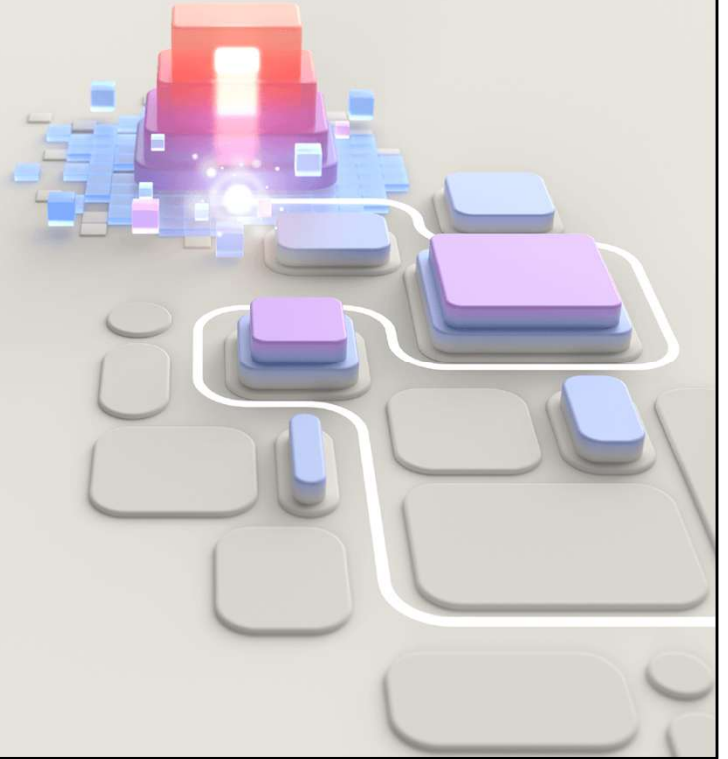


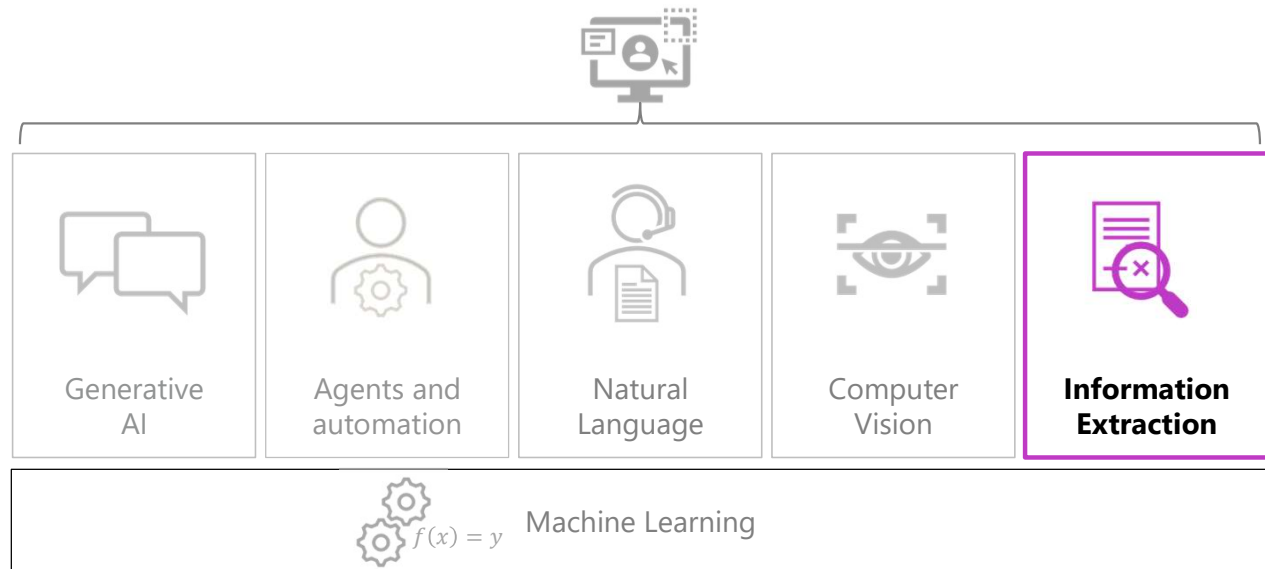


Introduction to AI in Azure: Information Extraction



© Copyright Microsoft Corporation. All rights reserved.


Our focus



© Copyright Microsoft Corporation. All rights reserved.

