Use Gen AI to Identify and Validate Documents

Instructor-led lab guide

Last Updated: Friday, October 17, 2025

Paul Pacholski
pacholsk@ca.ibm.com

Harrish Narayanan
harrish.narayanan@ibm.com

Table of Contents

1 Lab Introduction

This hands-on lab explores IBM Business Automation Workflow (BAW) enhanced by Generative AI (GenAI) to streamline business processes and improve decision-making. Participants will gain practical experience integrating GenAI capabilities into workflows to automate content generation, enhance decision logic, and optimize task management. In this lab, attendees will learn how to leverage AI-powered intelligent documents within IBM BAW.

1.1 Lab Scenario - TBC

In this lab, you will complete the implementation of a Human Task Activity (highlighted in green in the figure below), which is part of a procurement Process.
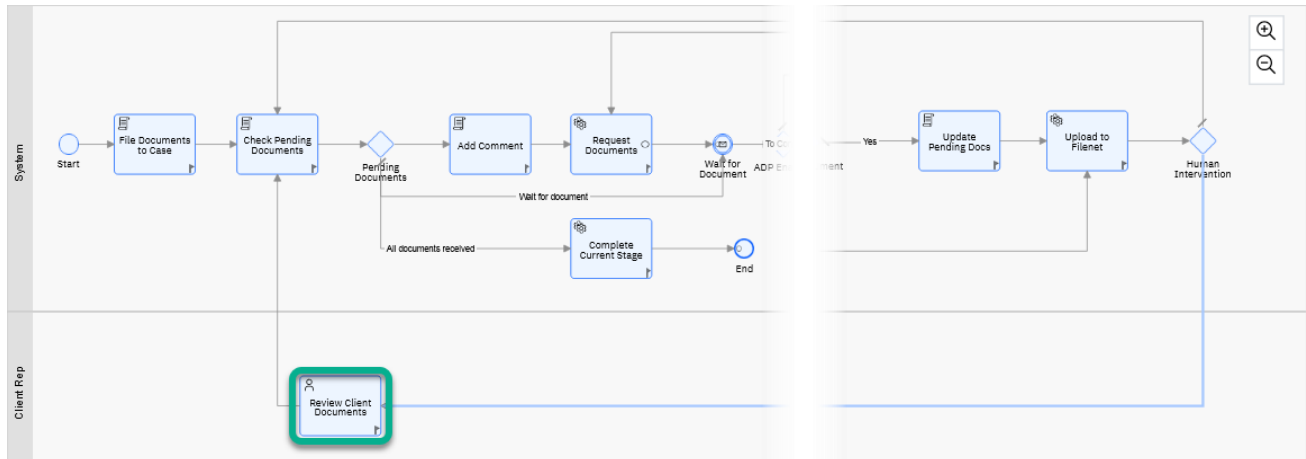
Figure 1. Client Onboarding Case: Review Documents Process

In this activity, a client onboarding specialist must review and validate the client documents. In this lab, you will create an automation that uses LLM to assist the client onboarding specialist. The LLM will automatically classify and summarize the document.

2 Lab Instructions

2.1 Download and Configure Project Files

_1. Open this document: https://ibm.ent.box.com/v/wxO-Agentic-SDK.

2.2 Reserve IBM watsonx Orchestrate ADK and DBA - US East Only

_1. Create a Tech Zone Reservation: **1.1 Reserve Tech Zone Environment**

> **Note:** This environment includes the Digital Business Automation VM, which provides the IBM Business Workflow runtime required for this lab.

2.3 BAW VM Setup

_1. Access the IBM DBA VM: **1.3.1 Access the DBA VM**

Note: To access the DBA VM you can use either the **Console** method

Virtual Machines

| Name	OS	IP	Status	
68eff80d926501b4673f452a-DBA-4-2022	Microsoft Windows Server 2022 (64-bit)	10.0.0.6	✓ Running	Console 🖥 ⋮
68eff80d926501b4673f452a-WxO-Lite-TechJam	Microsoft Windows Server 2022 (64-bit)	10.0.0.2	✓ Running	Console 🖥 ⋮

Or the **Remote Desktop** method.

_2. Start the BAW Server: **1.3.2 Start IBM Business Automation Workflow Server.**
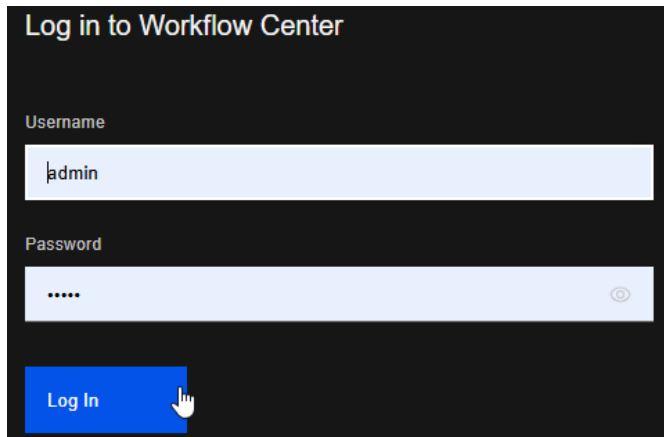
2.4 Open Web Process Designer

_1. Start the **Firefox** Web Browser.

_2. From the Web browser toolbar, select the **BAW > Workflow Center - Solution**.
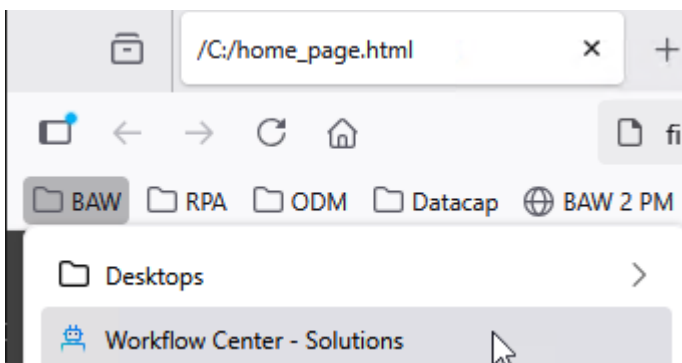
_1. For Username, enter **'admin',** and for password, enter **'admin'**. Then, click **Log In**.
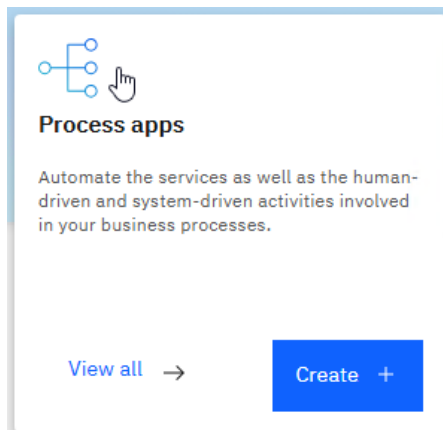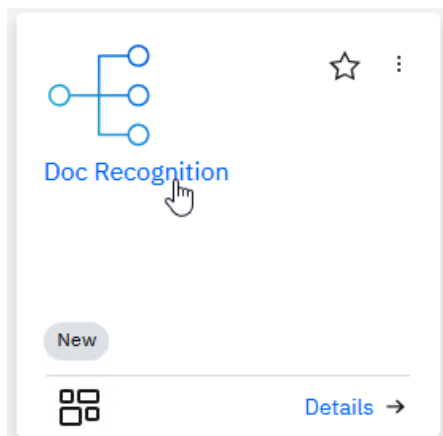

2.5 Open Process App

_1. Click **Process apps.**

_2. Click **Doc Recognition**.

2.6 Author GenAI Service Flow

You will now create a Service Flow with a GenAI Activity to assist knowledge workers in verifying and summarizing customer-submitted documents during the onboarding process.

For example, given a utility bill, Gen Ai will classify and summarize it as shown below.

The Gen AI Activity you will create shortly will use a one-shot prompt (no training) to enable it to summarize and classify three types of documents: Utility Bills, Back Statements, and Annual Reports.
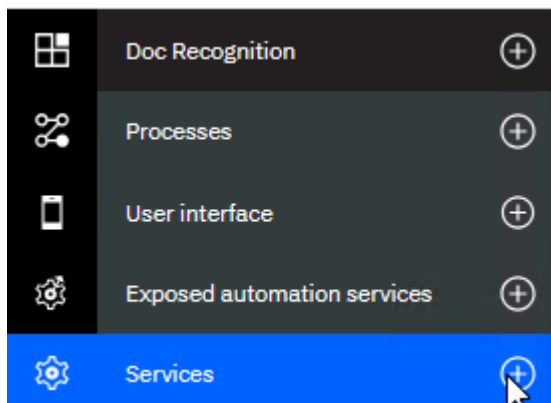
2.6.1 Create Service Flow

