Generative AI with SAP

Chetna Khaparde

Abstract

Generative Artificial Intelligence or GenAI is all over the news and gaining worldwide popularity.It falls under the broad category of machine learning. Generative AI enables quick generation of new content such as text, images, animation, sound, 3D models based on input prompts.

SAP is integrating Generative AI with extensive industry-specific data and deep process knowledge to create innovative AI capabilities for the SAP applications.For example, creating compelling job descriptions and interview questions, developing an AI copilot called "Joule" to complete tasks using natural language processing, expediting freight verification and documentation, generating ready-to-use process models and KPI recommendations, etc.This paper focuses on using Generative AI in SAP Business Technology Platform (SAP BTP) to process business documents in applications such as expediting freight verification, automatically processing payables, invoices, or payment notes while making sure that invoices and payables match.

Keywords:

SAP;
Generative AI;
GenAI;
Artificial Intelligence;
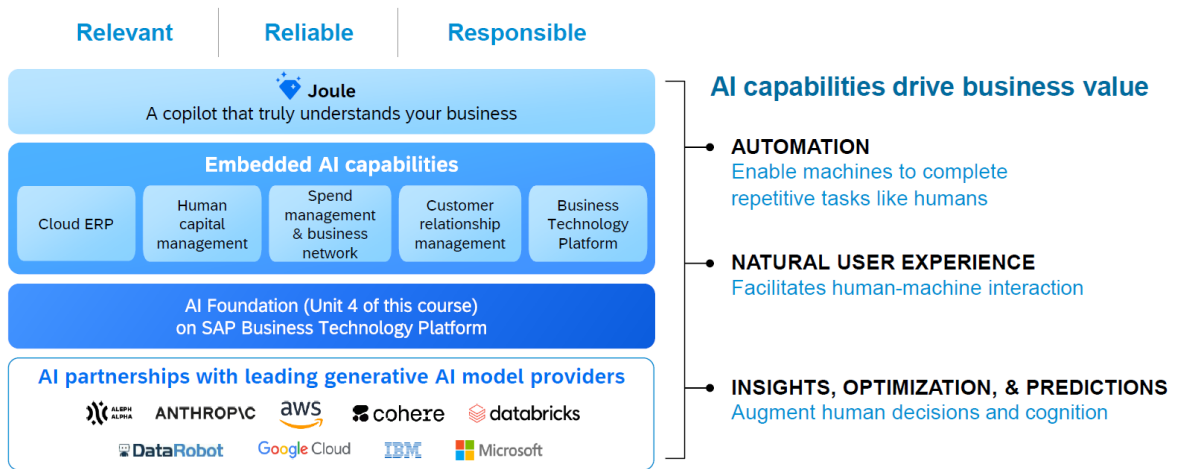AI.

Author correspondence:

Chetna Khaparde,
Advisor ERP Functional Analyst,
Gainwell Technologies LLC
Email: chetnakhaparde@gmail.com

1. Introduction

Extraction of data from business documents used to be a time-consuming and error-prone manual process. Manual data extraction could be either manual data entry or copying from source documents and pasting into the desired document. Both these methods are labor-intensive, high cost, limit scalability, and subject to human errors. Given the volume of business documents that businesses need to deal with today, it is important to embrace automation. Different automation technologies available today use Artificial Intelligence (AI) and Machine Learning (ML)to provide effective tools for document data extraction. One such technology is the Document Processing and Information Extraction using SAP BTP.

SAP offers AI foundation on Business Technology Platform (BTP) that has access to a variety of Foundation Models offered by renowned vendors such as – AWS, Microsoft, Google Cloud, IBM, etc.
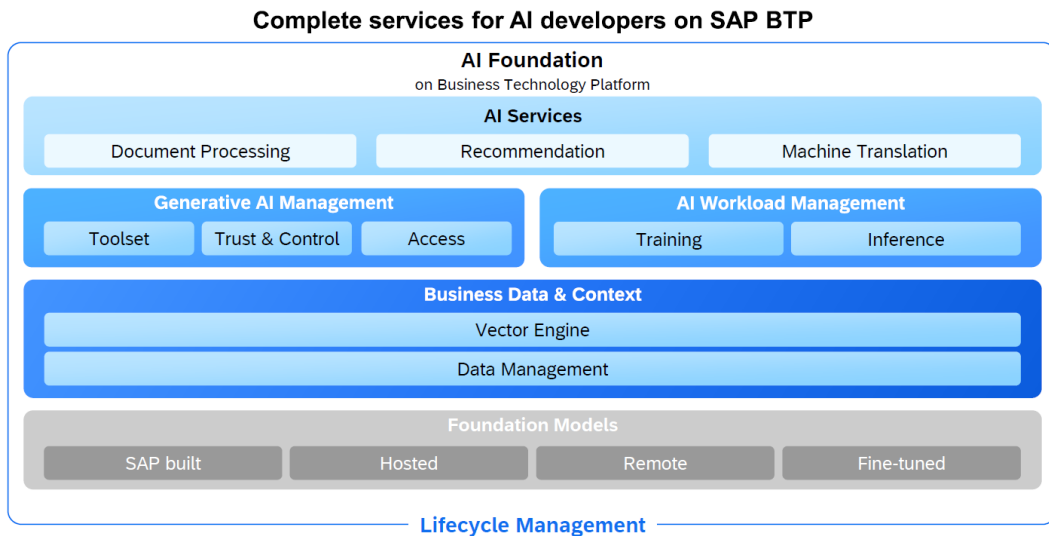
SAP Business AI product portfolio

