Use prebuilt Document intelligence models

Learn what data you can analyze by choosing **prebuilt Forms Analyzer models** and how to deploy these models in a **Document intelligence solution**.

Learning objectives

In this module, you'll learn to:

- Identify business problems that you can solve by using prebuilt models in **Forms Analyzer**.
- Analyze forms by using the **General Document, Read, and Layout models**.
- Analyze forms by using **financial, ID, and tax prebuilt models**.

Introduction

Many forms and documents that businesses handle are common across disparate companies in different sectors. For example, most companies use invoices and receipts. Microsoft **Azure AI Document Intelligence** includes **prebuilt models** so you can handle common document types easily.

Imagine you conduct polls for private companies and political parties. Participants submit their responses as paper forms or as online PDFs. You've decided to deploy **Azure AI Document Intelligence** to streamline data entry. You want to know if you can use the prebuilt models to generate meaningful data from your forms.

In this module, you'll learn about the capabilities of the **prebuilt models in Azure AI Document Intelligence** and how to use them.

Understand prebuilt models

Prebuilt models in Azure AI Document Intelligence enable you to **extract data from common forms and documents without training your own models**.

In your polling company, polling forms are unique to each survey project, but you also use invoices and receipts to record financial transactions and you have many unstructured documents. You want to know how much work is required to extract names, addresses, amounts, and other information from these documents.

Here, you learn how prebuilt models can help you analyze common document types.

What are prebuilt models?

The general approach used in AI solutions is to provide a large quantity of sample data and then train an optimized model by trying different data features, parameters, and statistical treatments. The combination that best predicts the values that interest you constitute the trained model, and you can use this model to predict values from new data.

Many of the forms that businesses use from day to day are of a few common types. For example, most businesses issue or receive invoices and receipts. Any business that has employees in the United States must use the W-2 tax declaration form. Also you often have more general documents that you might want to extract data from. For these cases, Microsoft helps you by providing prebuilt models. **Prebuilt models are already trained on large numbers of their target form type**.

If you want to use Document Intelligence to extract data from one of these common forms or documents, you can choose to use a prebuilt model and you don't have to train your own. Because Microsoft trains these models on a large corpus of examples, you can expect them to provide accurate and reliable results when dealing with their intended forms.

Several of the prebuilt models are trained on specific form types:

- **Invoice** model. Extracts common fields and their values from invoices.
- **Receipt** model. Extracts common fields and their values from receipts.
- **US Tax** model. Unified US tax model that can extract from forms such as W-2, 1098, 1099, and 1040.
- **ID document** model. Extracts common fields and their values from US drivers' licenses, European Union IDs and drivers license, and international passports.
- **Business card** model. Extracts common fields and their values from business cards.
- **Health insurance card** model. Extracts common fields and their values from health insurance cards.
- **Marriage certificate**. Extracts information from marriage certificates.
- **Credit/Debit card** model. Extracts common information from bank cards.
- **Mortgage documents**. Extracts information from mortgage closing disclosure, Uniform Residential Loan Application (Form 1003), Appraisal (Form 1004), Validation of Employment (Form 1005), and Uniform Underwriting and Transmittal Summary (Form 1008).
- **Bank statement** model. Extracts account information including beginning and ending balances, transaction details from bank statements.
- **Pay Stub** model. Extracts wages, hours, deductions, net pay, and other common pay stub fields.
- **Check** model. Extracts payee, amount, date, and other relevant information from checks.

The other models are designed to extract values from documents with less specific structures:

- **Read model**. Extracts text and languages from documents.
- **General document model**. Extract text, keys, values, entities, and selection marks from documents.
- **Layout model**. Extracts text and structure information from documents.