You will create a Service flow that, given a text extracted from a digital document, will classify and summarize it.
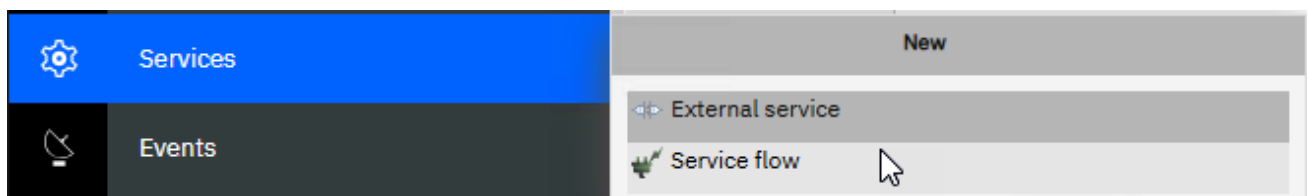
Note that we have already created a solution Service Flow for you. You can use it as a reference.
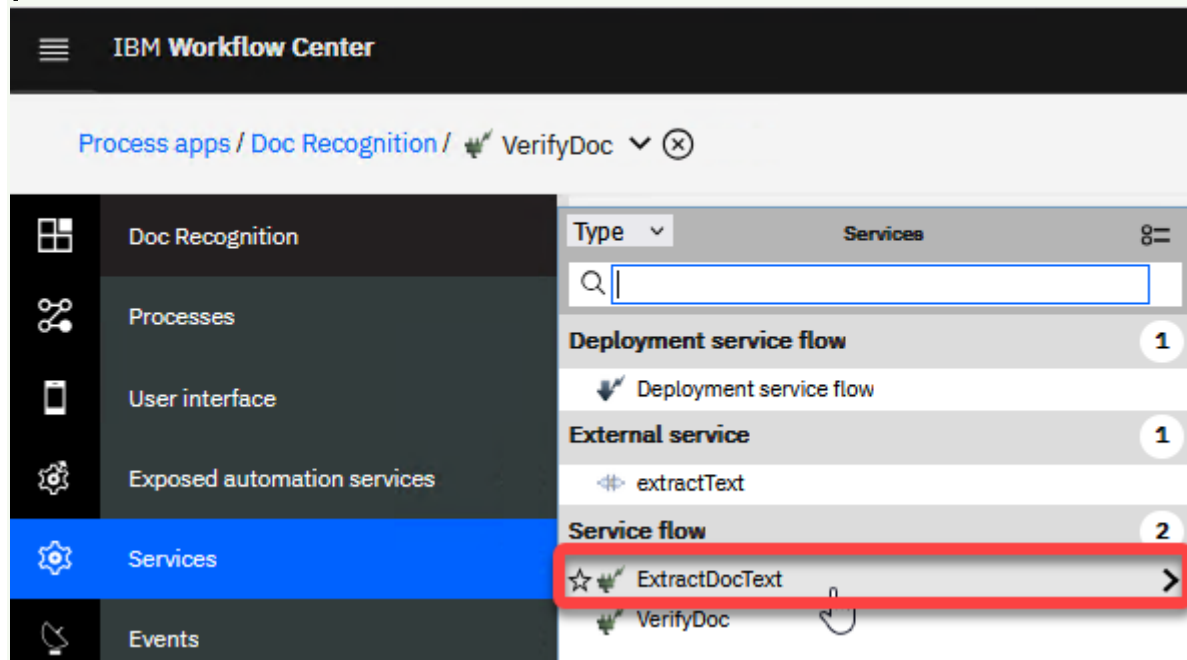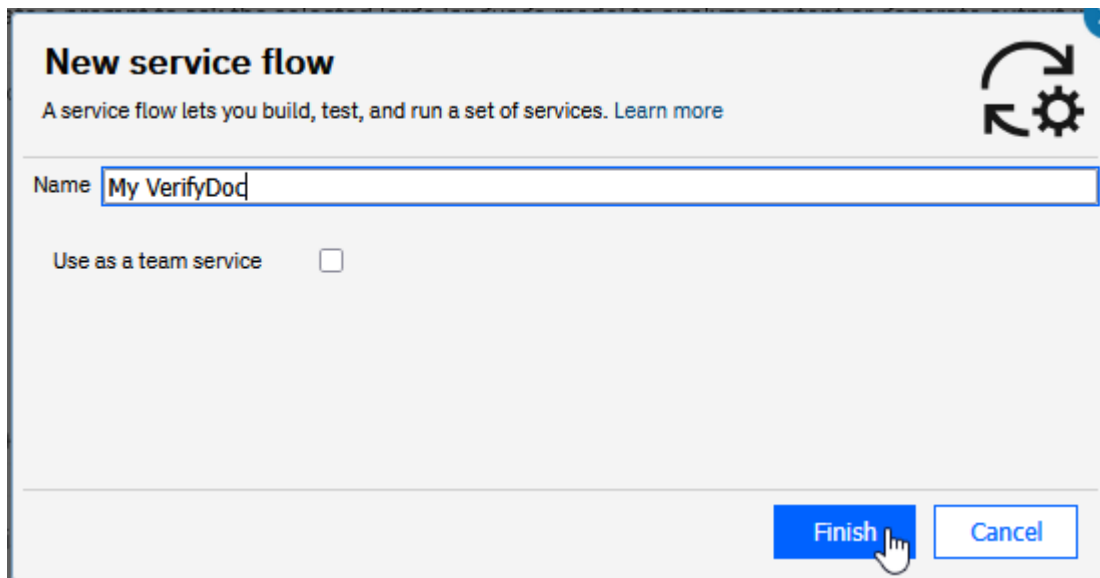

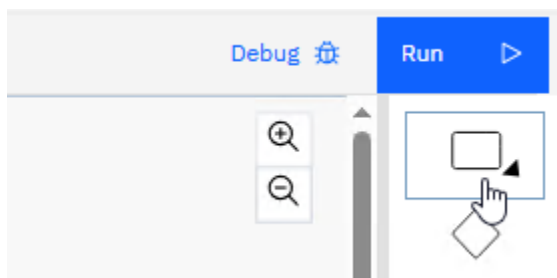_1. Click the **+** on *Services.*

_2. Click **Service flow**.
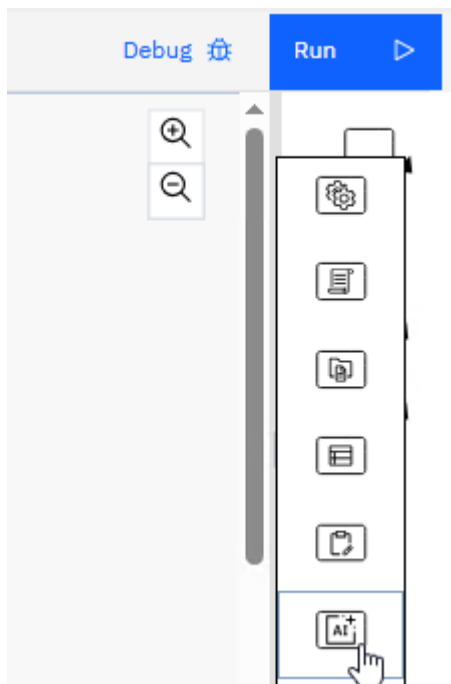
_3. Enter **My VerifyDoc** and click **Finish**.

2.6.2 Add GenAI Activity

_1. Click the **Activity** palette.

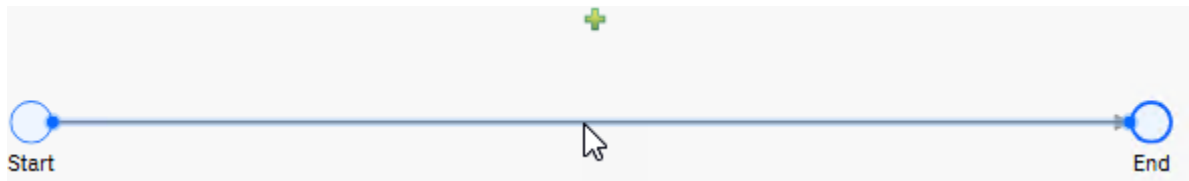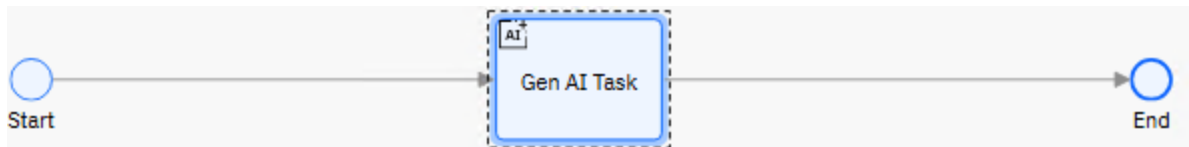
_2. Select the **Generative AI** activity.

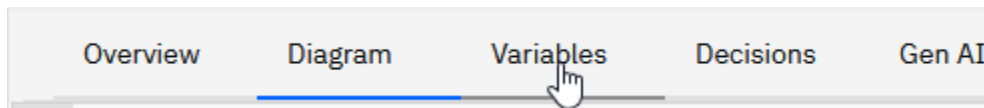_3. While holding the left mouse button, **drag and drop** it onto the wire between the Start and End nodes.

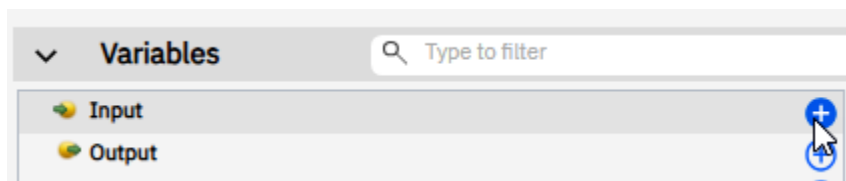_4. Ensure the Gen AI Task AI Activity has been added as shown below:
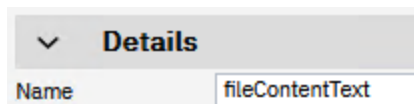

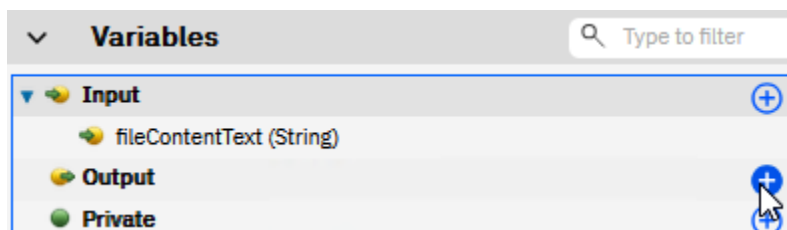### 2.6.3 Create Service Flow Variables

_1. Click the **Variables** tab.
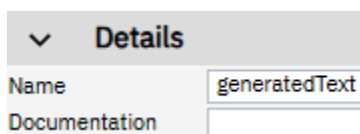

_2. Click the **+ icon** on the input.