Relevant | **Reliable** | **Responsible**

Joule
A copilot that truly understands your business

Embedded AI capabilities

| Cloud ERP | Human capital management | Spend management & business network | Customer relationship management | Business Technology Platform |

AI Foundation (Unit 4 of this course)
on SAP Business Technology Platform

AI partnerships with leading generative AI model providers

ALEPH ALPHA · ANTHROP\C · aws · cohere · databricks · DataRobot · Google Cloud · IBM · Microsoft

AI capabilities drive business value

- **AUTOMATION**
  Enable machines to complete repetitive tasks like humans

- **NATURAL USER EXPERIENCE**
  Facilitates human-machine interaction

- **INSIGHTS, OPTIMIZATION, & PREDICTIONS**
  Augment human decisions and cognition

3

SAP BTP AI Foundation is a starting point for SAP developers to build and extend applications leveraging the vast AI offerings.

AI Foundation

Complete services for AI developers on SAP BTP

AI Foundation
on Business Technology Platform

AI Services

| Document Processing | Recommendation | Machine Translation |

| **Generative AI Management** | | | **AI Workload Management** | |
| Toolset | Trust & Control | Access | Training | Inference |

Business Data & Context

Vector Engine

Data Management

Foundation Models

| SAP built | Hosted | Remote | Fine-tuned |

Lifecycle Management

The AI Foundation includes everything needed from embedding business-ready AI into SAP applications as extensions, like ready-to-use AI services. It has access to the business data and, most importantly, it has access to the foundation models, which are either SAP built, or partner built. It is designed to have security and governance in mind. Besides the possibilities of using different foundation models being hosted, it is possible to leverage grounding capabilities via the vector engine and the data management piece, to ensure that the enterprise data gives a proper context to your LLM so that hallucinations are mitigated.

1.1 Generative AI Hub in SAP AI Core

Generative AI Hub Architecture Overview
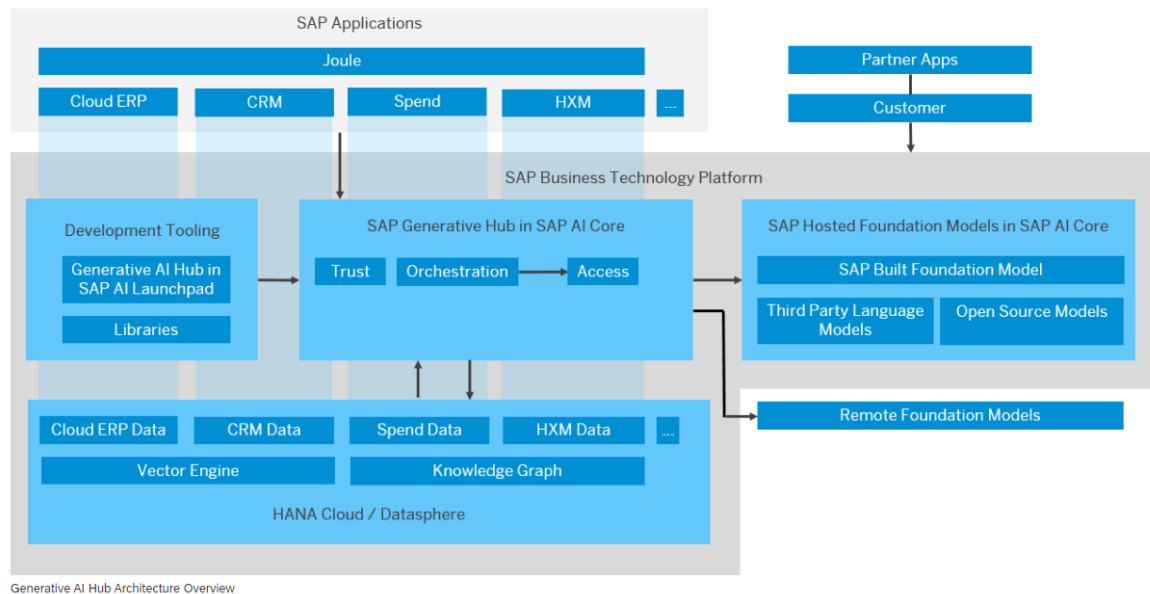
The generative AI hub incorporates generative AI into your AI activities in SAP AI Core and SAP AI Launchpad. Generative AI hub consist of the following main components-

1) **Development Tooling** – Purpose-build tools and services for the development phase including playground for experiments and testing, APIs with the SAP applications, and libraries and SDKs.

