Azure AI Document Intelligence uses Azure AI Services to analyze the content of scanned forms and convert them into data. It can recognize text values in both common forms and forms that are unique to your business.

**Azure AI Document Intelligence**

* **Purpose**:
  + Automates data extraction from completed forms, reducing the need for manual data entry.
  + Improves accuracy and reduces costs associated with manual data entry.

**Planning Azure AI Document Intelligence Resources**

**1. Choosing Your Development Language**

Select the language that best fits your existing systems and team expertise. Supported languages include:

* C#/.NET
* Java
* Python
* JavaScript

**2. Creating Azure AI Document Intelligence Resources**

To set up Azure AI Document Intelligence, follow these steps:

1. **Sign in to Azure Portal**:
   * Navigate to [Azure Portal](https://portal.azure.com/).
2. **Create a New Resource**:
   * Select "Create a resource" in the left-hand menu.
   * In the "Search services and marketplace" box, type "Document Intelligence" and press Enter.
3. **Configure the Resource**:
   * On the Document Intelligence page, select "Create".
   * **Project Details**:
     + Select your Subscription.
     + Choose an existing Resource group or create a new one.
   * **Instance Details**:
     + Select a Region near your users.
     + Enter a unique Name for the resource.
     + Choose a Pricing tier (Free (F0) or Standard (S0)).
   * Select "Review + create".
4. **Deploy the Resource**:
   * If validation tests pass, select "Create". Azure will deploy the new Azure AI Document Intelligence resource.

**3. Obtaining Connection Details**

After creating the resource, you need to obtain the endpoint and access key to connect your application:

1. **Navigate to the Resource**:
   * In the Azure portal, go to the Azure AI Document Intelligence resource you created.
2. **Access Keys and Endpoint**:
   * Under "Resource Management", select "Keys and Endpoint".
   * Copy either KEY 1 or KEY 2 and the Endpoint URL. These will be used to authenticate and connect your application to Azure AI Document Intelligence.

**Models in Azure AI Document Intelligence**

* **Prebuilt Models**: For common document types, enabling quick setup.
  + **General Models**: Read, General document, Layout.
  + **Specific Models**: Invoice, Receipt, W-2 US tax declaration, ID Document, Business card, Health insurance card.
* **Custom Models**: For unique or specific forms, trained using example documents.
* **Composed Models**: Combine multiple custom models to handle different document types with a single service.

**Integration with Azure AI Vision**

* **Azure AI Vision OCR**: Extracts simple words and text from images.
* **Azure AI Document Intelligence**: Provides contextual analysis, identifying key/value pairs, tables, and fields for comprehensive document understanding.

**Tools for Azure AI Document Intelligence**

* **Azure AI Document Intelligence Studio**: Visual tool for exploring features without coding.
* **APIs**: Available in multiple programming languages (C#/.NET, Java, Python, JavaScript) for integration into applications.
* **RESTful Web Service**: For use with other programming languages.

**Choosing a Model Type for Azure AI Document Intelligence**

**1. Prebuilt Models**

These are ready-to-use models provided by Azure, designed to handle common document types. They require no training and are quick to implement.

**General Document Analysis Models:**

* **Read:** Extracts words and lines from printed and handwritten documents, and detects the language used.
* **General Document:** Extracts key-value pairs and tables.
* **Layout:** Extracts text, tables, and structural information from forms, including checkboxes and radio buttons.

**Specific Document Type Models:**

* **Invoice:** Extracts key information from sales invoices (available in English and Spanish).
* **Receipt:** Extracts data from printed and handwritten receipts.
* **W-2:** Extracts data from the U.S. government's W-2 tax forms.
* **ID Document:** Extracts data from U.S. driver's licenses and international passports.
* **Business Card:** Extracts names and contact details from business cards.

**2. Custom Models**

Use custom models if your forms are unique and not covered by prebuilt models. Custom models are trained using examples of your specific forms.

**Training Custom Models:**

* Supply at least five examples of the completed form for training.
* More examples lead to greater confidence levels.
* Include a range of possible inputs in your training set (e.g., both handwritten and printed text).

**Types of Custom Models:**

* **Custom Template Models:** Suitable for forms with consistent visual templates. Train a model for each form variation if there are different templates.
* **Custom Neural Models:** Suitable for a range of document structures, from structured to unstructured. Best for English and other Latin-based languages.

**3. Composed Models**

A composed model combines multiple custom models. It's useful when you have various forms or don't know the document type in advance.

**Using Composed Models:**

* Automatically classifies and uses the most appropriate custom model for each document.
* Up to 100 custom models can be combined in a single composed model with the Standard pricing tier (up to 5 with the Free tier).
* The docType property in the results indicates which custom model was used for analysis.