_3. For *Name*, enter **fileContentText**
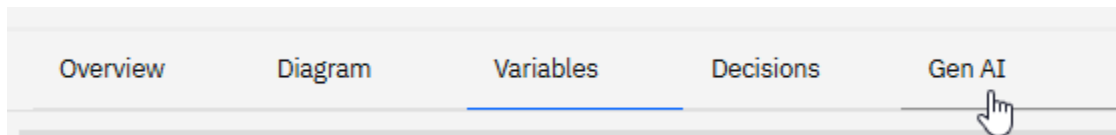

_4. Click the **+ icon** on the *output*.

_5. For *Name*, enter **generatedText.**
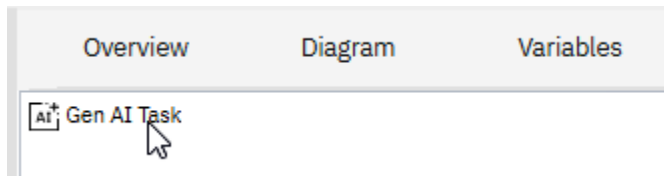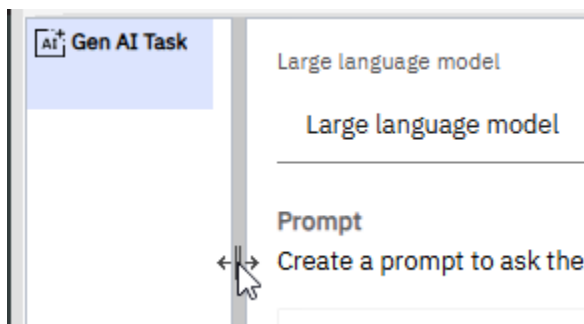
2.6.4 Author the GenAI Activity

_1. Click the **Gen AI** tab.

| Overview | Diagram | Variables | Decisions | Gen AI |

_2. Click **Gen AI Task**.

| Overview | Diagram | Variables |

[AI] Gen AI Task

_3. **Capture the divider and drag it to the left** to expand the Gen AI Editor.

[AI] **Gen AI Task**

Large language model

Large language model

Prompt
Create a prompt to ask the

2.6.5 Select the LLM

The IBM Business Automation Workflow is connected to an IBM watsonx.ai instance, which provides a selection of LLMs.

_1. From the Large language model dropdown, select **meta-llama/llama-3-2-11b-vision-instruct.**

Large language model

Large language model

⚙ **ibm/granite-3-8b-instruct**   8192 max tokens
The Granite model series is a family of IBM-trained, dense decoder-only models, which are particularly well-suited for generative tasks.

⚙ **ibm/granite-4-h-small**   0 max tokens
Granite-4.0-H-Small is a 30B parameter long-context instruct model finetuned from Granite-4.0-H-Small-Base using a combination of open

⚙ **ibm/granite-8b-code-instruct**   16384 max tokens
The Granite model series is a family of IBM-trained, dense decoder-only models, which are particularly well-suited for generative tasks.

⚙ **ibm/granite-guardian-3-8b**   8192 max tokens
The Granite model series is a family of IBM-trained, dense decoder-only models, which are particularly well-suited for generative tasks.

⚙ **ibm/granite-vision-3-2-2b**   8192 max tokens
Granite 3.2 Vision is a image-text-in, text-out model capable of understanding images like charts for enterprise use cases for computer

⚙ **meta-llama/llama-2-13b-chat**   4095 max tokens
Llama-2-13b-chat is an auto-regressive language model that uses an optimized transformer architecture.

⚙ **meta-llama/llama-3-2-11b-vision-instruct** 8192 max tokens
Llama-3-2-11b-vision-instruc is an auto-regressive language model that uses an optimized transformer architecture.

_2. Verify that the **meta-llama/llama-3-2-11b-vision-instruct** is your LLM.

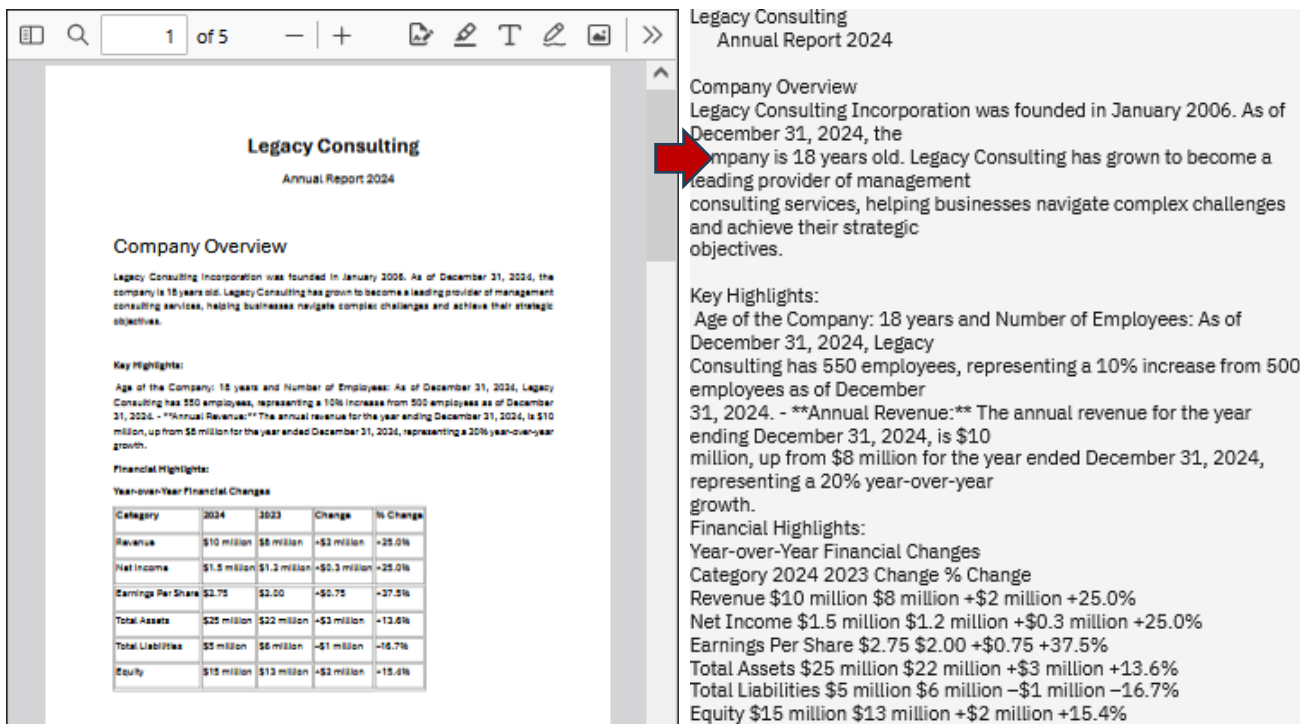Large language model

⚬ meta-llama/llama-3-2-11b-vision-instruct ⌄

> **Why did we choose this LLM over another?**
>
> The Meta-Llama 3.2-11B-Vision-Instruct model is a multimodal instruction-tuned LLM that can process both text and images. It excels at visual question answering, image captioning, document analysis, and multimodal chat — any scenario where understanding and reasoning over images (with or without accompanying text) is required. It has ~11 billion parameters and supports long-context reasoning (up to ~128K tokens).
>
> However, in this lab, to conserve the precious tokens, we will not supply raw images as the input to this model. Instead, we will first extract text from the image and use it as the input.

2.6.6 Create Input to LLM

The input variable will provide the LLM with the text extracted from the PDF quote file.

_1. Click **Add variables +** in the Variables section.

Variables

Add variables that you can use in your prompt. T

No variables

Select and add variables to use in your prompt.

Add variables +

_2. From the dropdown, select **"fileContentText"** and then click **"Add."**

_3. In the **Value column**, enter the following text for the quote Variable.

```
Banking Information       Focus Corp This banking information provided here will be used
by Focus Corp to withdraw the monthly services fee. Client Name Automation Elite Inc.
Bank Name  The Digital Bank Bank Phone Number +1-289-333-1234 Bank Address 2190 New
Orchard Road, Armonk, New York - 10504 Routing Number 91373084 Account Number 1179476345
Account Type  Checking   Savings   Other By submitting this form, you authorize Focus
Corp to withdraw the monthly services fee from he specified bank account. Date 4/21/2021
```

Variable	Value	
fileContentText	Corp to withdraw the monthly services fee from he specified|	🗑

2.6.7 Define Prompt Input
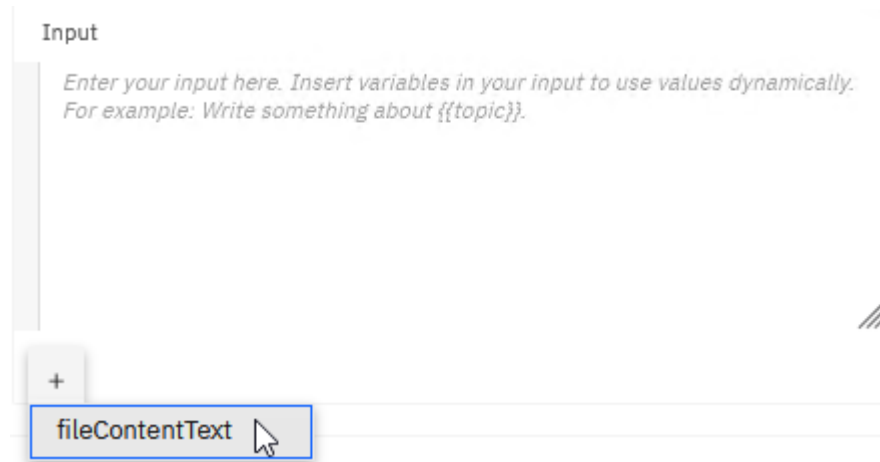
You will define the input the LLM will receive every time you call this Service Flow.

