**CUSTOM OCR MODEL TRAINING**

**Pre-requisite:**

1. Push Invoices to Azure Blob Container.

* Go to Azure From Recognizer -> Select Storage Browser: -> Blob containers

1. Add container -> Provide Name for container
2. Inside the Container Created, upload Training dataset.
3. From context Menu, Generate SAS
4. Provide permissions, set expiry time (as per reference video)
5. Click generate SAS token and copy SAS URL.

**Custom OCR Training**

|  |  |
| --- | --- |
| UI for Training | [Home - Form OCR Testing Tool (fott-2-1.azurewebsites.net)](https://fott-2-1.azurewebsites.net/) |

* 1. First make a connection to Blob Container
     1. Click on the Connection (\*plugin symbol)
     2. Display name: This should be your Dataset Name
     3. Paste the SAS URL copied during prerequisite
     4. Save Connection
  2. Go to Home -> Click [Use Custom to train a model with labels….]
     1. Create new project
     2. Display Name: This should be your Project Name
     3. Source connection: Your Connection Name would pop up here.
     4. **Form recognizer Service URL and API key are confidential** [Ask Sachin]
  3. Labelling the training dataset
     1. Create Custom Tags
     2. To choose any invoice and click on the invoice and wait for layout analysis to be completed.
     3. Select the value from the invoice and map the values to respective tags by clicking on the tag name.
     4. For Line Item details use table tag:

**Note**\*: Choose column dynamic type, when the line items count is not known in advanced else can used static column type

* 1. Provide name to Column as per requirement and mention column type.
  3. Repeat the steps for all invoices in training dataset

Train the model:

1. Click on the web symbol to Train
2. Provide Model Name
3. Click Train

Test the Model

* 1. Click on the Analyze tab on left panel of UI.
  2. Upload testing invoice to check for testing accuracy.

**Reference Link :**

* [Azure Form Recognizer | Custom Model | Invoice data extraction | #3 - YouTube](https://www.youtube.com/watch?v=hQ2NeO4c9iI)
* [Custom document models - Form Recognizer - Azure Applied AI Services | Microsoft Learn](https://learn.microsoft.com/en-us/azure/applied-ai-services/form-recognizer/concept-custom?view=form-recog-3.0.0&tabs=extraction%2Cclassification)