**Summary**

* **Prebuilt Models:** Quick implementation for common document types. No training required.
* **Custom Models:** Tailored for unique documents. Requires training with specific examples.
* **Composed Models:** Combines multiple custom models for handling various document types seamlessly.

Many forms and documents that your business handles 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.

**Key Points about Prebuilt Models in Azure AI Document Intelligence**

**1. What are Prebuilt Models?**

* Prebuilt models are AI models trained on large datasets to handle common document types like invoices, receipts, and W-2 tax forms.
* They provide accurate and reliable results without requiring custom training.

**2. Types of Prebuilt Models:**

* **Invoice Model:** Extracts fields like CustomerName and InvoiceTotal from invoices.
* **Receipt Model:** Extracts common fields from receipts.
* **W2 Model:** Extracts data from U.S. W-2 tax forms.
* **ID Document Model:** Extracts fields from U.S. driver's licenses and international passports.
* **Business Card Model:** Extracts names and contact details from business cards.
* **Health Insurance Card Model:** Extracts data from health insurance cards.
* **Read Model:** Extracts text and language from documents.
* **General Document Model:** Extracts text, key-value pairs, entities, and selection marks from documents.
* **Layout Model:** Extracts text, tables, and structure from documents.

**3. Features of Prebuilt Models:**

* **Text Extraction:** Extracts printed and handwritten text lines and words.
* **Key-Value Pairs:** Extracts labels and their responses, like "Weight: 31 kg".
* **Entities:** Extracts complex data structures such as people, locations, and dates.
* **Selection Marks:** Extracts radio buttons and checkboxes.
* **Tables:** Extracts data from tables, including cell data, column/row counts, and headers.
* **Fields:** Specific fields identified for certain form types (e.g., InvoiceTotal on invoices).

**4. Input Requirements:**

* Supported formats: JPEG, PNG, BMP, TIFF, PDF.
* File size: <500 MB for standard tier, <4 MB for free tier.
* Image dimensions: 50x50 to 10,000x10,000 pixels.
* PDF dimensions: Less than 17x17 inches or A3 size.
* PDF files: Should not be password protected; text-embedded PDFs are preferred for accuracy.
* Page limits: First 2000 pages analyzed in standard tier, first two pages in free tier.

**5. Comparison of Prebuilt Models:**

| **Model** | **Text Extraction** | **Key-Value Pairs** | **Entities** | **Selection Marks** | **Tables** | **Fields** |
| --- | --- | --- | --- | --- | --- | --- |
| Read | X |  |  |  |  |  |
| General Document | X | X | X | X | X |  |
| Layout | X |  |  |  | X | X |
| Invoice | X | X |  |  | X | X |
| Receipt | X | X |  |  |  | X |
| W2 | X | X |  |  | X | X |
| ID Document | X | X |  |  |  | X |
| Business Card | X | X |  |  |  | X |

**6. Trying Prebuilt Models with Azure AI Document Intelligence Studio:**

* Use the Azure AI Document Intelligence Studio to visually explore how prebuilt models work with your documents.
* Microsoft provides sample documents, or you can upload your own to analyze.

**7. Using APIs to Call Prebuilt Models:**

* Azure AI Document Intelligence offers RESTful web services accessible from any language supporting web service calls.
* Simplified with APIs available for C#, Java, Python, and JavaScript.
* Requires connection and authentication using the service endpoint and API key obtained from the Azure portal.

**Using Azure AI Document Intelligence Models: Read, General Document, and Layout**

When dealing with documents that have unpredictable structures, the Read, General Document, and Layout models in Azure AI Document Intelligence can be very useful. Here’s how each model can help:

**1. Read Model**

**Purpose:**

* Extracts printed and handwritten text from documents and images.
* Detects the language of the text.
* Classifies text as handwritten or printed.

**Use Cases:**

* Ideal for documents with no fixed or predictable structure.
* Extracts words and lines of text from documents like specifications, tenders, and statements of work.

**Features:**

* Supports multiple languages for printed text.
* Handwriting detection is limited to Latin languages.
* Can specify a range of pages for multi-page PDF or TIFF files using the pages parameter.

**2. General Document Model**

**Purpose:**

* Extends the Read model by adding the detection of key-value pairs, entities, selection marks, and tables.
* Suitable for structured, semi-structured, and unstructured documents.

**Use Cases:**

* Analyzing documents that include a mix of structured information (like tables) and unstructured text.
* Extracting rich data including key-value pairs and entities from documents such as contracts or reports.