_1. In the *Input section* of the Prompt definition, click the **Add variable +** button.
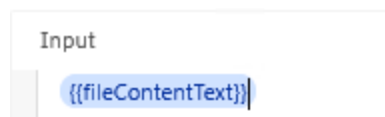
_2. Select **fileContentText.**.

_3. You should now see that the input to the LLM is the *fileContentText* variable you will pass to this Service Flow.

2.6.8 Define Prompt Context

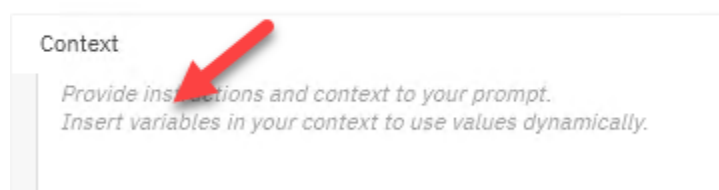You will define a prompt to instruct the LLM to categorize and summarize a document.

We will use a one-shot prompt with no training examples!

_1. In the **Context** section of the *Prompt* definition, add the **text below**.

```
Identify the document as Bill, Banking Information, or Annual Report.
Provide a summary of the document, including only the names or company names, addresses,
and dates.
Do not include any amounts.
```
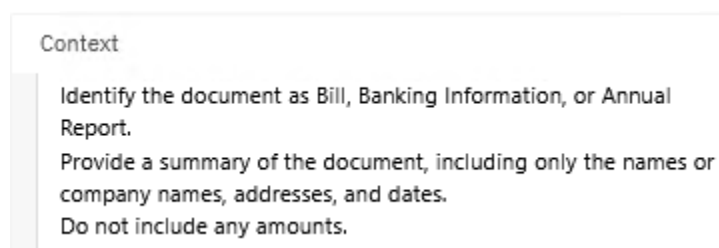

_2. The Context section should look exactly as shown below.

2.6.9 Define Parameters

To limit the length of the generated output, we can specify the minimum and maximum number of tokens that the LLM can generate from the Parameters pane. In our case, we will need more than the default 50 tonnes.

What is a token?

A token is a sequence of characters that holds semantic meaning for a model. Before LLM processing, the words in your prompt text are converted into tokens. Similarly, the raw output from the model is also in the form of tokens. These output tokens are then converted back into words to be displayed as the final result.

Each LLM has a maximum limit on the total number of tokens, which includes both the tokens in the input prompt and the tokens in the generated output. You can view those limits in the LLM descriptions in the Large language model dropdown menu. For more information about tokens, see Tokens and tokenization

_1. In the *Parameters* section, for *Maximum generated tokens*, enter **250**.

Parameters ⓘ

Constrain the length of the generated output.

Minimum generated tokens

1

Maximum generated tokens

250

2.6.10 Test the Gen AI Task

We can use the default value for the input quote variable to test our Gen AI Task.

Variable	Value	
fileContentText	Corp to withdraw the monthly services fee from he specified	🗑

_1. Click the **Generate** button.

Parameters ⓘ

Constrain the length of the generated output.

Generate ⟳ RAW

_2. You should now see that the document has been correctly identified and labeled as "Banking Information," with the document summary included below.

> Note the **Save as example** button. If you decide to train the model with a set of documents (multi-shot prompt), you can use it to add this run to the training example section. Also note that 99 tokens were generated, which is why we increased the maximum generated tokens to 250.

2.6.11 Map Gen AI Task Activity to Service Flow Variables

We will map the Gen AI Task output variables to the Service Flow's output variables. This will allow the callers of the Service Flow to interact with the Gen AI Task.

> Note that we do not need to map the qute variable explicitly because it has already been associated as input to the Gen AI Task in the Gen AI Editor.

_1. Click the **Diagram** tab.
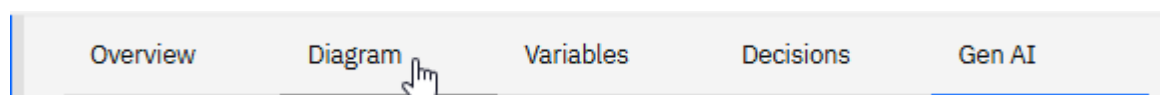
_2. Select the **Gen AI Task** activity.
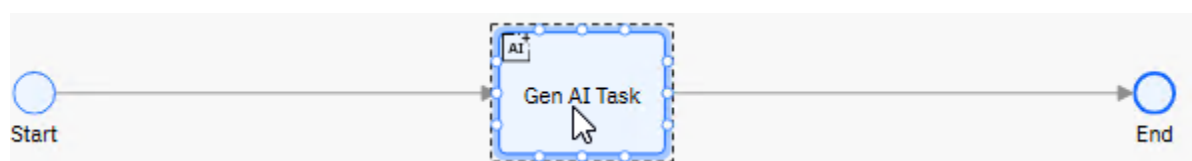
_3. Click the **Data Mapping** tab.

_4. On the *Generated text (string)* variable, click the **variable picker icon.**

_5. Select **generatedText.**

_6. Click the **"Magic Wand" icon.**

_7. Unselect the **"Generated text"** variable and click **Finish**.

This action will create private variables to map the Gen AI Task variables. We can use private variables since they don't need to be mapped as the Output of the Service Flow. The users of the Service Flows are not interested in the internal workings of the LLM!

_8. Click the **"Finish editing" icon** to ensure that all your work is saved.
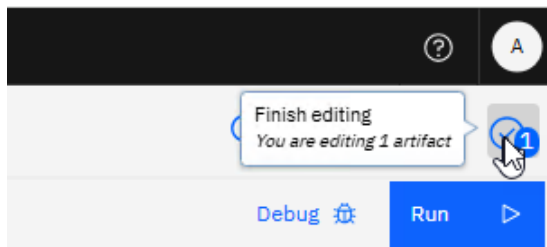
2.7 Add Gen AI Service Flow to Human Task

The Gen AI Service Flow (My Gen AI Document Processing) you just authored is intended for use in the Human Task invoked from the Porcurmente Porcess.

In this lab, we have partially authored the Human Activity (the official term is Client-side Human Service – CSHS) for you. Your task is to add and configure the "My VerifyDoc" Service Flow that you have just authored (My Gen AI Document Processing).

We have also created a partially implemented Human Task Activity for you: *My Document Identification and Validation.* You will need to configure it to invoke the *My VerifyDoc* Service Flow you just created.
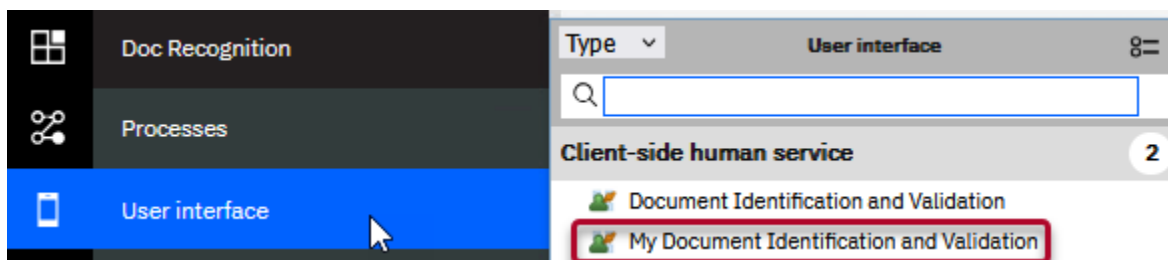
2.7.1 Examine the "My Document Identification and Validation" Client Side Human Service

_1. Click **User interface,** and then click **My Document Identification and Validation.**
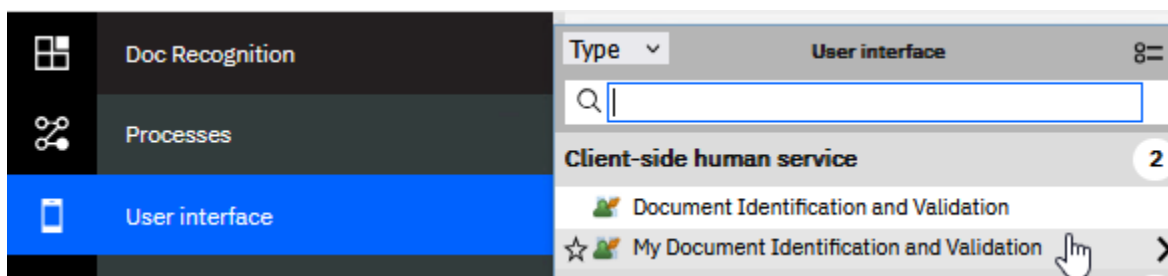
_2. Click the **Run** button.
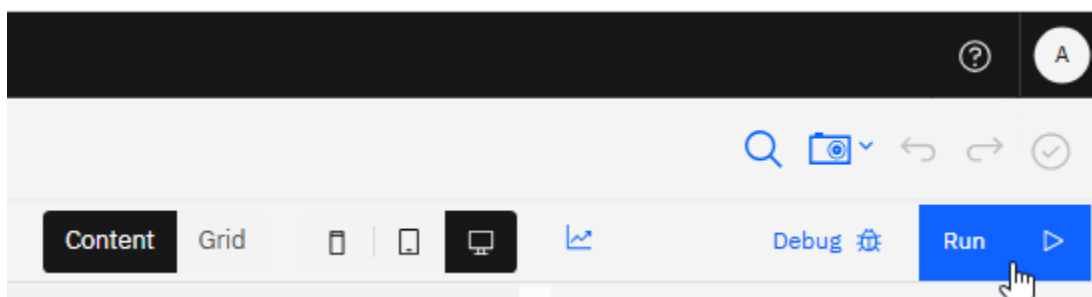
_3. Click the **"Open Document"** button, select **Banking Information – Automation Elite Inc.pdf**, and then clikc **Open.**

_4. Click the **Extract Text** button to extract text from the PDF document.