2) **Access to Foundation Models** – Instant access to the top-rated foundation models from multiple providers.SAP managed legal and commercial framework allows direct and instant access to any foundation model of our choice; we can also switch the foundation model in future to upgrade to the best suited technology.

3) **Trust**– Preserving security and privacy of the business data.Applications are typically native and embedded with SAP security.

4) **Control**–The HANA Cloud Vector Store Capabilities provide grounding capabilities to minimize hallucinations to model responses.

1.2 Document Information Extraction – Features, Compliance and Security

Document Information Extraction is an SAP BTP service that is used to extract information from business documents. It can process large volumes of business documents that have content in headers and line items, such as tables. The Large Language Models support more than 40 languages. The information, either structured or semi-structured can be extracted using pretrained machine learning models.

The following features are available for Document Information Extraction-

1) **Automate information extraction** – Automate the extraction of relevant information from business documents. The Document API takes document files as input and returns header fields and line items as structured data.

2) **Automate data enrichment** - Match a business document to enrichment data records based on the information extracted from the document. The Enrichment Data API takes document files as input and returns the ID of the matching enrichment data records.

3) **Benefit from multitenancy support** - Use this service in tenant-aware (multitenant) applications. Run them on a shared compute unit that can be used by multiple consumers (tenants).

Document Information Extraction ensures cloud security at multiple levels-

1) For the complete list of compliance and security standards that the Document Information Extraction service is compliant with, see SAP Business Technology Platform ISO Certificates - Compliance Document Finder | SAP Trust Center.

2) For the complete list of Service Organizational Control (SOC) audit reports available for the Document Information Extraction service, see SAP Business Technology Platform SOC Reports - Compliance Document Finder | SAP Trust Center.

2. Research Method

For the Document Information Extraction,we can use APIs or the Document Information Extraction UI to consume the service. In this example we will use the UI. An Enterprise Account with SAP with the entitlement "Document Information Extraction" with Service Plan "Premium Edition" or a BTP Trial Account is needed.