**Features:**

* **Key-Value Pairs:** Identifies pairs of related text, such as labels and values.
* **Entities:** Recognizes various types of entities such as people, organizations, dates, etc.
* **Selection Marks:** Identifies checkboxes and radio buttons.
* **Tables:** Extracts tables and their structure.

**Entity Types:**

* Person (names)
* PersonType (job titles or roles)
* Location (geographical entities)
* Organization (companies, groups)
* Event (gatherings, anniversaries)
* Product (goods or services)
* Skill (capabilities)
* Address (physical mailing addresses)
* Phone Number (contact numbers)
* Email (email addresses)
* URL (web addresses)
* IP Address (network addresses)
* DateTime (dates and times)
* Quantity (measurements and units)

**3. Layout Model**

**Purpose:**

* Extracts text and detailed structure information from documents, including selection marks and tables.
* Handles digitized documents with complex layouts.

**Use Cases:**

* Suitable for documents with rich structures, like forms and detailed reports.
* Useful for digitized documents that may have been scanned at odd angles or have complex table structures.

**Features:**

* **Text Extraction:** Extracts all text within the document.
* **Selection Marks:** Identifies and extracts checkboxes and radio buttons, including their selection status.
* **Tables:** Extracts comprehensive table information, including:
  + Content text of each cell.
  + Size and position of each cell’s bounding box.
  + Header information.
  + Row and column indices.

**Financial, ID, and Tax Models in Azure AI Document Intelligence**

For a polling company managing diverse documents like invoices, receipts, and identity documents, Azure AI Document Intelligence provides specialized models to streamline data extraction and reduce manual data entry. Below are the key models tailored for handling financial, identity, and tax documents:

**Invoice Model**

**Purpose:**

* Extracts detailed information from invoices, handling various formats and poor-quality scans.

**Extractable Fields:**

* **Customer Information:** Name, reference ID.
* **Order Details:** Purchase order number, invoice and due dates.
* **Vendor Information:** Name, tax ID, physical address.
* **Customer Information:** Similar details as the vendor.
* **Addresses:** Billing and shipping addresses.
* **Financial Amounts:** Total tax, invoice total, amount due.
* **Line Items:** Product/service descriptions, product codes, unit prices, quantities, tax, and line totals.

**Benefits:**

* Handles poorly scanned and creased documents.
* Provides structured data from various invoice formats.

**Receipt Model**

**Purpose:**

* Extracts key details from receipts, focusing on amounts paid and transactional information.

**Extractable Fields:**

* **Merchant Information:** Name, phone number, address.
* **Financial Details:** Receipt total, tax, tip.
* **Transaction Time:** Date and time.
* **Line Items:** Item names, quantities, unit prices, total prices.

**Note:**

* Supports single-page hotel receipts, extracting additional fields like arrival and departure dates.

**Benefits:**

* Accurate extraction even from low-quality scans.
* Efficiently captures transactional details for financial records.

**ID Document Model**

**Purpose:**

* Analyzes and extracts information from United States driver's licenses and international passports.

**Extractable Fields:**

* **Personal Information:** First and last names, sex, date of birth, nationality.
* **Document Details:** Country/region of issuance, document number, machine-readable zone.
* **Additional Information:** Endorsements, restrictions, vehicle classifications.

**Important Consideration:**

* Data extracted is sensitive; ensure compliance with data protection laws and obtain necessary permissions for data storage and processing.

**Benefits:**

* Accurate extraction of key identity details.
* Supports compliance with identity verification processes.

**Business Card Model**

**Purpose:**

* Extracts contact information from business cards, which often have diverse designs and formats.

**Extractable Fields:**

* **Personal Information:** First and last names.
* **Contact Details:** Postal addresses, email, website addresses, various phone numbers.

**Benefits:**

* Efficiently captures contact information.
* Handles varied formats and design elements.

**W-2 Model**

**Purpose:**

* Extracts information from the IRS W-2 form used for reporting employees' wages and taxes withheld in the United States.

**Extractable Fields:**

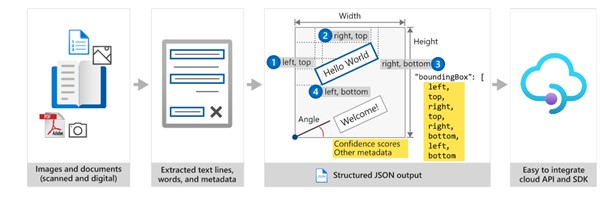
* **Employer Information:** Name, address.
* **Employee Information:** Name, address, social security number.
* **Tax Information:** Details about taxes paid by the employee.

**Benefits:**

* Accurate extraction of detailed tax information.
* Simplifies the processing of tax-related documents.