In this section we will focus on Information extraction.

Agenda



- Introduction to AI-powered information extraction concepts
- Get started with AI-powered information extraction in Microsoft Foundry

© Copyright Microsoft Corporation. All rights reserved.

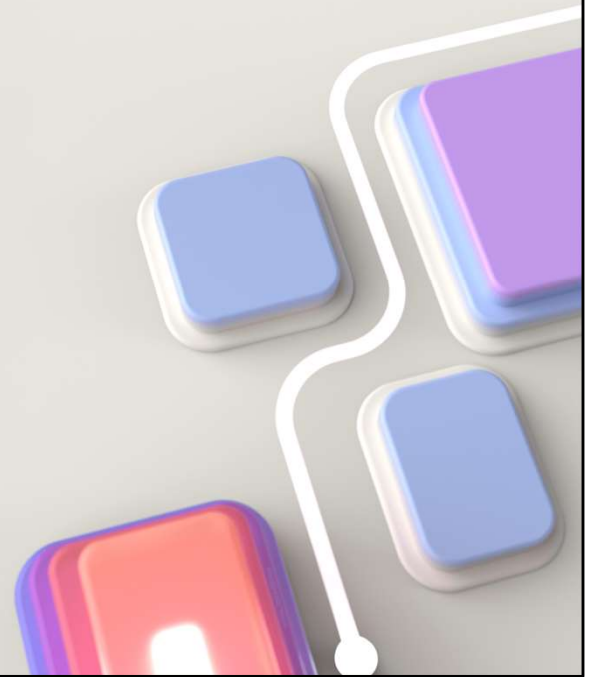
Time estimates:

- Introduction to AI-powered information extraction concepts: 25 minutes (including lab exercise)
- Get started with AI-powered information extraction in Microsoft Foundry: 30 minutes (including demo)

Introduction to AI-powered information extraction concepts

<https://aka.ms/mslearn-ai-info>

© Copyright Microsoft Corporation. All rights reserved.

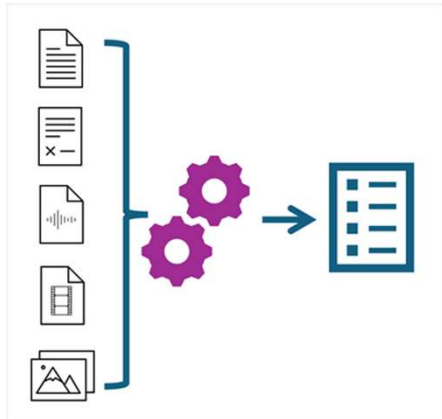


This content is meant to be taught in companion to:

- <https://learn.microsoft.com/training/modules/introduction-information-extraction/>
- <https://learn.microsoft.com/training/modules/ai-information-extraction/>

Use the link on the slide to see the Microsoft Learn learning module from which this section is derived.

Information extraction processes



Step	Description
Source Identification	Determine where the information resides and if it needs to be digitized.
Extraction	Leverages many techniques based on machine learning to understand and extract data from digitized content.
Transformation & Structuring	Extracted data is transformed into structured formats like JSON or tables.
Storage & Integration	The processed data is then stored in databases, data lakes, or analytics platforms for further use.

© Copyright Microsoft Corporation. All rights reserved.

Bilden visar de fyra huvudsakliga stegen i en **informationsutvinningsprocess**, alltså hur data samlas in, tolkas, struktureras och lagras så att den kan användas i digitala system. Den här processen är central inom områden som dokumentautomatisering, AI, affärsanalys och dataintegration.

1. Source Identification

Först identifieras var informationen finns – det kan vara dokument, bilder, formulär eller ljudfiler. Man avgör också om materialet redan är digitalt eller behöver skannas in. Exempel: ett papperskontrakt som först måste digitaliseras för att kunna behandlas.

2. Extraction

Här används AI och maskininlärning för att känna igen innehåll. Till exempel kan OCR (Optical Character Recognition) känna igen handskriven text, eller NLP (Natural Language Processing) identifiera nyckelord i ett avtal. AI kan även förstå layout och kontext, inte bara enskilda ord.

Exempel: Ett kvitto analyseras för att hitta belopp, datum och butiksnamn.

3. Transformation & Structuring

När data är extraherad omvandlas den till ett **strukturerat format**, som t.ex. en JSON-fil eller en tabell. Det gör informationen lättare att söka i, filtrera och analysera.