Document Contents:

Banking Information   Focus Corp

This banking information provided here will be used by Focus Corp to withdraw the monthly
services fee.

Client Name
Automation Elite Inc.

Bank Name
The Digital Bank

Bank Phone Number
+1-289-333-1234

Bank Address
2190 New Orchard Road, Armonk, New York - 10504

Routing Number
91373084

Account Number
1179476345

Account Type
 Checking Savings Other

By submitting this form, you authorize Focus Corp to withdraw the monthly services fee from
the specified bank account.

Date
4/21/2021

>|≡   Extract Text

_5. Click the "**Generate Summary**" button, and note that nothing happens.

Gen AI Generated Document Summary:

?

💡 Generate Summary
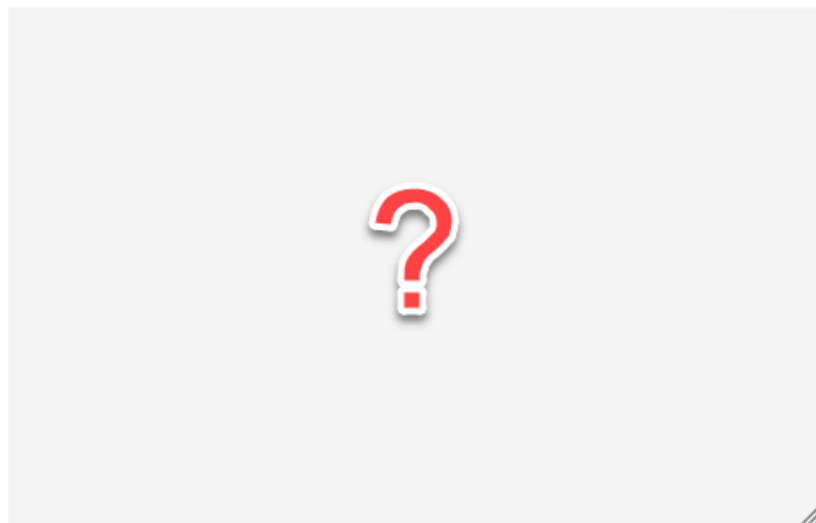
In the next few lab steps, you will configure the Client-side Human Service UI to send the document
text to the Gen AI Service Flow you created and display the document classification and summary.

_6. Click the **X** button to close the Chrome Web Browser Window.

2.7.2 Add "My VerifyDoc" Service Flow

_1. Click the **Diagram** tab.
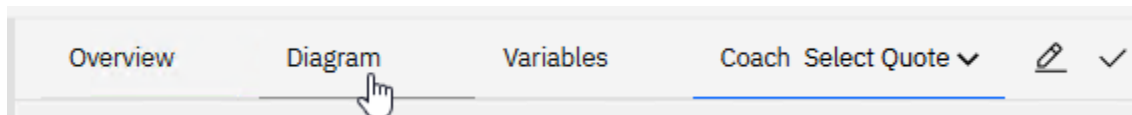
_2. Click **"Services,"** then drag and drop **"My VerifyDoc"** onto the Diagram.
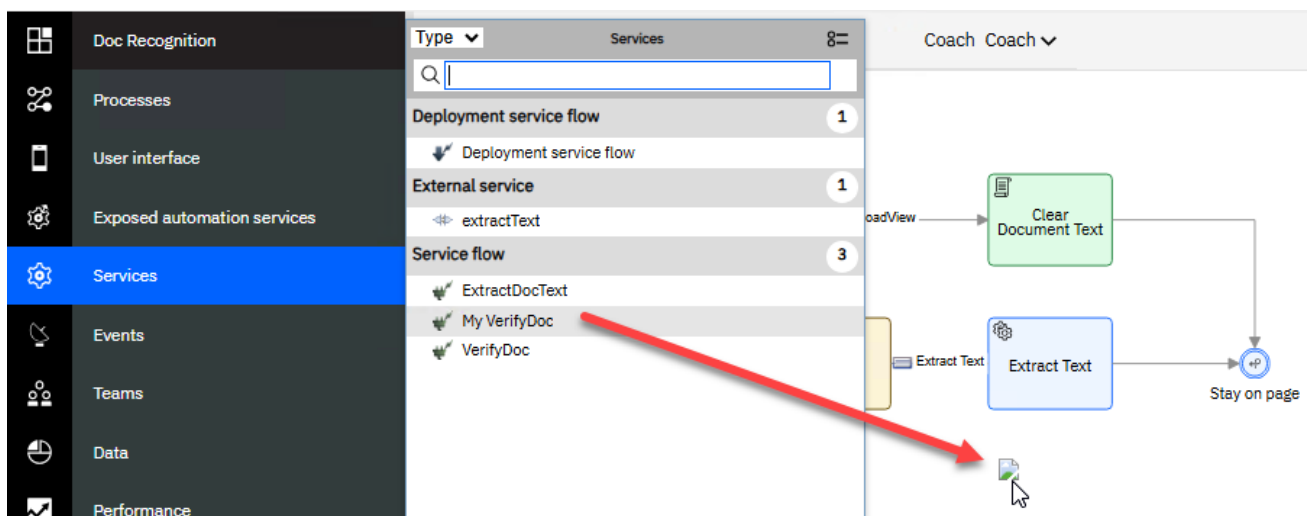
_3. Please verify that the Diagram now includes the **My VerifyDoc** Service Flow you dragged.
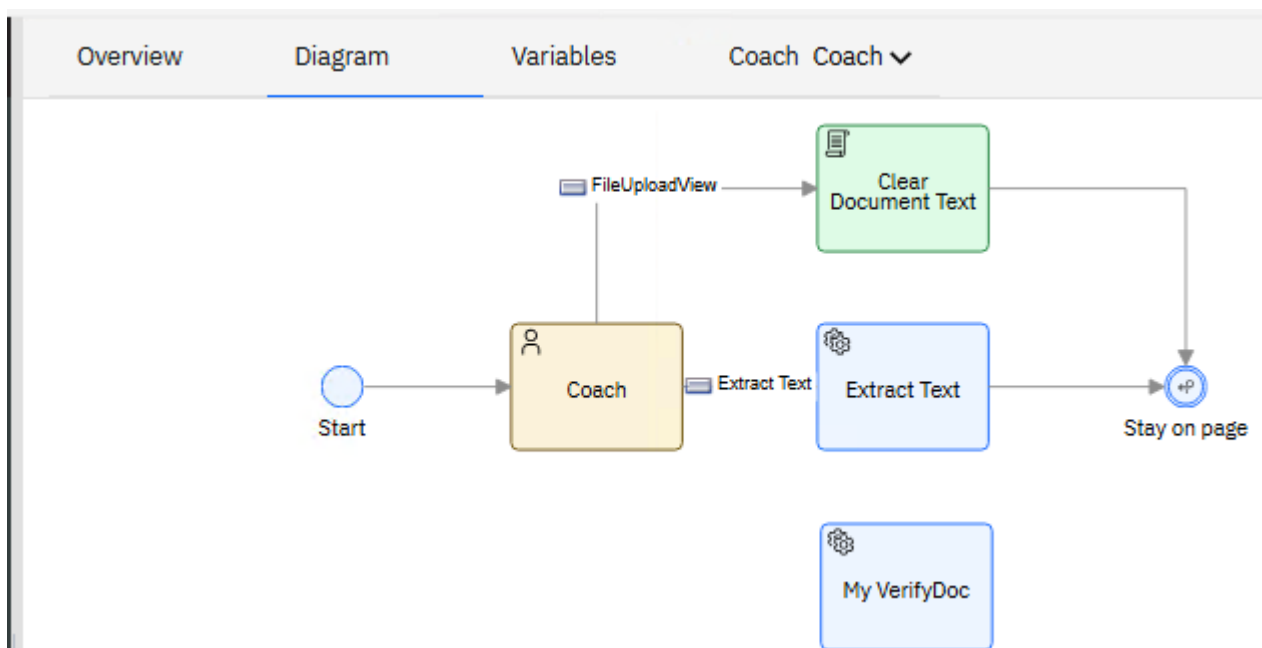
_4. Hover over the bottom right boundary of the **Coach** Activity – you should see a connecting arrow as shown below.
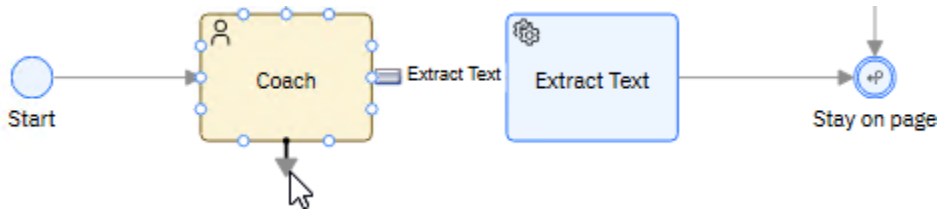
_5. While holding down the left mouse button on the connecting arrow, make a connection to **My VerifyDoc** Activity.
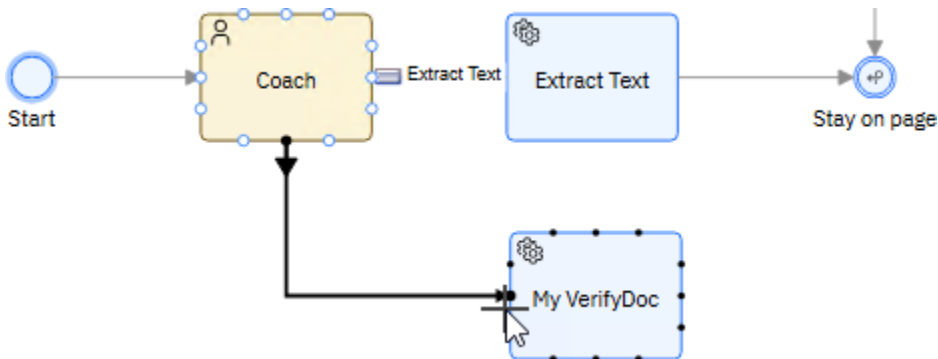
_6. Connect **My VerifyDoc** Activity to **Stay on Page**. Use the method described in steps 4-5.
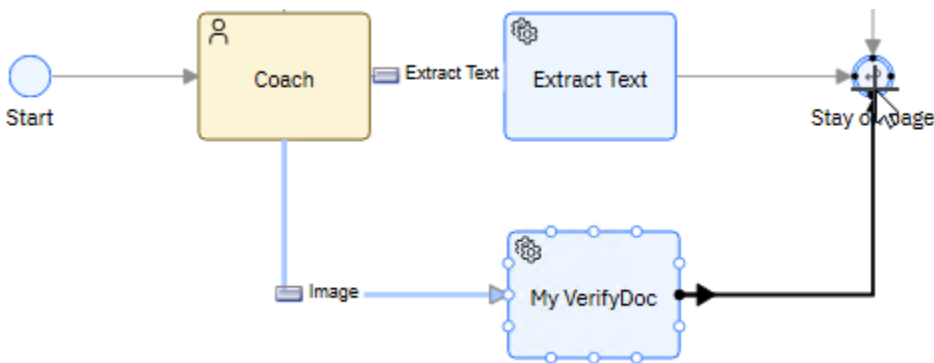
This connection ensures that after you execute the *My GenAI Document Processing* Activity, control returns to the UI.