Steps to set up the trial account are as follows-

Create an Enterprise Account on SAP BTP- SAP BTP cockpit.
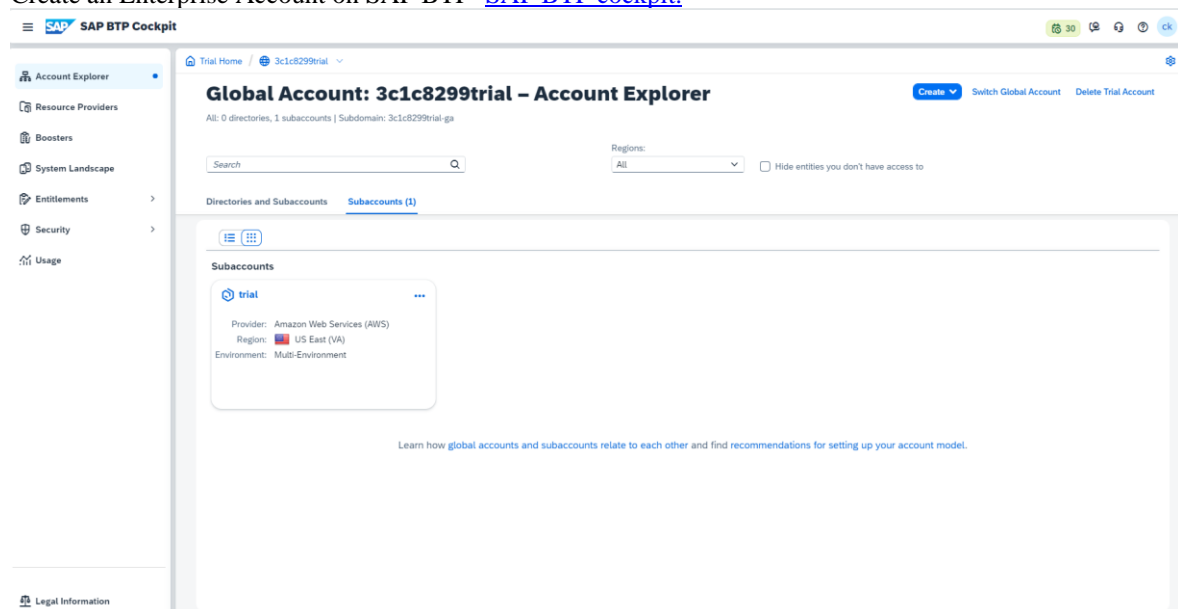
Select booster Set up account for Document Information Extraction to automatically assign entitlements, update your subaccount (or create a new one), create a service instance, subscribe to and access the Document Information Extraction UI. On the navigation side bar, click Boosters. Search for Document Information Extraction and click the tile to access the booster.
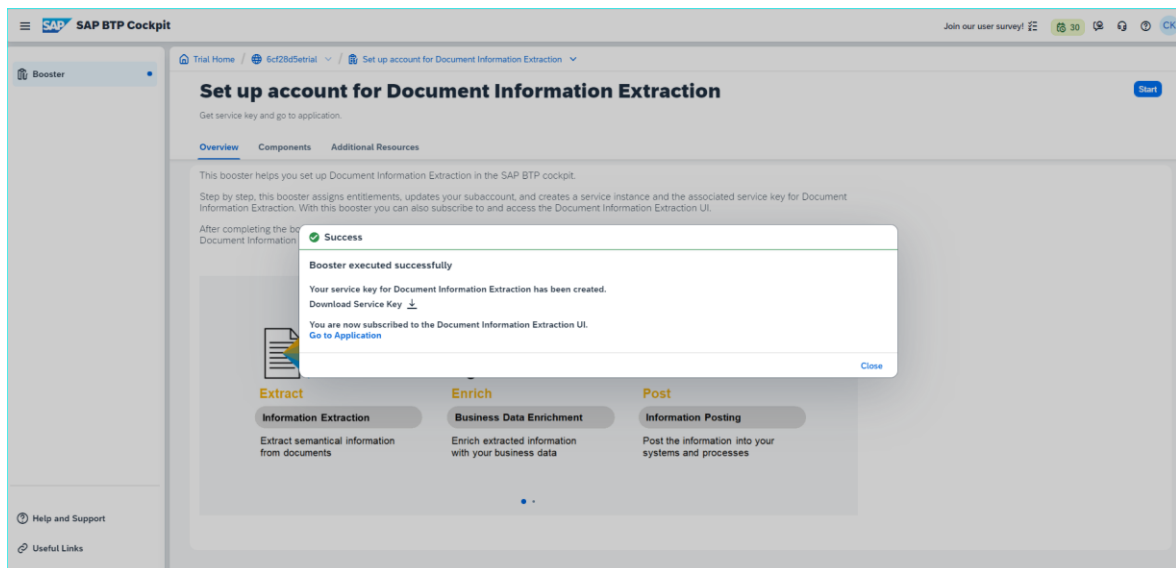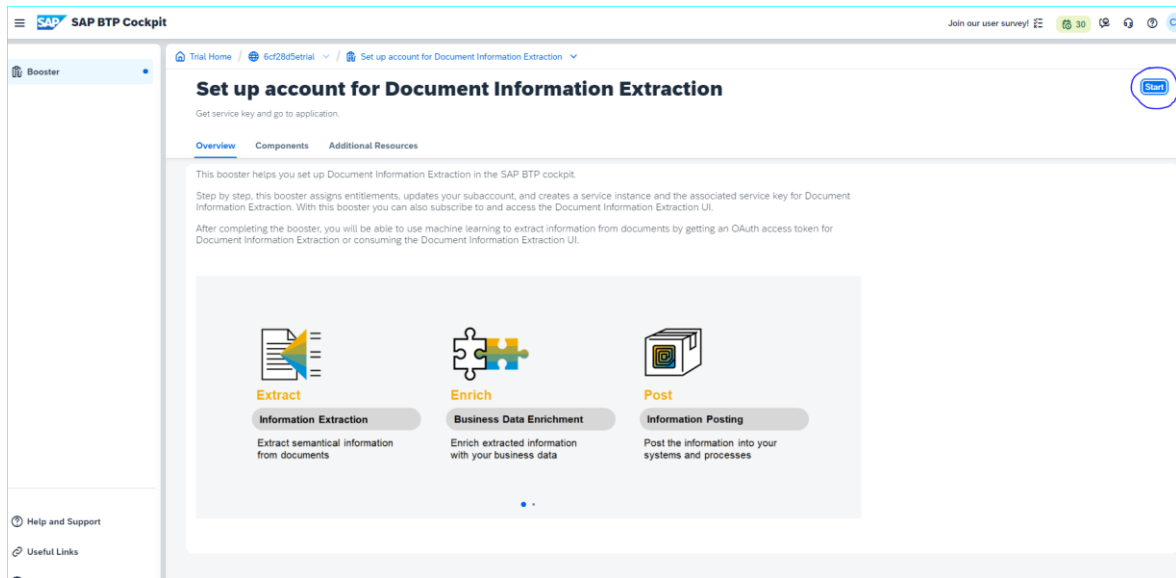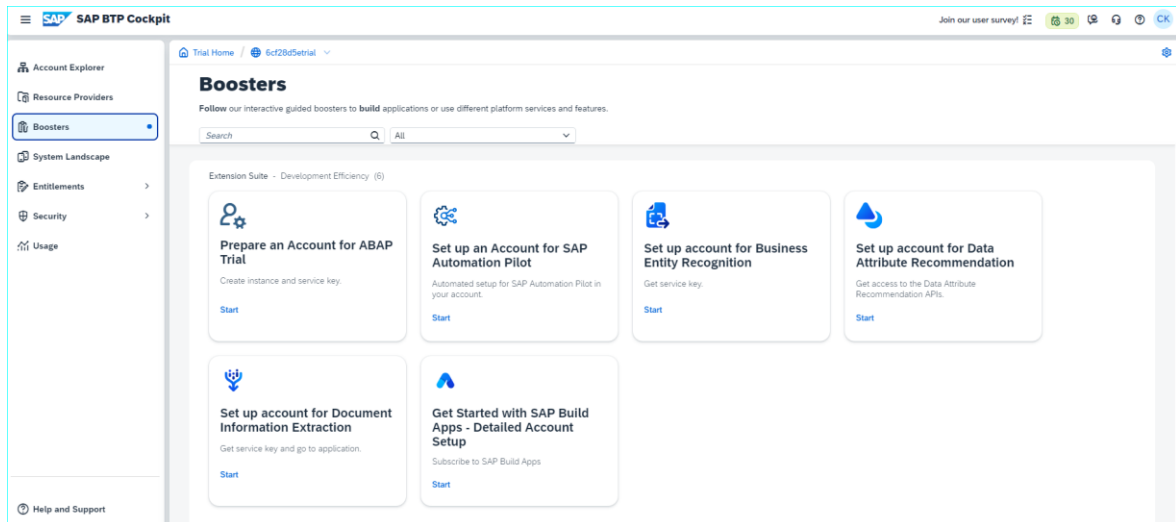
