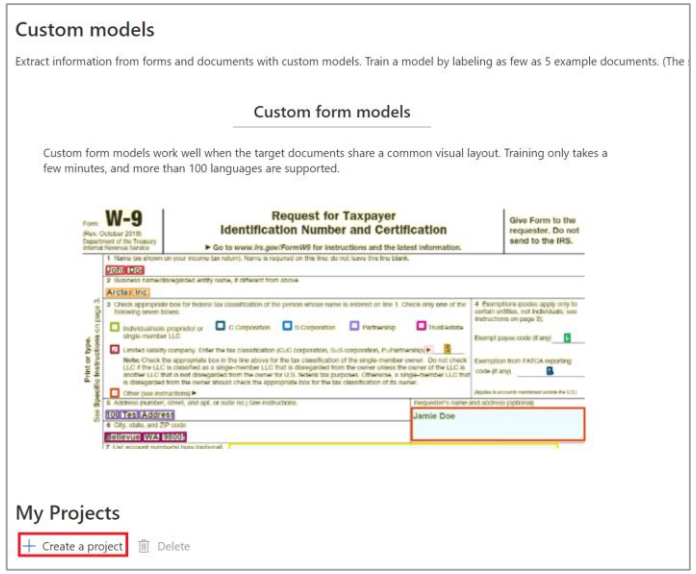
Azure AI Document Intelligence

© Copyright Microsoft Corporation. All rights reserved.

AI Document Intelligence services







- **Document analysis**
  - Returns structured data representations.
  - Regions of interest and relationships.
  - Configure **Analyze options** for free and chargeable analysis
- **Prebuilt models**
  - Invoices
  - Receipts
  - ID
  - Recognizes and extracts key-value pairs.
- **Custom models**
  - Train models with at least five sample data.
  - Identify fields of interest to your organization.

© Copyright Microsoft Corporation. All rights reserved.

Analyzing forms with the *Document Intelligence s*ervice

Extract information from scanned forms in image or PDF format.

- Use the pre-trained models for common document types, such as invoices, receipts, IDs, etc.
- Train a custom model using your own forms.

Models perform *semantic recognition* of form fields – not just text extraction.

Forms Recognizer

© Copyright Microsoft Corporation. All rights reserved.