Exempel: Ett skannat formulär omvandlas till en databaspost med kolumner som namn, adress, och beställningsnummer.

4. Storage & Integration

Slutligen lagras datan i databaser, data lakes eller andra analysplattformar där den kan användas vidare – till exempel i BI-system, automatiska beslut eller rapportering.

Exempel: Insamlade kunddata från formulär skickas direkt till ett CRM-system.

Sammanfattningsvis visar bilden hur rå, ostrukturerad information förvandlas till användbar data – en nyckelprocess i dagens datadrivna verksamheter.

Examples of information extraction scenarios include:

A company needs to process employee expense claims, and has to extract expense descriptions and amounts from scanned receipts.

A customer service agency wants to analyze recorded support calls to identify common problems and resolutions.

A historical society needs to extract and store data from census records in scanned historical documents.

A tourist organization wants to analyze video footage and images taken at popular sites to help estimate visitor volumes and improve capacity planning for tours.

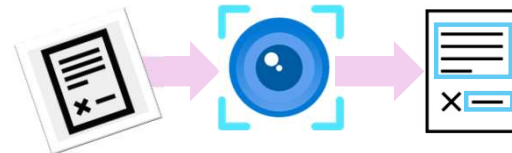
A finance department in a large corporation wants to automate accounts-payable processing by routing invoices received centrally to the appropriate departments for payment.

A marketing organization wants to analyze a large volume of digital images and documents, extracting and indexing the extracted data so it can be easily searched.

Understand the extraction of data from images

Use optical character recognition (OCR) to:

- Extract contact information from scanned business cards or conference badges
- Capture information from IDs to complete electronic forms
- Scan and store recipes or other text
- Photograph a sign or storefront so you can submit it to a translation service
- Digitize handwritten notes



© Copyright Microsoft Corporation. All rights reserved.

AI-powered information extraction replaces the need to manually inspect each piece of content for insights. Computer vision can extract insights from images to describe the people, places, things, and words they depict. Computer vision is made possible by machine learning models that are trained to recognize features based on large volumes of existing images. Machine learning models process images by transforming the images into numerical information. At its core, vision models perform calculations on the numerical information, which result in predictions of what's in the images.

Optical Character Recognition (OCR) helps computers recognize that an element in an image contains text. OCR is the foundation of processing text in images and uses machine learning models that are trained to recognize individual shapes as letters, numerals, punctuation, or other elements of text. Much of the early work on implementing this kind of capability was performed by postal services to support automatic sorting of mail based on postal codes. Since then, the state-of-the-art for reading text has moved on, and we have models that detect printed or handwritten text in an image and digitize it line-by-line and word-by-word.

Understand the extraction of data from forms

- The **field name** is the key or type of data entry.
- The **field description** is the definition of what the field name represents.
- The **value** corresponds with the field name and is the data specific to the content.

The data in forms is recognized with *bounding boxes*.

The image shows a sample invoice form. At the top left is a bicycle icon. The header 'Adventure Works Cycles' is underlined. To the right, 'Invoice No: 1234' and 'Date: 03/07/2028' are shown. The customer information section includes 'Customer Name: John Smith' and 'Address: 123 River Street, Marston, England, GL1 234'. Below this is a table with columns 'Item', 'Price', 'Quantity', and 'Item Total'. The table contains three rows: '38" Racing Bike (Red)' with price 1,299.00 and quantity 1; 'Cycling helmet (Black)' with price 25.99 and quantity 1; and 'Cycling shirt (L)' with price 42.50 and quantity 2. To the right of the table, a summary section shows 'Subtotal: 1,409.99', 'Tax: 140.99', 'Shipping: 35.00', and 'Total Due: 1,585.98'. Bounding boxes are drawn around various elements: the bicycle icon, the company name, the invoice number and date, the customer name and address, each item in the table, and the subtotal, tax, shipping, and total due.

Item	Price	Quantity	Item Total
38" Racing Bike (Red)	1,299.00	1	1,299.00
Cycling helmet (Black)	25.99	1	25.99
Cycling shirt (L)	42.50	2	85.00
Subtotal			1,409.99
Tax			140.99
Shipping			35.00
Total Due			1,585.98

© Copyright Microsoft Corporation. All rights reserved.

Bilden visar hur information extraheras automatiskt från formulär, som t.ex. fakturor.

Systemet identifierar **field names** (t.ex. *Customer Name*), som är nycklar eller etiketter.

Dessa beskrivs av **field descriptions**, som förklarar vad fältet betyder.