 **Azure Document Intelligence Overview:**

* Cloud-based AI service within Azure AI Services.
* Uses Optical Character Recognition (OCR) and deep learning models.
* Extracts text, key-value pairs, selection marks, and tables from documents.



 **OCR Capabilities:**

* Captures document structure using bounding boxes.
* Bounding boxes' locations recorded as coordinates relative to the page.
* Provides structured data with relationships from the original file.

 **Benefits:**

* High-accuracy data extraction without extensive model training.
* Trained on thousands of form examples.

 **Service Components:**

* **Document Analysis Models:**
  + Input formats: JPEG, PNG, PDF, and TIFF.
  + Output: JSON file with text locations, content, tables, selection marks, and structure.
* **Prebuilt Models:**
  + W-2 forms
  + Invoices
  + Receipts
  + ID documents
  + Business cards
* **Custom Models:**
  + Trainable on business-specific forms via Azure Document Intelligence Studio.

 **Accessing the Service:**

* **REST API**
* **Client Library SDKs:** Available in multiple programming languages.
* **Azure Document Intelligence Studio:** Visual tool for exploring and testing models.

**Preparation Steps:**

1. **Azure Resource Subscription:**
   * Ensure you have an active Azure subscription.
2. **Form Files for Data Extraction:**
   * Collect a selection of form files that you want to analyze.

**Subscribing to a Resource:**

1. **Accessing Services:**
   * **Azure AI Service Resource:**
     + Use if accessing multiple Azure AI Services.
     + Requires a multi-service subscription key.
   * **Azure Document Intelligence Resource:**
     + Use for Azure Document Intelligence only.
     + Requires a single-service subscription key.
     + Needed for Microsoft Entra authentication.
2. **Subscription Methods:**
   * Use the Azure portal or Azure CLI.
   * Learn more about CLI commands in the Azure documentation.

**File Input Requirements:**

1. **Accepted Formats:**
   * JPG, PNG, BMP, PDF (text or scanned), or TIFF.
2. **File Size Limits:**
   * Less than 500 MB for paid (S0) tier.
   * Less than 4 MB for free (F0) tier.
3. **Image Dimensions:**
   * Between 50 x 50 pixels and 10,000 x 10,000 pixels.
4. **Training Data Set Size:**
   * Must be 500 pages or less.

**Deciding on the Component to Use:**

1. **OCR Capabilities:**
   * **Models:** Layout, Read, General Document.
   * Purpose: Capture document analysis.
2. **Prebuilt Models:**
   * **Use Cases:** W-2s, Invoices, Receipts, ID documents, Health insurance, vaccination cards, business cards.
   * Note: These models do not need training and return a JSON output.
3. **Custom Models:**
   * **Use Cases:** Extract data from industry-specific forms.
   * **Requirement:** Train the model on sample documents.
   * **Output:** Analyzed new documents return a JSON output.

Azure Document Intelligence allows for supervised machine learning, enabling you to train custom models tailored to your specific form documents. Here’s a step-by-step guide on how to train these custom models:

**Steps to Train a Custom Model:**

1. **Prepare Sample Forms:**
   * Store sample forms in an Azure Blob container.
   * Include JSON files containing layout and label field information.
2. **Generate OCR and Label Files:**
   * Use Azure Document Intelligence's Analyze document function to generate an ocr.json file for each sample form.
   * Create a fields.json file that describes the fields you want to extract.
   * Create a labels.json file for each sample form, mapping the fields to their locations in the form.
3. **Generate SAS URL:**
   * Create a shared access security (SAS) URL for the Azure Blob container.
4. **Train the Model Using REST API or SDK:**
   * **Build Model:**
     + Use the Build model REST API function or the equivalent SDK method to train the model.
   * **Retrieve Model ID:**
     + Use the Get model REST API function or the equivalent SDK method to obtain the trained model ID.

**Alternative Method: Using Azure Document Intelligence Studio**

Azure Document Intelligence Studio offers a user-friendly interface for labeling and training models. You can choose between two types of custom models:

1. **Custom Template Models:**
   * **Features:** Accurately extract labeled key-value pairs, selection marks, tables, regions, and signatures.
   * **Training Time:** Only takes a few minutes.
   * **Language Support:** Supports more than 100 languages.
   * **Use Case:** Best for structured documents with predictable layouts.
2. **Custom Neural Models:**
   * **Features:** Deep learning models that combine layout and language features to accurately extract labeled fields.
   * **Use Case:** Ideal for semi-structured or unstructured documents where the layout is less predictable.