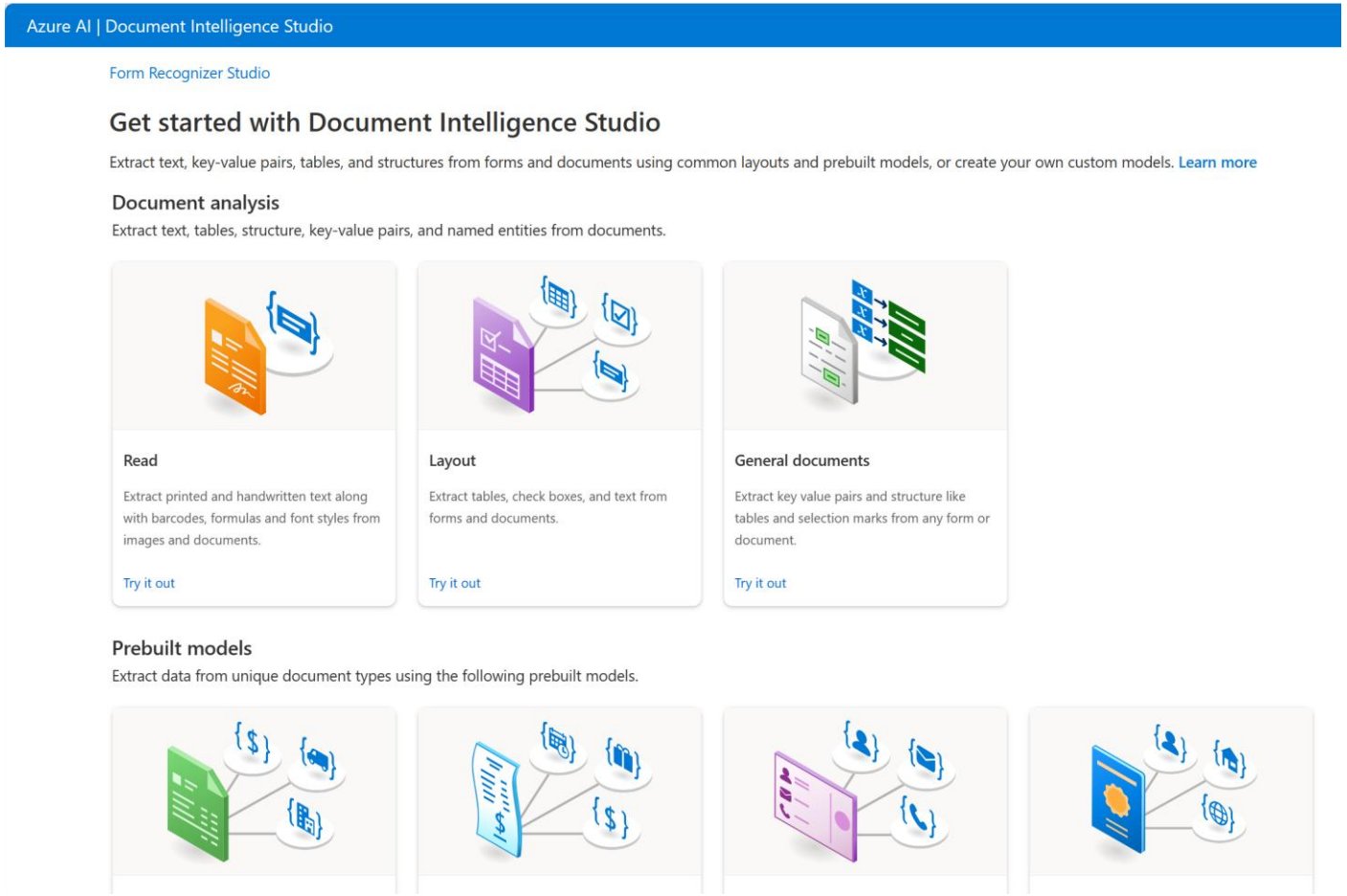
Document Intelligence Studio

Getting started with Document Intelligence

- Using a **no code** approach you can explore the functionality using samples and your own documents.

- First - **create a resource**
  - Document Intelligence resource
  - AI services resource
- Then **Enable** the resource in Document Intelligence Studio.

- **Getting Started** page: select a model to try.



© Copyright Microsoft Corporation. All rights reserved.

Exercise: Extract form data in Document Intelligence Studio

In this exercise, you will explore the **Azure AI Document Intelligence** service to analyze a receipt.

1. Use the hosted environment and Azure credentials provided for this exercise.

2. The instructions are also available on Learn: https://aka.ms/ai900-document-intelligence

© Copyright Microsoft Corporation. All rights reserved.

Azure AI Search

© Copyright Microsoft Corporation. All rights reserved.

What is knowledge mining?

Organizations have a lot of content

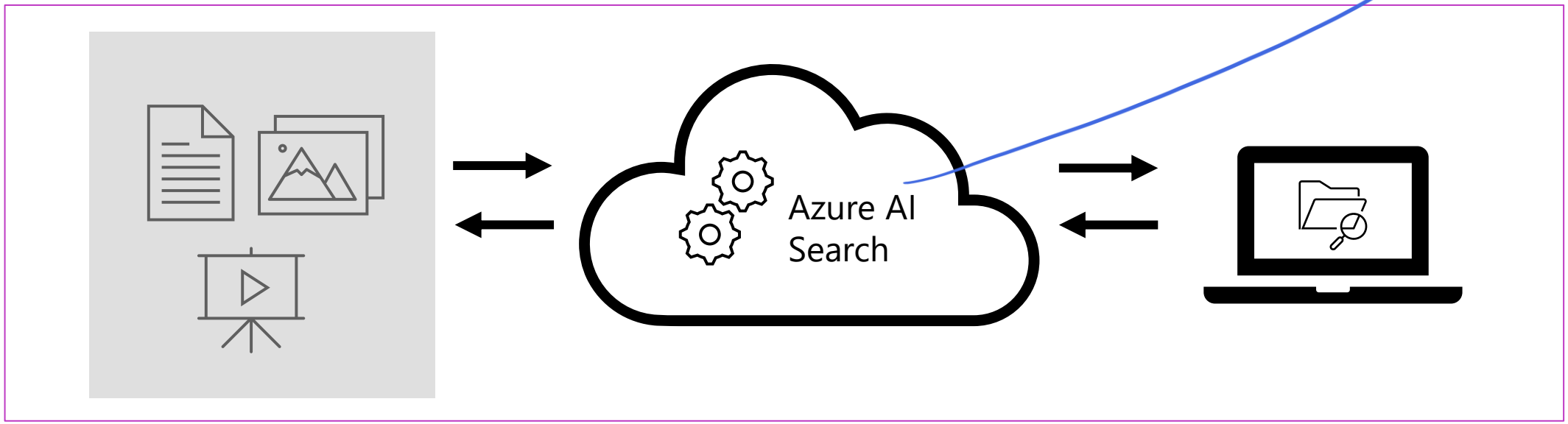
- Data is *locked away* in documents, PDFs, hand-written notes, etc.
- Time consuming and labour intensive to find data.
- Knowledge mining finds insights – *at scale*.