After the booster is executed, use Go to Application to navigate to the Document Information Extraction application-
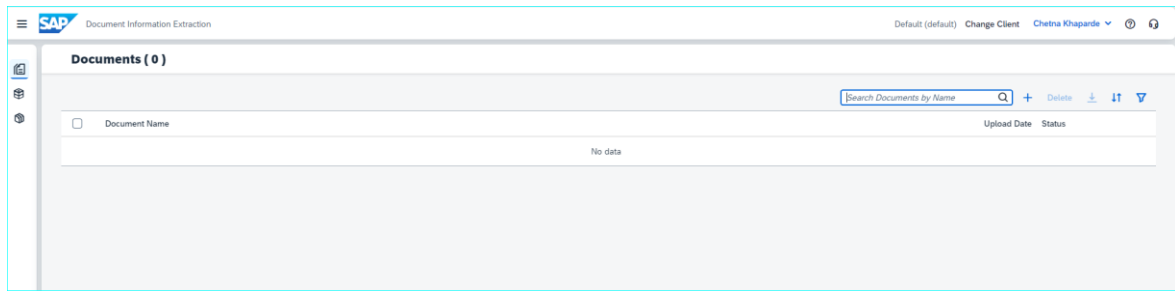
In the left navigation pane, click Schema Configuration.
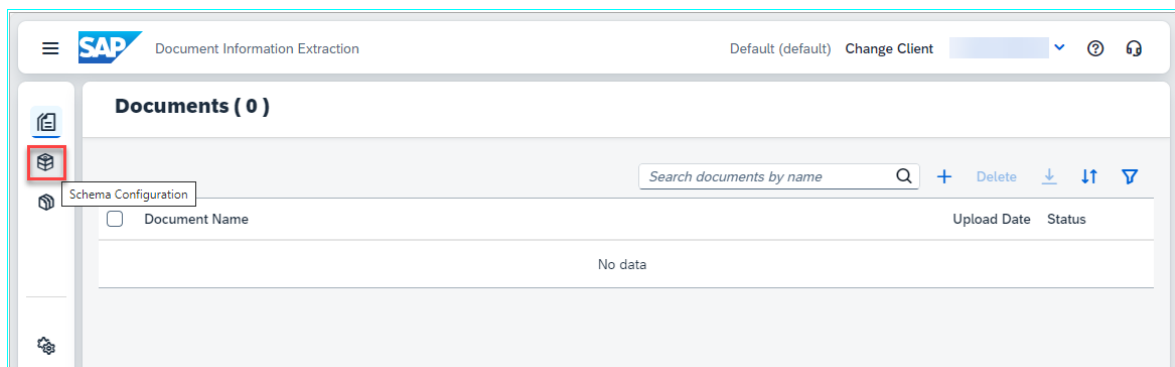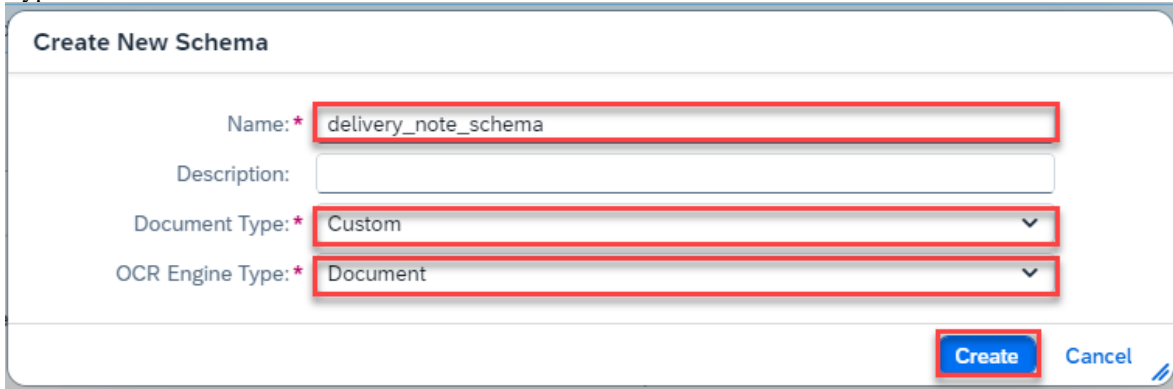
To create your own schema, click Create. A dialog opens-
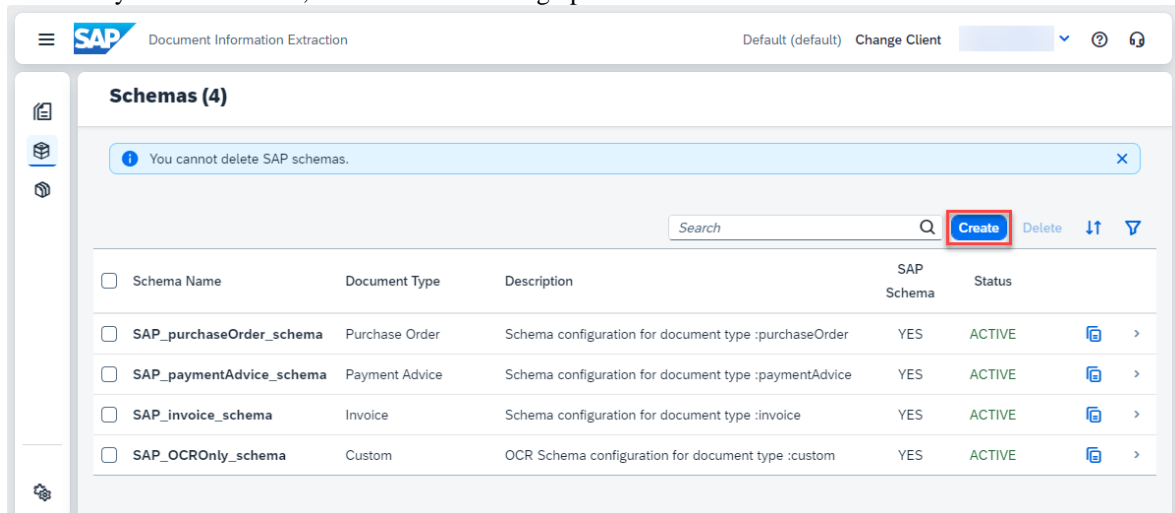
In the dialog, enter a name for your custom schema – for example, delivery_note_schema. Note that the name can't include blanks. Next, select Custom as your Document Type and Document as the OCR Engine