Features of prebuilt models

The prebuilt models are designed to extract different types of data from the documents and forms users submit. To select the right model for your requirements, you must understand these features:

- **Text extraction**. All the prebuilt models extract lines of text and words from hand-written and printed text.
- **Key-value pairs**. Many models extract spans of text within a document that identify a label or key and its response or value as key-values pairs. For example, *a typical key might be Weight and its value might be 31 kg.*
- **Entities**. Text that includes common, more complex data structures can be extracted as entities. **Entity types include people, locations, and dates**.
- **Selection marks**. Some models extract spans of text that indicate a choice as selection marks. These **marks include radio buttons and check boxes**.
- **Tables**. Many models can extract tables in scanned forms included the data contained in cells, the numbers of columns and rows, and column and row headings. Tables with merged cells are supported.
- **Fields**. Models trained for a specific form type identify the values of a fixed set of fields. For example, the Invoice model includes **CustomerName** and **InvoiceTotal** fields.

Also consider that prebuilt models are designed for and trained on generic document and form types. If you have an industry-specific or unique form type that you use often, you might be able to obtain more reliable and predictable results by using a custom model. However, **custom models take time to develop because you must invest the time and resources to train them on example forms before you can use it**. *The larger the number of example forms you provide for training, the better the model is at predicting form content accurately*.

Input requirements

The prebuilt models are flexible but you can help them to return accurate and helpful results by **submitting one clear photo or high-quality scan for each document**.

You must also comply with these requirements when you submit a form for analysis:

- The file must be in **JPEG, PNG, BMP, TIFF, or PDF** format. Additionally, the Read model can accept Microsoft Office files.
- The file must be smaller than **500 MB for the standard tier**, and **4 MB for the free tier**.
- Images must have dimensions **between 50 x 50 pixels and 10,000 x 10,000 pixels**.
- PDF documents must have dimensions **less than 17 x 17 inches or A3 paper size**.
- PDF documents must **not be protected with a password**.

> Note: If you can, submit *text-embedded PDF files* because they eliminate errors in character recognition.

PDF and TIFF files can have any number of pages but, in the **standard tier**, only the **first 2,000 pages** are analyzed. In the **free tier**, only the **first two pages** are analyzed.

Try out prebuilt models with Azure AI Document Intelligence Studio

Azure AI Document Intelligence is designed as a web service you can call using code in your custom applications. However, it's often helpful to explore the models and how they behave with your forms visually. You can perform such experiments by using Azure AI Document Intelligence Studio and use the experience to help design and write your code.

You can choose any of the prebuilt models in Azure AI Document Intelligence Studio. Microsoft provides some sample documents for use with each model or you can add your own documents and analyze them.
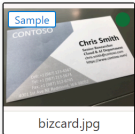

Calling prebuilt models by using APIs

Because **Azure AI Document Intelligence implements RESTful web services**, you can use web service calls from any language that supports them. However, when you use Microsoft's Azure AI Document Intelligence APIs, security and session management is simplified and you have to write less code.

Azure AI Document Intelligence is available for:

- C# and other .NET languages.

- Java.
- Python.
- JavaScript.

Whenever you want to call Azure AI Document Intelligence, you must start by **connecting and authenticating with the service in your Azure subscription**. To make that connection, you need:

- The service **endpoint**. This value is the URL where the service is published.
- The **API key**. This value is a unique key that grants access.

You obtain both of these values from the Azure portal.

Because the service can take a few seconds to respond, it's best to use **asynchronous** calls to submit a form and then obtain results from the analysis:

```
poller = document_analysis_client.begin_analyze_document(          "prebuilt-layout", AnalyzeDocumentRequest(url_source=docUrl      ))
result: AnalyzeResult = poller.result()
```

The details you can extract from these results depend on the model you used.

Learn more

- [What is Azure AI Document Intelligence?](#)
- [Azure AI Document Intelligence models](#)

Use the General Document, Read, and Layout models

If you want to extract text, languages, and other information from documents with unpredictable structures, you can use the **read, general document, or layout models**.

In your polling company, customers and partners often send specifications, tenders, statements of work, and other documents with unpredictable structures. You want to know if Azure AI Document Intelligence can analyze and extract values from these documents.

Here, you'll learn about the **prebuilt models** that Microsoft provides **for general documents**.