- **Azure AI Search** is Azure's AI-powered knowledge mining platform.
- Solutions provide access through, for example:
  - Search, eg bots to answer questions
  - Dashboards
  - Business applications
  - Availability for further analysis

© Copyright Microsoft Corporation. All rights reserved.

M365 Copilot → RAG (Prompt Anreicherung)

What is Azure AI Search?



Data ingestion

- Azure Blob Storage containers
- Azure SQL Database
- Documents in Cosmos DB
- Azure Data Lake Storage Gen2
- Azure Table Storage

AI enrichment & indexing

- AI enables deeper understanding
  - Extract information & patterns
- Azure AI services
  - Vision, Natural Language Processing, etc.
- Indexing makes content searchable

Exploration

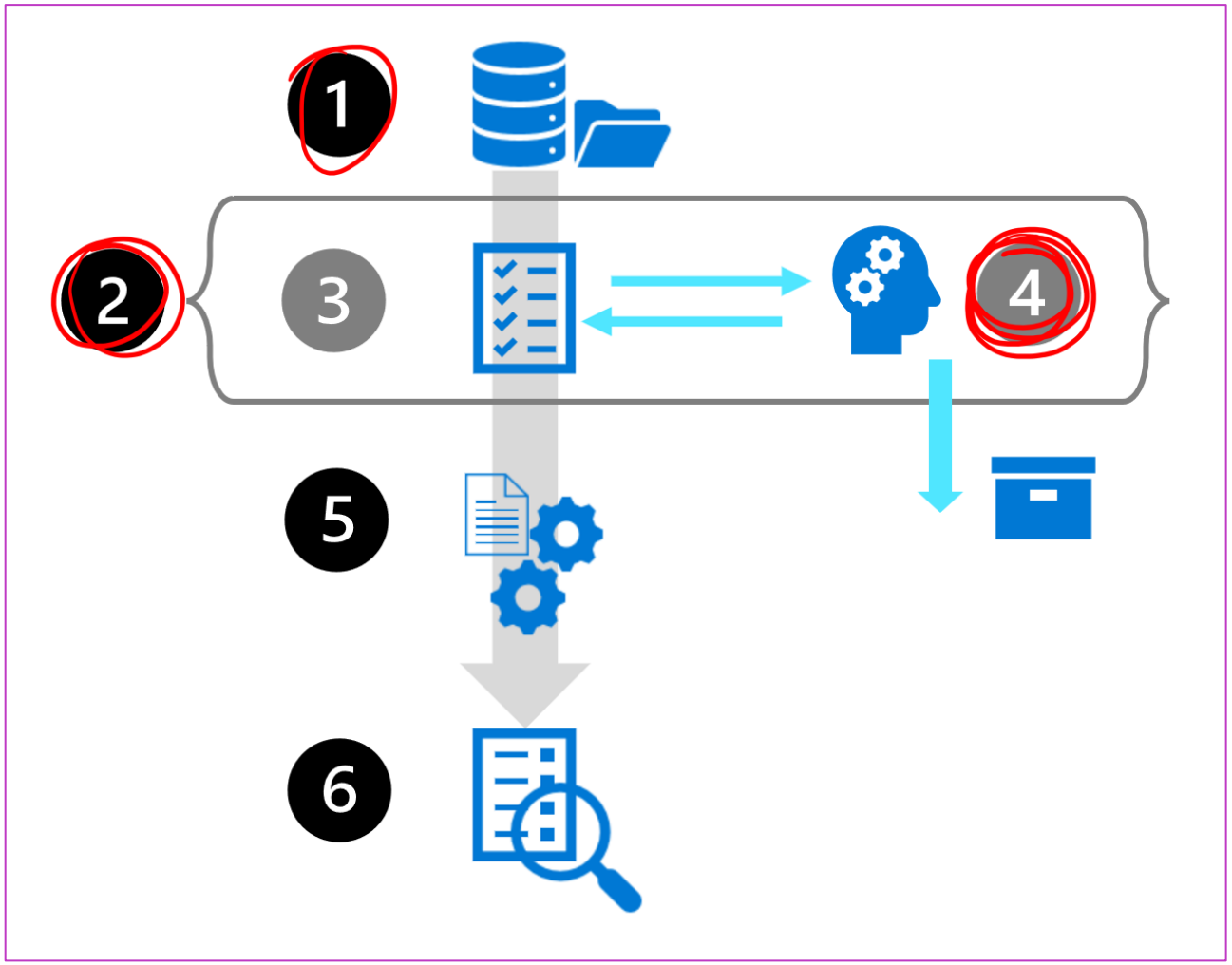
- Search performed on indexes
- Results used:
  - Within applications
  - Create data visualizations