Det viktiga är **values** – de faktiska uppgifterna, t.ex. "John Smith" eller "1,299.00".

För att hitta och läsa data använder AI:n så kallade **bounding boxes**, som ritar rutor runt relevant text i bilden. Det gör att systemet kan matcha etiketter med tillhörande värden, även om layouten varierar.

Exempel:

Invoice No: → 1234

Item: → "38" Racing Bike (Red)"

Total Due: → 1,585.98

Detta används vid automatisering av bokföring, kunddataimport eller kvittoregistrering. Det sparar tid och minskar risken för fel jämfört med manuell inmatning.

Forms and other documents have text data with *semantic meaning*. Semantic meaning refers to the intended meaning or interpretation of words, phrases, or symbols in a given context. Semantic meaning goes beyond just the literal definition of a word (syntax) and focuses on what the word or sentence actually conveys.

Document intelligence describes AI capabilities that process text and attach semantic meaning to the extracted text. As an extension of optical character recognition (OCR), document intelligence automates the process of extracting and understanding information.

Consider an organization that needs to process large numbers of receipts for expenses claims, project costs, and other accounting purposes. Using document intelligence, the company can take a scanned image of a receipt,

digitize the text with OCR, and extract semantic meaning. The semantic meaning of data in forms can be described in field-value pairs.

The **field name** is the key or type of data entry.

The **field description** is the definition of what the field name represents.

The **value** corresponds with the field name and is the data specific to the content.

For example, in an invoice, the fields recognized may include:

Name, address, and telephone number of the merchant

Date and time of the purchase

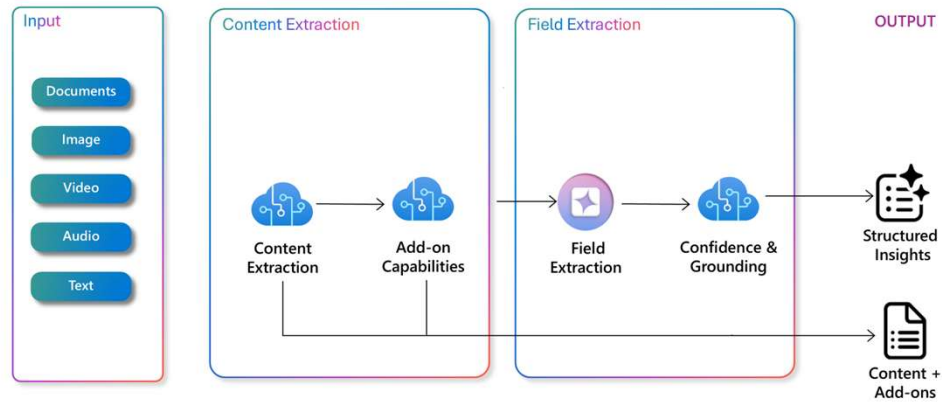
Name, quantity, and price of each item purchased

Total, subtotals, and tax values

The data in forms is recognized with *bounding boxes*.

Understand multimodal data extraction

AI-powered information extraction techniques can be combined to perform data extraction on multiple modalities of content, from documents to video and audio.



© Copyright Microsoft Corporation. All rights reserved.

Den här bilden visar hur AI kan extrahera information från olika typer av innehåll – som dokument, bilder, video, ljud och text. Först läser systemet innehållet (**content extraction**), sedan plockas viktiga fält ut (**field extraction**) och kontrolleras för noggrannhet (**confidence & grounding**).

Slutresultatet blir **strukturerad data**, redo för användning i t.ex. databaser eller analyser.

Exempel: En videoanalys kan hitta talade produktnamn, en bild kan avslöja ett kvittobelopp och ett dokument kan ge kundens adress – allt samlat till ett tydligt, användbart resultat.

AI-powered information extraction techniques can be combined to perform data extraction on multiple modalities of content, from documents to video and audio. Using multimodal data extraction can help with digital asset management, workflow automation, generating further insights, and more.