2.7.3 Associate "Generate Summary" Button with "My VerifyDoc" Service Flow

By default, the Generate Summary Button in the Coach was not associated with My VerifyDoc Service Flow. Let's correct this.

_1. Note that the button associated to *Coach* and *My VerifyDoc is the Image button.*
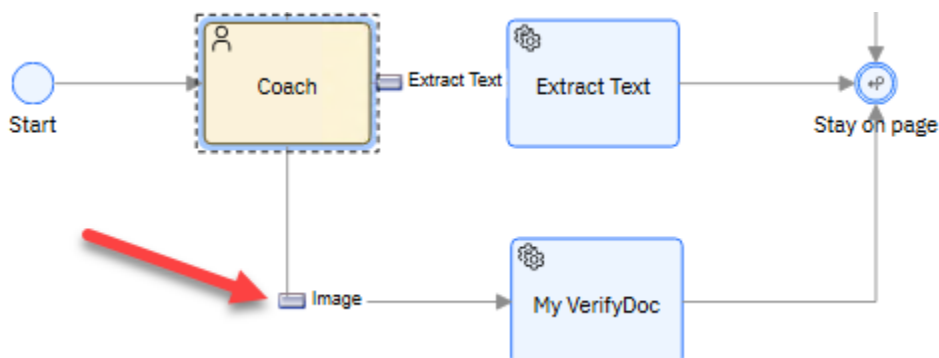
Note that the Image button is the Focus Corp logo!
We don't want the Gen AI service to start when we click the logo image.
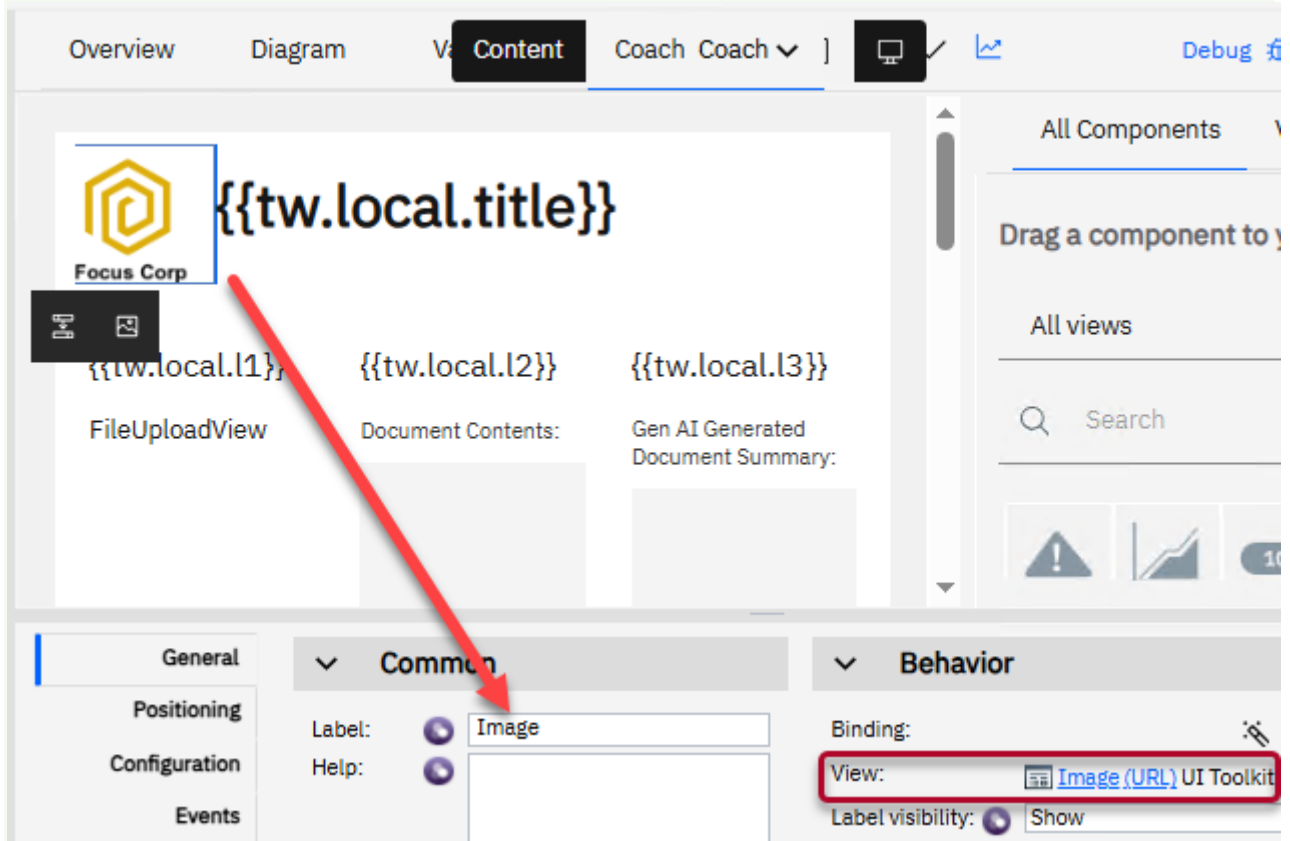
_2. Click the **Image** link.
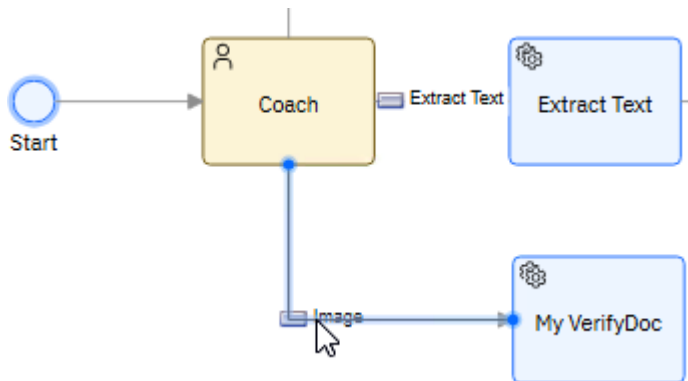
_3. Click the "**Select…**" button.
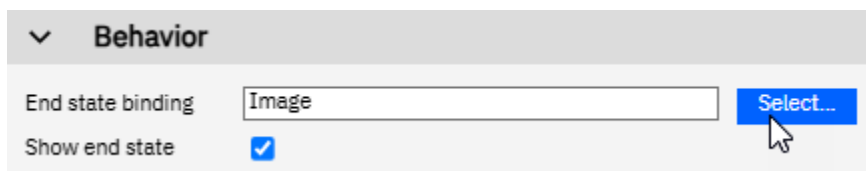
_4. Select the **Generate Summary** button and click **Finish**.
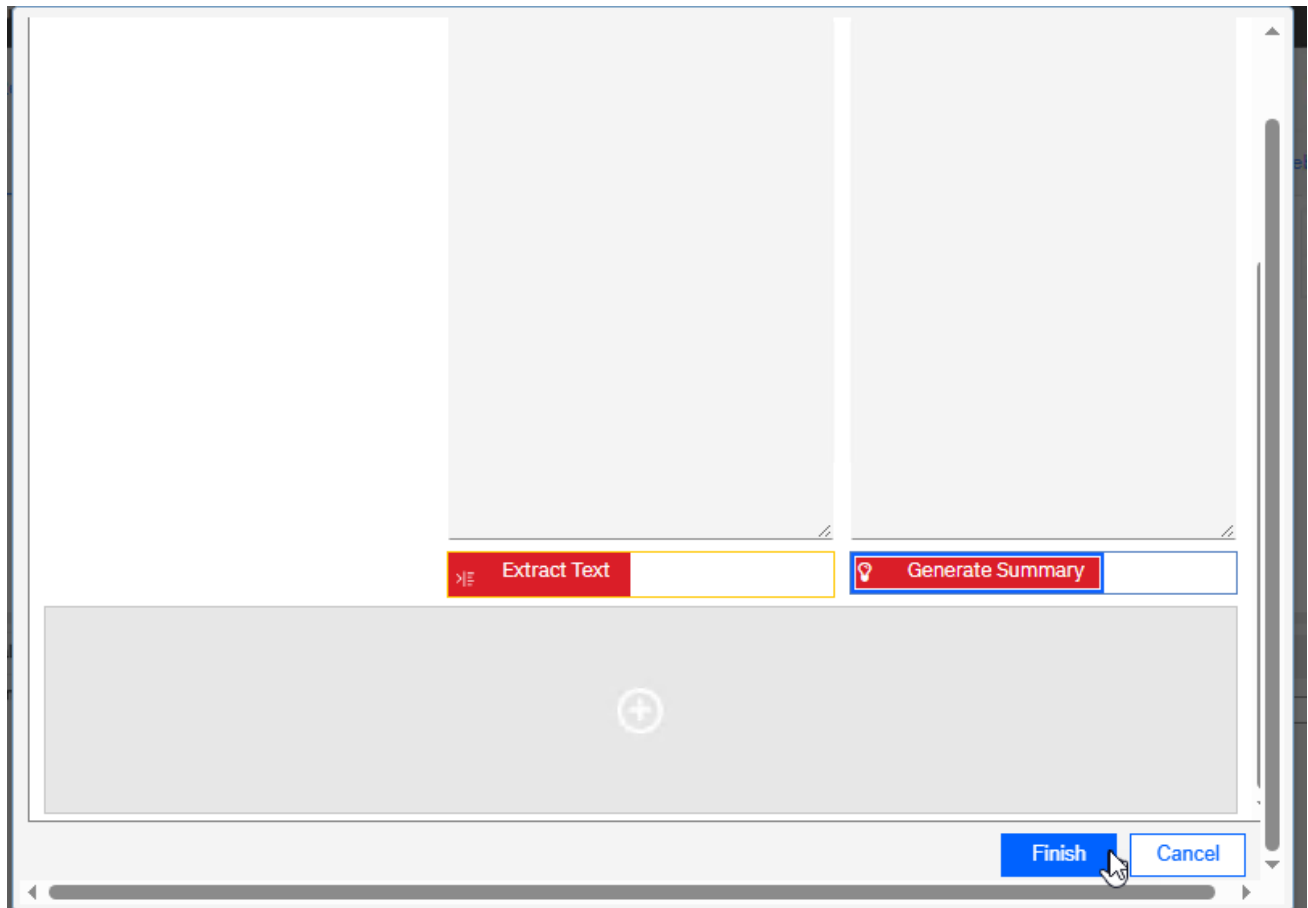
_5. Verify the **link** is now associated with the **Generate Summary** button.
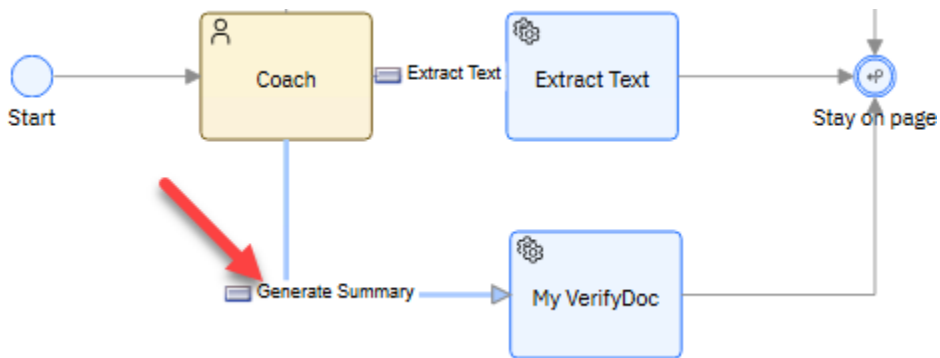
2.7.4 Map "My VerifyDoc" Service Flow Interface