Type.Click Create to create the schema.
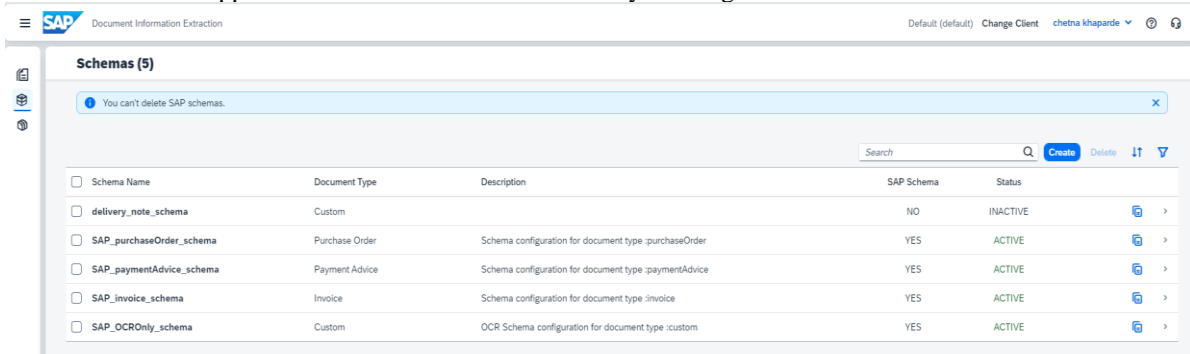
Your schema now appears in the list. Access the schema by clicking on it.
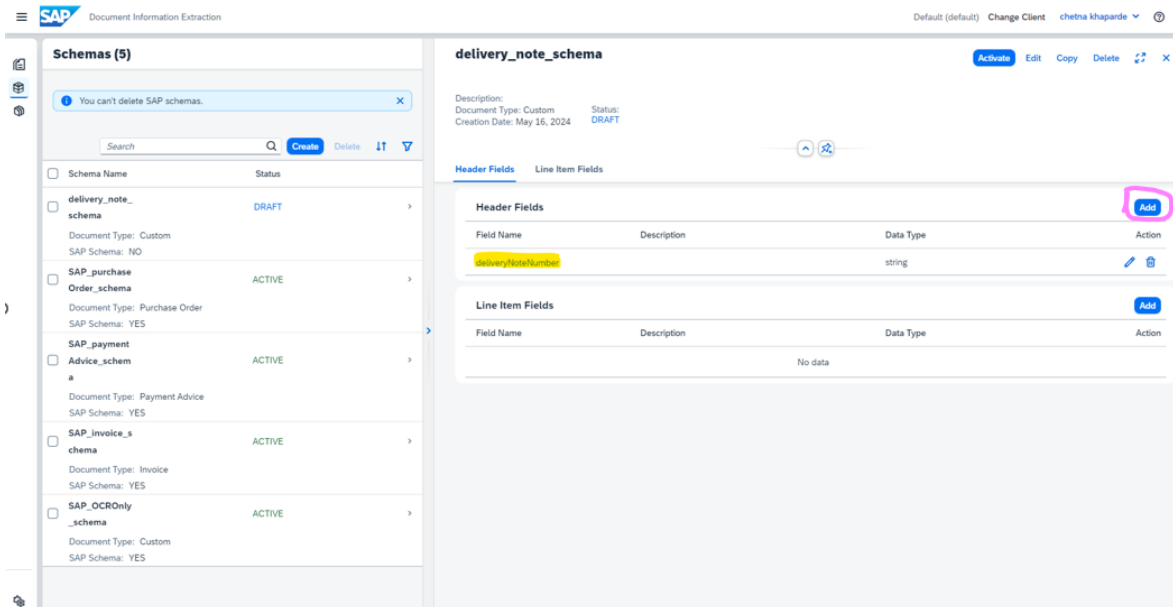
Click on "Add" to add fields for extraction-