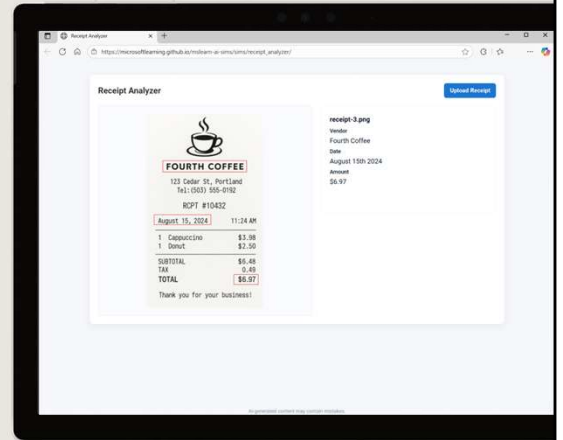
Exercise

Explore AI information extraction scenarios

In this exercise, you'll explore applications that extract information

Start the exercise at:

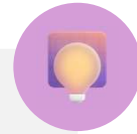
<https://go.microsoft.com/fwlink/?linkid=2334119>



© Copyright Microsoft Corporation. All rights reserved.

Note: The applications used in this exercise are *simulations* - there are no actual AI models or services behind them. However, it's based on real capabilities you can implement with [Azure AI Foundry](#); and in particular, the [Azure AI Document Intelligence](#) and [Azure AI Content Understanding](#) services.

Knowledge check



- 1** What is the primary role of machine learning in information extraction?
 - ☐ To store extracted data in a database
 - ☐ To convert structured data into unstructured formats
 - ☒ To predict fields and values
- 2** Which of the following best describes a "field" in the context of data extraction?
 - ☐ A visual marker used to highlight important text
 - ☒ A key that identifies the type of data being extracted
 - ☐ A storage location for raw content
- 3** How does generative AI enhance the data extraction process?
 - ☒ By allowing users to define custom fields and generate values from unstructured content
 - ☐ By converting JSON data into images
 - ☐ By generating new documents from scratch

© Copyright Microsoft Corporation. All rights reserved.

Här är korta svar med **max 10 ord per motivering**:

1. Maskininlärningens roll i information extraction

- ☒ **Fel:** Lagring är inte maskininlärningens uppgift.
- ☒ **Fel:** Målet är strukturering, inte ostrukturering.
- ☒ **Rätt:** Identifierar och förutsäger relevanta fält och värden.

2. Vad är ett "field" i dataextraktion?

- ☒ **Fel:** Ett fält är inte en visuell markering.
- ☒ **Rätt:** Beskriver vilken typ av information som extraheras.
- ☒ **Fel:** Ett fält är inte en lagringsplats.

3. Hur förbättrar generativ AI dataextraktion?

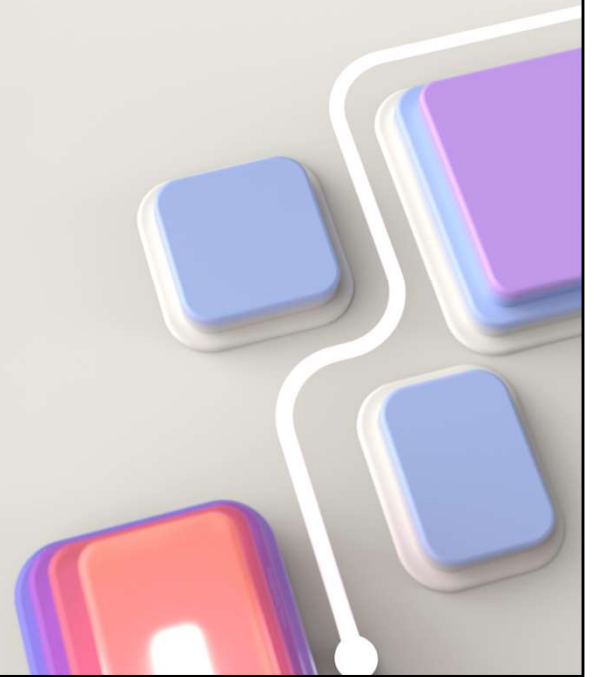
- ☒ **Rätt:** Möjliggör flexibla fält och värdegenerering från text.
- ☒ **Fel:** JSON till bild är irrelevant här.
- ☒ **Fel:** Dataextraktion skapar inte helt nya dokument.

Allow students a few minutes to think about the questions, then reveal the correct answers.

Get started with AI-powered information extraction in Microsoft Foundry

<https://aka.ms/mslearn-azure-info-extract>

© Copyright Microsoft Corporation. All rights reserved.



Use the link on the slide to see the Microsoft Learn learning module from which this section is derived.