Before we perform variable mapping, let's examine the figure below to understand the flow and the variables involved.
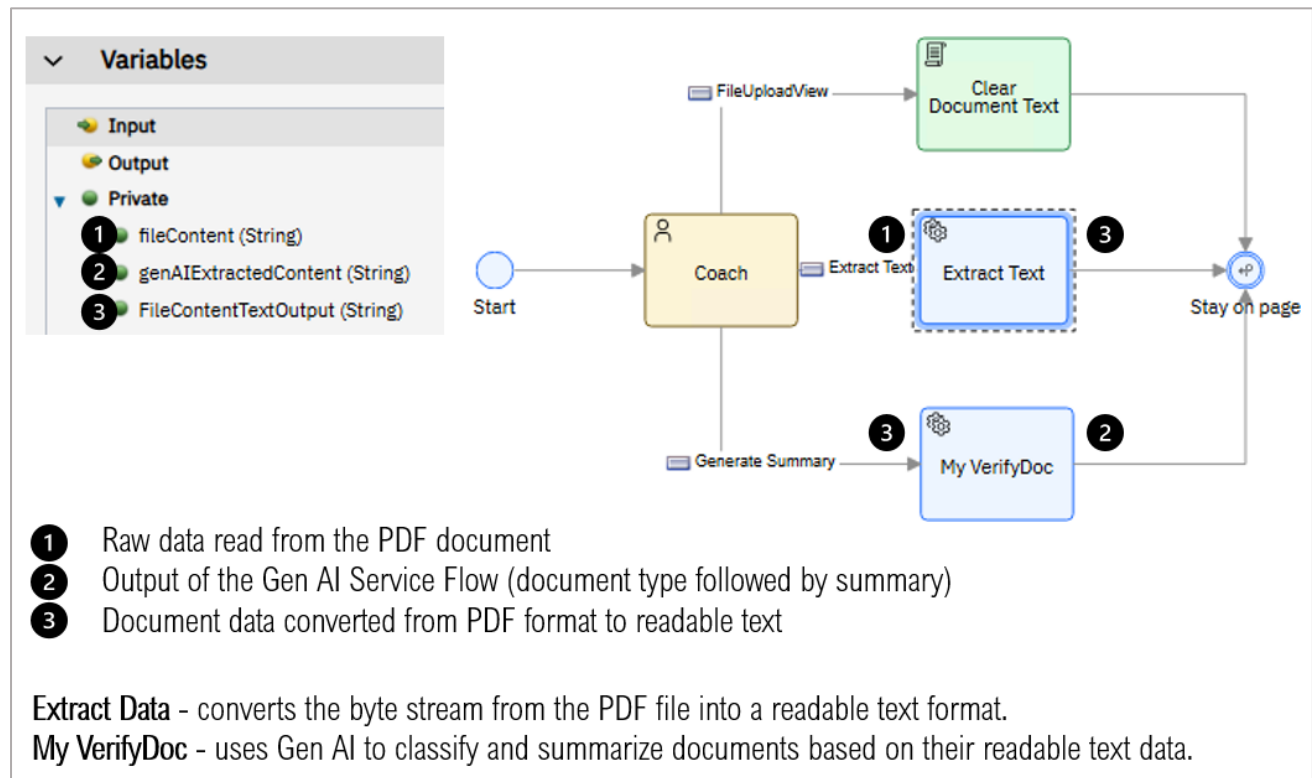
Figure 2. Service Flows and Variables

_1. Click the **"My VerifyDoc"** Activity.
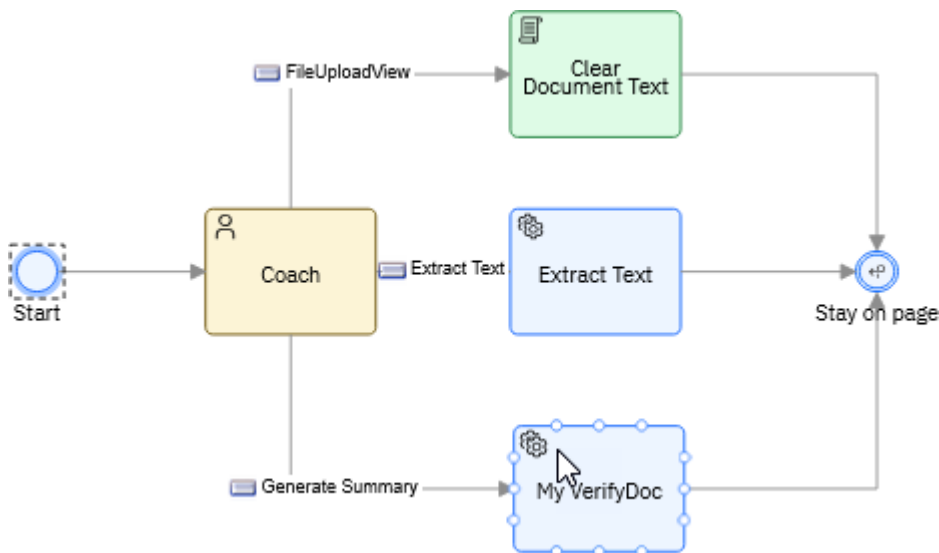
_2. Click the **Data mapping** tab.
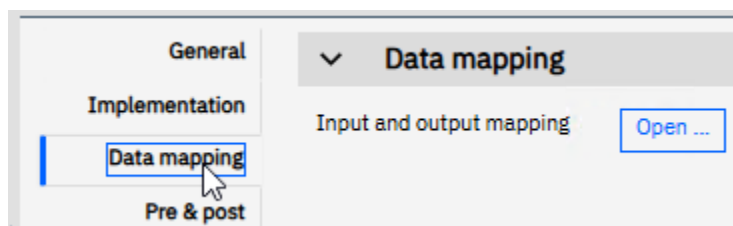
_3. Click **Open…** to open the Mapping Editor.
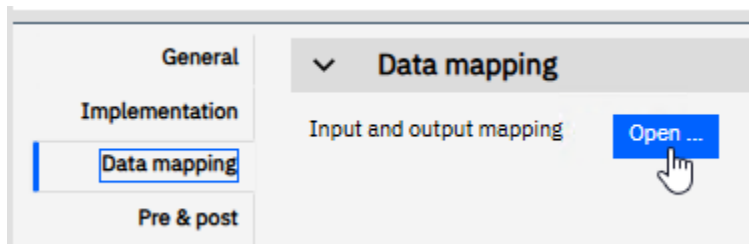
_4. Click the **Select a variable** icon.
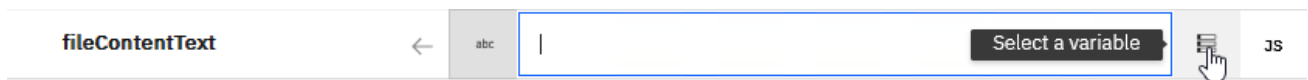
_5. Select the **FileContentTextOutput** variable.