You must enter a field name and data type for each custom field. The available data types are string, number, date, discount, currency, and country/region. Default extractors aren't available for custom documents. You can also optionally add a field label (user-friendly name) and a description.A description is an optional entry. It can be useful if you want to include an explanation or some additional context for a field.You can also use a description for other purposes, such as categorizing fields. For example, in the description of the field limitedContract in work contracts, you could specify yes, if the

contract is limited and no if the contract is not. Or you could specify that the line-item field skill Type in a résumé can be technical or language.
See Step 5 for examples of schemas that use the description field.

As your first header field, add the number of the delivery note.
1. Enter the name for your field – for example, deliveryNoteNumber.
2. Select string as the Data Type.
3. Use auto as the Setup Type and click Add.

Note that when you use the setup type auto without a default extractor, LLMs are used to extract the information from the document. The setup type manual supports extraction using a template.

The field now appears in your list of header fields, where you can see all the information that you've just entered. You can edit or delete the field by clicking the respective icons on the right.

Click Add again to open the Add Data Field dialog.
1. Enter the name for your second header field – for example, purchaseOrderNumber.
2. Select string as the Data Type.
3. Use auto as the Setup Type and click Add.

Add Data Field

Field Name: * purchaseOrderNumber

Field Label:
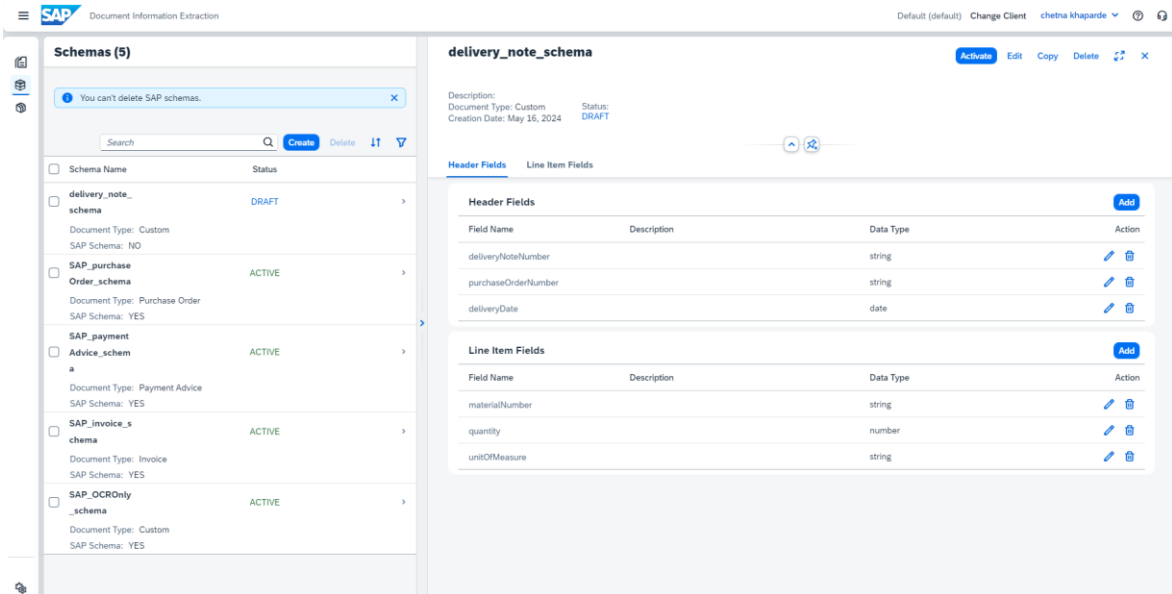
Description:

Data Type: * string ⌄

Setup Type: * auto ⌄

Add   Cancel

Now, go ahead and add the remaining header fields and line-item fields shown in the table and image below. Pay attention to the different data types and notice that the last three fields are line-item fields (not header fields). Feel free to extend or reduce the list of fields.