Foundry options for reading text



Azure Vision Image Analysis

- Identify text and its location in scanned documents
- Find and read text in photographs
- Combine with other image analysis features to implement a digital asset management (DAM) solution



Azure Document Intelligence

- Design to support form processing by extracting fields and associated values from documents
- Use prebuilt models for common document types
- Create custom models for your specific requirements



Azure Content Understanding

- Use multimodal content extraction capabilities to analyze documents, forms audio, video, and images.
- Create custom analyzers to extract specific content or fields tailored to business needs

© Copyright Microsoft Corporation. All rights reserved.

Bilden visar en slide med titeln **"Foundry options for reading text"**. Den presenterar tre olika Azure-tjänster som kan användas för att läsa och analysera text i dokument och andra typer av material.

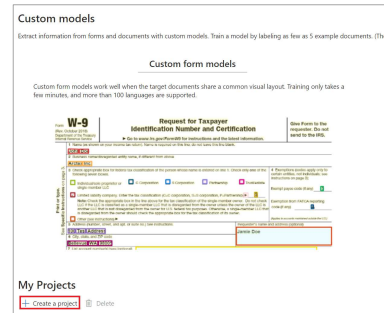
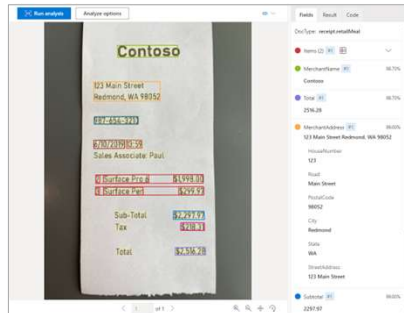
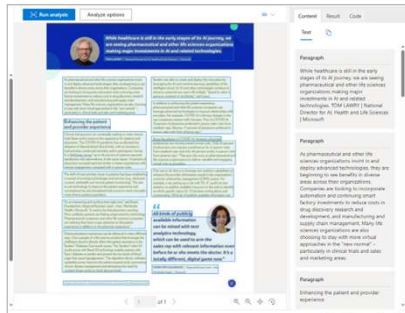
Den första tjänsten är **Azure Vision Image Analysis**, som handlar om att hitta och tolka text i bilder, till exempel i fotografier eller inskannade dokument. Den kan också användas tillsammans med andra bildanalysfunktioner i lösningar för digital hantering av bild- och dokumentresurser.

Den andra tjänsten är **Azure Document Intelligence**, som är inriktad på dokument och formulär. Den används för att extrahera information som fält och värden ur dokument, med stöd av färdiga modeller för vanliga dokumenttyper och möjlighet att skapa egna modeller vid behov.

Den tredje tjänsten är **Azure Content Understanding**, som är bredare och bygger på multimodal analys. Den kan hantera inte bara dokument och bilder utan även ljud och video, och gör det möjligt att skapa anpassade analysverktyg för att plocka ut specifikt innehåll utifrån verksamhetens krav.

Sliden sammanfattar alltså tre olika alternativ inom Azure Foundry för textläsning och innehållsanalys.

Azure Document Intelligence service



• 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.

Bilden visar vad **Azure AI Document Intelligence service** kan göra – en kraftfull tjänst för att läsa, tolka och strukturera information från dokument.

1. Document analysis

Tjänsten analyserar dokument och returnerar strukturerade data. Den hittar automatiskt viktiga områden, t.ex. rubriker, stycken eller tabeller, och visar relationer i innehållet. Det finns olika **analyze options** beroende på om du vill göra en gratis eller betald analys.

Exempel: En rapport i PDF kan analyseras så att varje avsnitt, citat eller faktaruta separeras och blir sökbar data.

2. Prebuilt models

Azure har färdiga modeller som direkt kan läsa vanliga dokumenttyper som **kvitton, ID-handlingar och fakturor**. Dessa modeller känner igen nyckel-värde-par, som t.ex. "Datum: 2025-08-01" eller "Total: 257.50".

Exempel: Ett inskannat kvitto analyseras automatiskt och ger dig belopp, plats, moms och datum som strukturerad data.

3. Custom models

Om du har egna formulär – t.ex. interna HR-blanketter – kan du träna en modell med så lite som fem exempel. Du markerar själv vilka fält som är viktiga.

Exempel: Ett företag skannar in sina anställningsavtal. En custom model kan tränas att extrahera t.ex. "Startdatum", "Lön" och "Roll" från varje dokument.