**Key Components for Custom Model Training:**

* **Sample Forms:** The documents you want the model to learn from.
* **OCR Files (ocr.json):** Generated by analyzing the document to capture the layout and text.
* **Fields File (fields.json):** Defines the fields you aim to extract.
* **Labels File (labels.json):** Maps fields to their corresponding locations in each form.
* **SAS URL:** Provides secure access to your Azure Blob container for training data.

**Summary:**

* **Preparation:** Collect and store sample forms and JSON files in Azure Blob.
* **Generate Files:** Create ocr.json, fields.json, and labels.json.
* **SAS URL:** Generate a secure URL for accessing the container.
* **Train Model:** Use REST API/SDK or Azure Document Intelligence Studio.
* **Choose Model Type:** Select between Custom Template Models for structured documents and Custom Neural Models for semi-structured/unstructured documents.

If you've created and trained custom models in Azure AI Document Intelligence, you can combine them into a single composite model and publish that as a single service. Composite models help when there are multiple versions of a form in use or when users find it difficult to keep track of the correct model for each form. They can also assist when you want customers to upload different types of documents to a single location for analysis and you don't know which type was uploaded.

**Composed Models in Azure AI Document Intelligence:**

* **Purpose:**
  + Useful when handling multiple custom models and form versions.
  + Automatically selects the best custom model for each form submitted.
* **Usage:**
  + Create and assemble custom models into a composed model using Azure AI Document Intelligence Studio or the StartCreateComposedModelAsync() method.
  + Submit forms for analysis using the composed model's ID.
  + The docType field in the results indicates which custom model was used.
* **Model Limits by Tier:**
  + **Free (F0) Tier:**
    - Custom Template: 500 models
    - Custom Neural: 100 models
    - Composed: 5 models
  + **Standard (S0) Tier:**
    - Custom Template: 5000 models
    - Custom Neural: 500 models
    - Composed: 200 models
  + Maximum of 100 custom models can be added to a single composed model.
* **Custom Model Compatibility:**
  + Custom template models are composable with other custom template models (versions 3.0 and 2.1).
  + Custom neural models are composable with other custom neural models.
  + Custom neural models cannot be composed with custom template models.

Azure AI Search can index content in many different formats to ensure users can locate the information they need to do their jobs. It's also extensible - skills can be added to add extra processing during indexing. If you add a skill that calls Azure AI Document Intelligence, you can use your Azure AI Document Intelligence models to enrich your AI Search index.

**Azure AI Search Enrichment Pipelines and Integration with Azure AI Document Intelligence:**

**Azure AI Search Indexing Process**

1. **Document Cracking:** Extracts content from files.
2. **Field Mappings:** Maps fields like titles, names, and dates for indexing.
3. **Skillset Execution:** Applies AI skills to enrich content.
4. **Output Field Mappings:** Maps output from skillsets to index fields.
5. **Push to Index:** Stores processed data in the search index.

**AI Search Skillsets**

* **Built-In Skills:** Pretrained AI models provided by Microsoft.
  + Key Phrase Extraction
  + Language Detection
  + Merge
  + Sentiment Analysis
  + Translation
  + Image Analysis
  + Optical Character Recognition (OCR)
* **Custom Skills:**
  + **Azure Machine Learning (AML) Custom Skills:** Enrich index using AML models.
  + **Custom Web API Skills:** Call web services to enrich the index.
    - Must accept JSON input and return JSON output.
    - Output should have a top-level values entity containing objects with recordId, data, warnings, and errors.

**Integrating Azure AI Document Intelligence with AI Search**

1. **Create Azure AI Document Intelligence Resource:** In your Azure subscription.
2. **Configure Models in Azure AI Document Intelligence:**
   * Use prebuilt models (e.g., Invoice, Business Card) or train custom models.
3. **Develop and Deploy a Web Service:**
   * Can use Azure Functions to host this service.
4. **Add Custom Web API Skill to AI Search Skillset:**
   * Configure to send requests to the web service.

**Application Example**

* **Polling Company Use Case:**
  + Users submit queries to locate completed polling forms by voter ID.
  + Train Azure AI Document Intelligence model to extract voter IDs from forms.
  + Ensure extracted voter IDs are included in AI Search index.
  + Use custom skills to call the Document Intelligence model from the search service.

**Summary**

* **Indexing Content:** Efficiently locate documents by indexing words and fields.
* **Skillset Execution:** Enhance indexing with built-in and custom AI skills.
* **Custom Web API Skills:** Integrate external AI services, like Document Intelligence, to enrich data.
* **Azure Functions:** Host web services for custom skills integration.
* **Enrichment Pipelines:** Automate selection of appropriate models for analysis.