Using the read model

The **Azure AI Document Intelligence read model** extracts printed and handwritten text from documents and images. It's used to provide *text extraction* in all the other prebuilt models.

The **read model** can also **detect the language** that a line of text is written in and **classify whether it's handwritten or printed text**.

> Note: The read model supports more languages for printed text than handwritten text. Check the [documentation](#) to see the current list of supported languages.

For multi-page PDF or TIFF files, you can use the `pages` parameter in your request to fix a page range for the analysis.

The read model is ideal if you want to extract words and lines from documents with no fixed or predictable structure.

Using the general document model

The general document model extends the functionality of the read model by adding the **detection of key-value pairs, entities, selection marks, and tables**. The model can extract these values from **structured, semi-structured, and unstructured documents**.

The general document model is the only prebuilt model to support entity extraction. It can recognize entities such as **people, organizations, and dates** and it runs against the whole document, not just key-value pairs. This approach ensures that, **when structural complexity has prevented the model extracting a key-value pair, an entity can be extracted instead**. Remember, however, that sometimes a single piece of text *might return both a key-value pair and an entity*.

The types of entities you can detect include:

- `Person` . The name of a person.
- `PersonType` . A job title or role.
- `Location` . Buildings, geographical features, geopolitical entities.
- `Organization` . Companies, government bodies, sports clubs, musical bands, and other groups.
- `Event` . Social gatherings, historical events, anniversaries.
- `Product` . Objects bought and sold.
- `Skill` . A capability belonging to a person.
- `Address` . Mailing address for a physical location.
- `Phone number` . Dialing codes and numbers for mobile phones and landlines.
- `Email` . Email addresses.
- `URL` . Webpage addresses.
- `IP Address` . Network addresses for computer hardware.
- `DateTime` . Calendar dates and times of day.
- `Quantity` . Numerical measurements with their units.

Using the layout model

As well as extracting text, the layout model returns **selection marks** and **tables** from the input image or PDF file. It's a good model to use when you need **rich information about the structure of a document**.

When you digitize a document, it can be at an odd angle. Tables can have complicated structures with or without headers, cells that span columns or rows, and incomplete columns or rows. The layout model can handle all of these difficulties to extract the complete document structure.

For example, each **table cell** is extracted with:

- Its content text.
- The size and position of its bounding box.
- If it's part of a header column.
- Indexes to indicate its row and column position in the table.

Selection marks are extracted with their **bounding box, a confidence indicator, and whether they're selected or not**.

Learn more

- Language support for Azure AI Document Intelligence
- Azure AI Document Intelligence read model
- Azure AI Document Intelligence general document model
- Azure AI Document Intelligence layout model

Use financial, ID, and tax models

Azure AI Document Intelligence includes some prebuilt models that are trained on common form types. You can use these models to obtain the values of common fields from **invoices, receipts, business cards**, and more.

In your polling company, invoices and receipts are often submitted as photos or scans of the paper documents. Sometimes the scan is poor and the paper is creased or damaged. You want to know if Azure AI Document Intelligence can get this information into your databases more efficiently than manual data entry.

Here, you'll learn about the prebuilt models that handle financial, identity, and tax documents.

Using the invoice model

Your business both issues invoices and receives them from partner organization. There might be many different formats on paper or in digitized forms and some will have been scanned poorly at odd angles or from creased paper.

The invoice model in Azure AI Document Intelligence can handle these challenges and uses the features of the read model to extract text from invoice scans. In addition, it extracts specific fields that are commonly used on invoices including:

- Customer name and reference ID
- Purchase order number
- Invoice and due dates
- Details about the vendor, such as name, tax ID, physical address.
- Similar details about the customer.
- Billing and shipping addresses.
- Amounts such as total tax, invoice total, and amount due.

Invoices also feature lines, usually in a table, each of which is one purchased item. **For each line**, the invoice model identifies details including:

- The description and product code of the product or service invoiced.
- Amounts such as the unit price, the quantity of items, the tax incurred, and the line total.