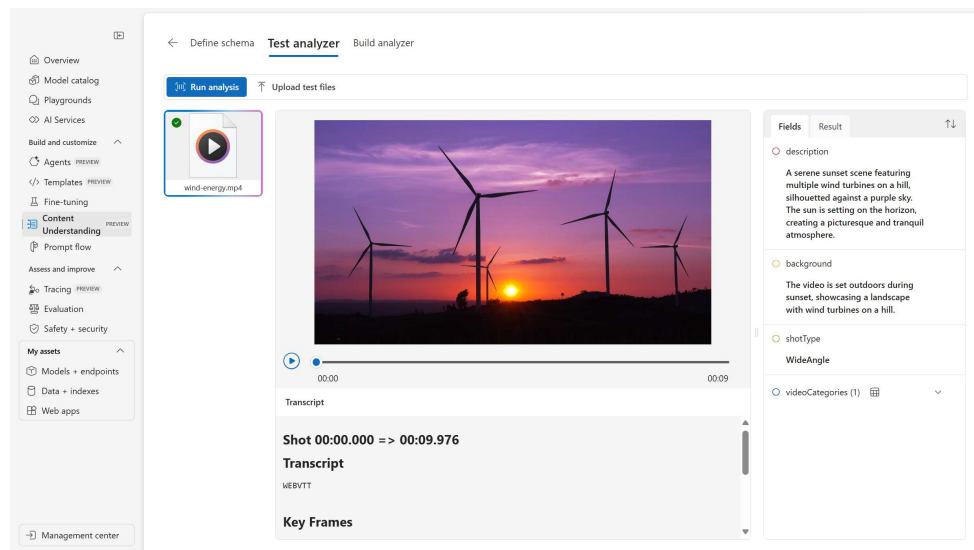
Tjänsten sparar tid, minskar fel och gör ostrukturerad dokumentinformation direkt användbar i databaser, appar och arbetsflöden.

Document Intelligence automatically reads, **understands**, and analyses business documents

Explain that Document Intelligence is made up of three related services:

- Document analysis
- Prebuilt models – machine learning models trained for specific types of forms, such as receipts (including hotel receipts, gas receipts, etc), invoices, ID.
 - Text and field names are identified
 - Available as JSON for input into a database
- Custom models
 - Train with as few as five sample forms
 - Requires an Azure storage account
 - Can customize the machine learning models for forms common in your organization

Azure Content Understanding service



© Copyright Microsoft Corporation. All rights reserved.

Azure Content Understanding uses state-of-the-art AI models to analyze content in multiple formats, including:

- Text-based forms and documents
- Audio
- Images
- Video

Azure Content Understanding's document analysis capabilities go beyond simple OCR-based text extraction to include schema-based extraction of fields and their values. In addition to text-based documents, Azure Content Understanding is capable of analyzing audio files to provide transcriptions, summaries, and other key insights. Azure Content Understanding supports analysis of images and video to extract information based on a custom schema. For example, you could analyze images of a video conference to extract details of attendance, location, and other information.

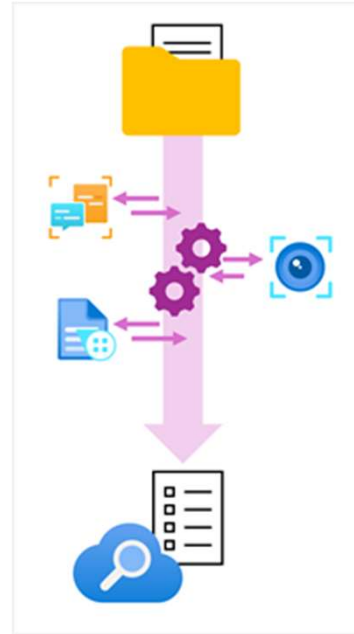
What is knowledge mining?

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.

- Azure AI Search results contain only your data, and can include new insights powered by AI
- Offered as Platform as a Service (PaaS) solution, meaning Microsoft manages the infrastructure and availability



© Copyright Microsoft Corporation. All rights reserved.

Knowledge mining innebär att använda AI för att hitta insikter i stora mängder ostrukturerad information, t.ex. dokument, PDF:er eller handskrivna anteckningar. I stället för att manuellt läsa igenom allt låter man AI identifiera relevant data snabbt och effektivt.

Azure AI Search är Microsofts plattform för detta. Den indexerar din data och skapar sökbara insikter, kombinerat med AI-tolkning. Resultatet är strukturerad och filtrerbar information som endast bygger på din egen data.

Tjänsten är en **PaaS-lösning**, vilket innebär att Microsoft sköter drift, säkerhet och tillgänglighet – du fokuserar på analysen.