_6. Click the **output mapping** tab.
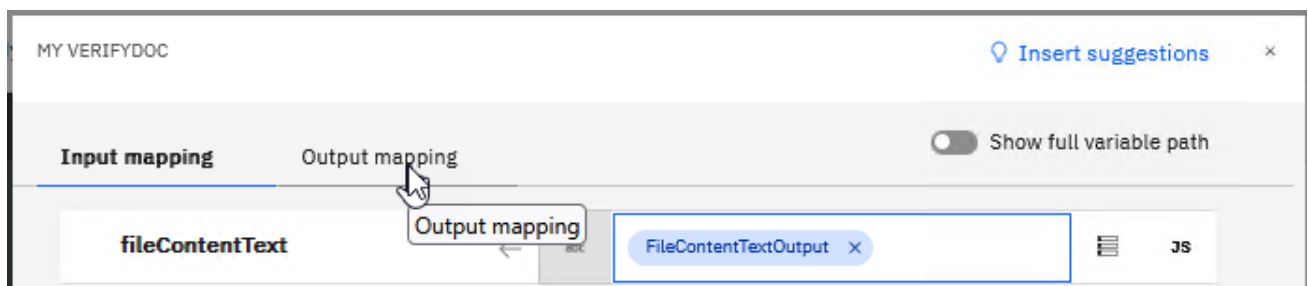
_7. Click the **Select a variable** icon.

_8. Select the **genAIExtractedContent** variable.

_9. Click **OK** to close the Mapping Editor.
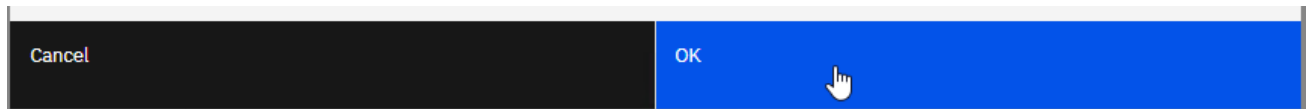
_10. Click the **"Finish editing" icon** to ensure that all your work is saved.
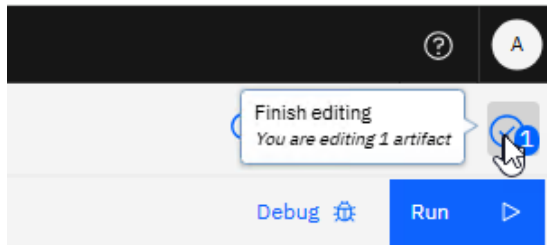
2.8 Test the "My Document Identification and Validation" CSHS

_1. Click the **Run** button.
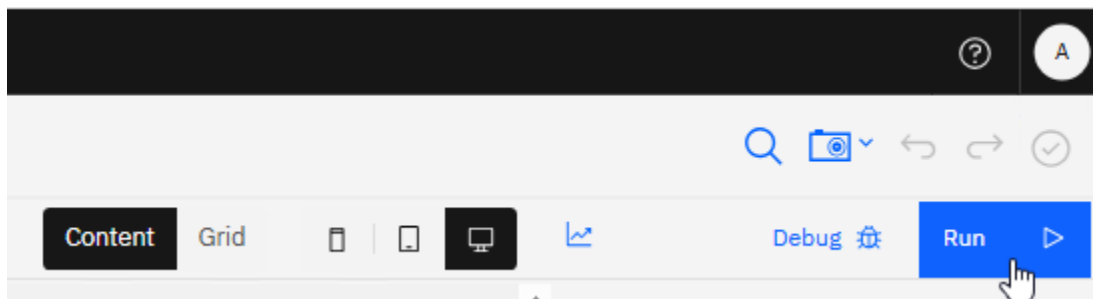
_2. Click the **Open Document** button.
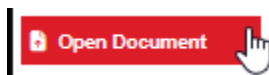
_3. From **C:\Documents,** select **Utility Bill - Automation Elite Inc.pdf** and click **Open.**
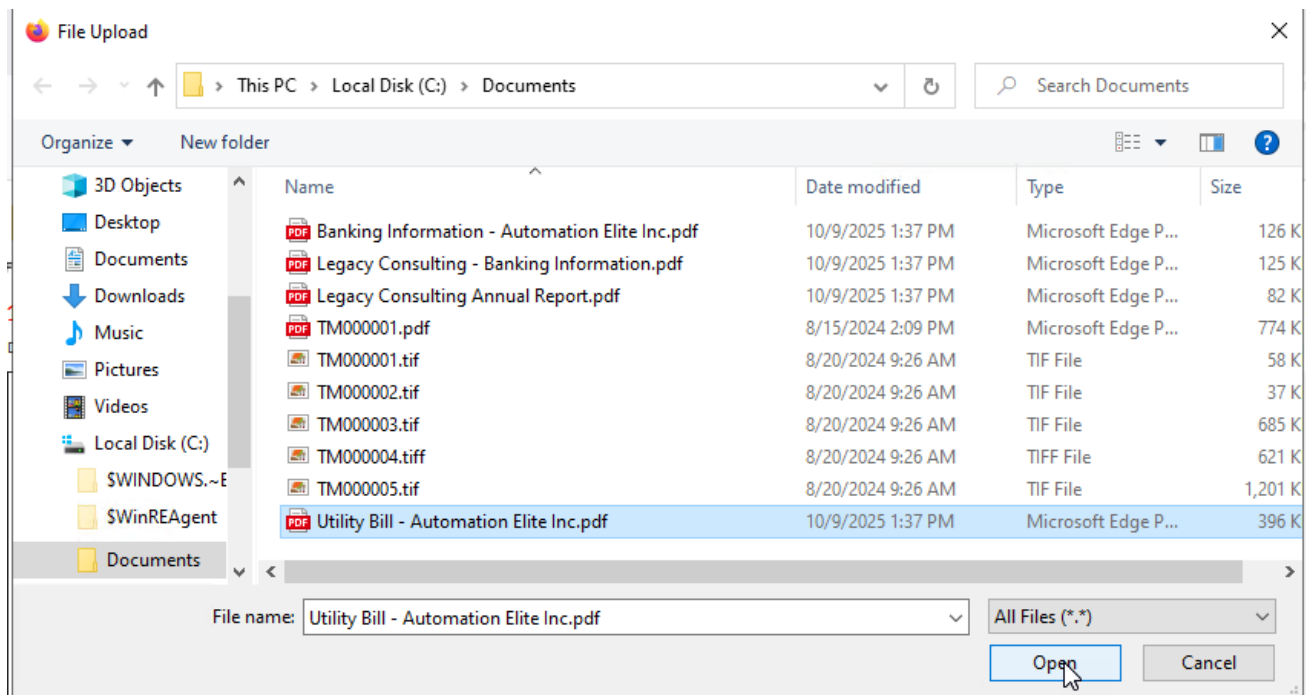
_4. Verify that the document image appears as shown below:

1. Upload Client Doucument

Document Image:

Open Document

_5. Click the **Extract Text** button.

_6. Verify that the extracted text appears as shown below:

2. Extract Text from Client Doucment

Document Contents:

```
Monthly Statement Summary
Previous Balance
Payment Received
New Charges - see details
Total Amount Due
$ XXX
-$ XXX
$ XXX
$ XXX
Company
Logo
Company
Name/Logo
Account No.   XXXX
Bill Period.   XXXXX
Page 1/2
Bill Period
Bill Due On
Account Number  XXXX
Bill Period
Bill Due On
Amount Due.    $
Amount Enclosed
COMPANY NAME
PO BOX XXX
CITY, STATE, ZIP CODE
9 0 2 3 8 4 0 5 7 9 2 8 3 6 4 0 1 8 9 3 7 4 2 9 3 6 1 0 9 1 4 7 0 8 1 3 9 4 8 3 7
4 7 0 3 9 1 9 3 4 0 3 1 9 7 1
Service For
Ways to Pay:
You may pay your bill by sending a check
using the enclosed return envelope. You may
also opt to pay with your credit or debit card
via our website. You can also enroll your
card for auto debit option for your
convenience.
Bill Saving Tips:
Internet speed quality depends on the
number of users in your home and what
you're using it for. Grandparents aren't going
to have the same need for internet speed as
```

>|≡   Extract Text

_7. Click the **Generate Summary** button.
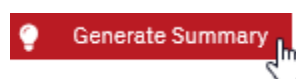
💡   Generate Summary

_8. Verify that the AI-generated output appears as shown below:

3. Generate Client Doucment Summary

Gen AI Generated Document Summary:

```
This document is a Bill.

Summary:

* Company: Automation Elite
* Address: 3974 Carson St, Lansing, MI 48911
* Account Number: 1109646555
* Bill Period: 04-01-2022 - 05-01-2022
* Bill Due On: 05-05-2022
* Contact Information: +1 800-354-8373, www.cableporium.com,
accessibility@company.com, or write to the address.
```

💡   Generate Summary

_9. If you like, repeat steps 2-8 to try the other four documents.

📄 Banking Information - Automation Elite Inc.pdf

📄 Legacy Consulting - Banking Information.pdf

📄 Legacy Consulting Annual Report.pdf

This marks the end of the lab.

Notices and disclaimers