© Copyright Microsoft Corporation. All rights reserved.

Elements of a search solution

Create a search index by moving data through the following indexing pipeline:

1. Start with a data source of original documents
2. Automate data movement with an **Indexer**.
3. Document cracking: the indexer opens files and extracts content.
4. Enrichment: the indexer moves data through AI enrichment by adding skills in a **skillset**. AI enriched content can be sent to a **knowledge store**
5. Push to index: the serialized JSON data populates the search index.
6. The result is a populated search index which can be explored through queries.



© Copyright Microsoft Corporation. All rights reserved.

Knowledge check

Which problem is Document Intelligence designed to solve?

- ☐ Reading text from large volumes of scanned documents
- ☐ Extracting key value pairs from forms, for input into a database
- ☐ Extracting positive and negative sentiment from scanned images

You want to extract information and perform semantic recognition on the extracted fields. What service would you use?

- ☐ AI Vision
- ☐ Optical Character Recognition
- ☐ AI Document Intelligence

Which two steps are part of building a AI Search solution?

- ☐ Creating databases of structured data, and indexing the data
- ☐ AI enrichment and indexing
- ☐ Identifying necessary skillsets to build indexes

© Copyright Microsoft Corporation. All rights reserved.

Knowledge check

Which problem is Document Intelligence designed to solve?

- ☐ Reading text from large volumes of scanned documents
- ☑ Extracting key value pairs from forms, for input into a database
- ☐ Extracting positive and negative sentiment from scanned images

You want to extract information and perform semantic recognition on the extracted fields. What service would you use?

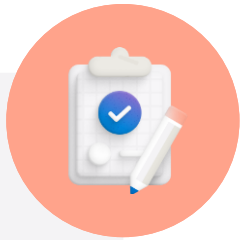
- ☐ AI Vision
- ☐ Optical Character Recognition
- ☑ AI Document Intelligence

Which two steps are part of building a AI Search solution?

- ☐ Creating databases of structured data, and indexing the data
- ☑ AI enrichment and indexing
- ☐ Identifying necessary skillsets to build indexes

© Copyright Microsoft Corporation. All rights reserved.

Summary



Azure AI Document Intelligence

- What is AI Document Intelligence?
- Three types of AI Document Intelligence services
- Analyzing specific forms with the AI Document Intelligence service

Azure AI Search

- What is knowledge mining?
- Three step process to creating a knowledge store with Azure AI Search

© Copyright Microsoft Corporation. All rights reserved.

© Copyright Microsoft Corporation. All rights reserved.