Using the receipt model

Receipts have similar fields and structures to invoices, but they record amounts paid instead of amounts charged. Azure AI Document Intelligence faces the same challenges of poor scanning or digitization but can reliably identify fields including:

- **Merchant details** such a name, phone number, and address.
- **Amounts** such as receipt total, tax, and tip.
- The **date and time of the transaction**.

As for invoices, receipts often include a table of items, each of which is a product or service purchased. **For each of these lines**, the model recognizes:

- The name of the item.
- The quantity of the item purchased.
- The unit price of the item.
- The total price for that quantity.

> Note: In Azure AI Document Intelligence v3.0 and later, the receipt model supports single-page hotel receipt processing. If a receipt is classified as a hotel receipt, the model extracts extra relevant fields such as **arrival and departure dates**.

Using the ID document model

The ID document model is trained to analyze two types of identity document:

- **United States drivers licenses**.
- **International passports**.

> Note: Only the biographical pages of passports can be analyzed. Visas and other travel documents are not supported.

The ID document model can extract fields including:

- First and last names.
- Personal information such as sex, date of birth, and nationality.
- The country and region where the document was issued.
- Unique numbers such as the document number and machine readable zone.
- Endorsements, restrictions, and vehicle classifications.

> **Important**: Since much of the data extracted by the ID document model is personal, it is of a sensitive nature and covered by data protection laws in most jurisdictions. Be sure that you have the permission of the individual to store their data and comply with all legal requirements in the way you handle this information.

Using the business card model

Business cards are a popular way to exchange contact information quickly and often include branding, unusual fonts, and graphic design elements. Fields that the business card model can extract include:

- First and last names.
- Postal addresses.
- Email and website addresses.
- Various telephone numbers.

Using other prebuilt models

Azure AI Document Intelligence offers several prebuilt models, with new models being released regularly. Before training a custom model, it's worth verifying if your use case can be analyzed accurately with one of these prebuilt models. Using a prebuilt model will benefit from rigorous testing, updated model versions, and reduced cost compared to a custom model.

Learn more

- Azure AI Document Intelligence model overview

Exercise - Analyze a document using Azure AI Document Intelligence

Use prebuilt Document Intelligence models

In this exercise, you'll set up an Azure AI Document Intelligence resource in your Azure subscription. You'll use both the Azure AI Document Intelligence Studio and C# or Python to submit forms to that resource for analysis.

Module assessment

1. You have a large set of documents with varying structures that contain customer name and address information. You want to extract entities for each customer. Which prebuilt model should you use? **General document model**.
2. You are using the prebuilt layout model to analyze a document with many checkboxes. You want to find out whether each box is checked or empty. What object should you use in the returned JSON code? **Selection marks**.
3. You submit a Word document to the Azure AI Document Intelligence general document model for analysis but you receive an error. The file is A4 size, contains 1 MB of data, and is not password-protected. How should you resolve the error? **Convert the document to PDF format**.

Summary

There are many document types that are common to most business and Azure AI Document Intelligence includes prebuilt models to handle them. If you have a collection of such forms that you want to analyze, you can extract data by using these prebuilt models and you don't have to train your own models. You can get up and running very quickly by submitting photos and scans to the most appropriate prebuilt model.

Now that you've completed this module, you can:

- Identify business problems that you can solve by using prebuilt models in Azure AI Document Intelligence.
- Analyze forms by using the General Document, Read, and Layout models.
- Analyze forms by using financial, ID, and tax prebuilt models.

Learn more

- What is Azure AI Document Intelligence?
- Azure AI Document Intelligence models
- Language support for Azure AI Document Intelligence
- Azure AI Document Intelligence read model
- Azure AI Document Intelligence general document model
- Azure AI Document Intelligence layout model
- Azure AI Document Intelligence invoice model
- Azure AI Document Intelligence receipt model
- Azure AI Document Intelligence ID document model
- Azure AI Document Intelligence business card model
- Azure AI Document Intelligence W-2 model