👉 **Azure Storage-konto** används ofta för att lagra källdokument, t.ex. Word-filer, skannade PDF:er och bilder. Azure AI Search kan sedan kopplas till detta konto för att automatiskt hämta och analysera innehållet. Det ger en sömlös koppling mellan lagring och AI-driven informationsutvinning.

Fundamentally, Azure AI Search is a cloud service for indexing and searching data. However, its use of AI *skills* to extract insights from multiple formats of data and the ability to integrate it with other Foundry Tools, including Azure Vision and Azure Document Intelligence make it a powerful platform for building digital asset management and knowledge mining solutions.

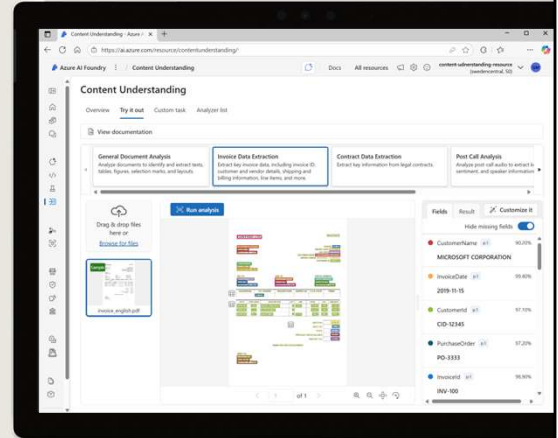
Exercise – If time permits

Extract data with Azure Content Understanding in Microsoft Foundry

In this exercise, you'll use Azure Content Understanding to extract information from an invoice.

Start the exercise at:

<https://go.microsoft.com/fwlink/?linkid=2320420>



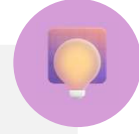
© Copyright Microsoft Corporation. All rights reserved.

You require an Azure subscription to perform this exercise – which may be provided by an Authorized Lab Host.

The exercise can take a substantial amount of time. Our recommendation in one-day deliveries is for the instructor to demonstrate the core tasks in this exercise; completing the exercise ahead of time

(<https://go.microsoft.com/fwlink/?linkid=2320420>) so you have a project to demonstrate – be sure to select a region in which Content Understanding is supported.

Knowledge check



- 1** Which service should you use to build a mobile app with which a user can take a photograph of a street sign and extract text to send to a translation service?
 - ☒ Azure Vision Image Analysis.
 - ☐ Azure Content Understanding
 - ☐ Azure Document Intelligence.
- 2** Which service should you use to analyze a recording of a customer service call to extract key points that were discussed?
 - ☐ Azure Vision Image Analysis
 - ☒ Azure Content Understanding
 - ☐ Azure Document Intelligence
- 3** Which service should you use to extract fields from standard US tax forms for automated processing?
 - ☐ Azure Vision Image Analysis
 - ☐ Azure Content Understanding
 - ☒ Azure Document Intelligence

© Copyright Microsoft Corporation. All rights reserved.

Här är korta svar med **max 10 ord per motivering**:

1. Fotografera gatuskylt och extrahera text

- ☒ **Rätt:** OCR från bilder stöds av Azure Vision Image Analysis.
- ☒ **Fel:** Avsedd för multimodalt innehåll, inte enkel bild-OCR.
- ☒ **Fel:** Fokuserar på dokument, inte spontana mobilbilder.

2. Analysera inspelat kundtjänstamtal

- ☒ **Fel:** Bildanalys hanterar inte ljudinspelningar.
- ☒ **Rätt:** Analyserar tal, sammanfattar och extraherar nyckelpunkter.
- ☒ **Fel:** Dokumentbaserad, inte avsedd för ljud.

3. Extrahera fält från amerikanska skatteblanketter

- ☒ **Fel:** Hanterar bilder, inte strukturerade formulärfält.
- ☒ **Fel:** För innehållsanalys, inte formulärextraktion.
- ☒ **Rätt:** Förtränade modeller för standardiserade formulär.

Allow students a few minutes to think about the questions, then reveal the correct answers.

Summary



Introduction to AI-powered information extraction concepts

- Information extraction uses a combination of OCR, machine learning, and generative AI to find fields and extract values

Get started with AI-powered information extraction in Microsoft Foundry

- Detect text and its location in images with Azure Vision
- Extract fields from complex forms using Azure Document Intelligence
- Extract information from multi-modal content using Azure Content Understanding
- Build knowledge mining solutions with Azure AI Search

© Copyright Microsoft Corporation. All rights reserved.