Field Type	Field Name	Data Type	Setup Type
header field	deliveryNoteNumber	string	auto
header field	purchaseOrderNumber	string	auto
header field	deliveryDate	date	auto
line item field	materialNumber	string	auto
line item field	quantity	number	auto
line item field	unitOfMeasure	string	auto

3. Results and Analysis (10pt)
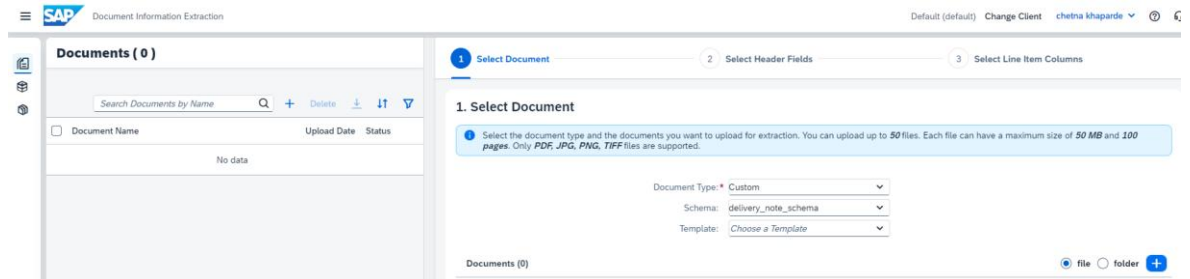
Access Document from the navigation on the left, then click '+' to upload the delivery note document.

Select Document Type – 'Custom' and created schema; click on upload.
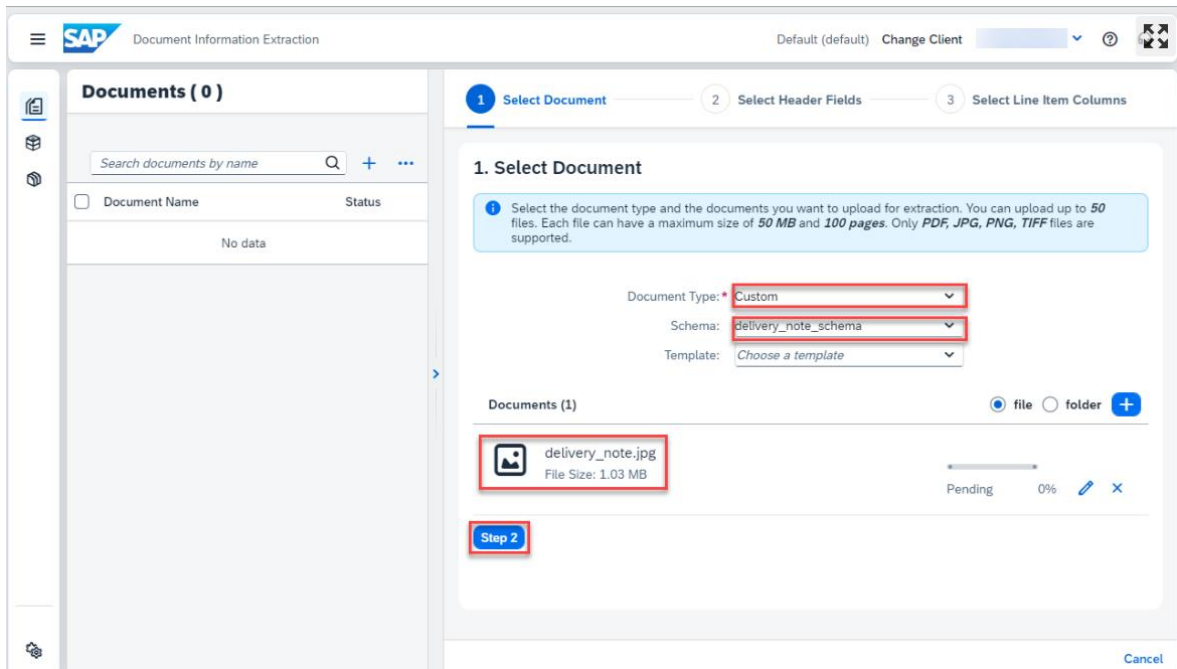
Click on Step 2.

Click on Step 3 and click Review. It is also possible to click on Edit to make any changes before confirming.
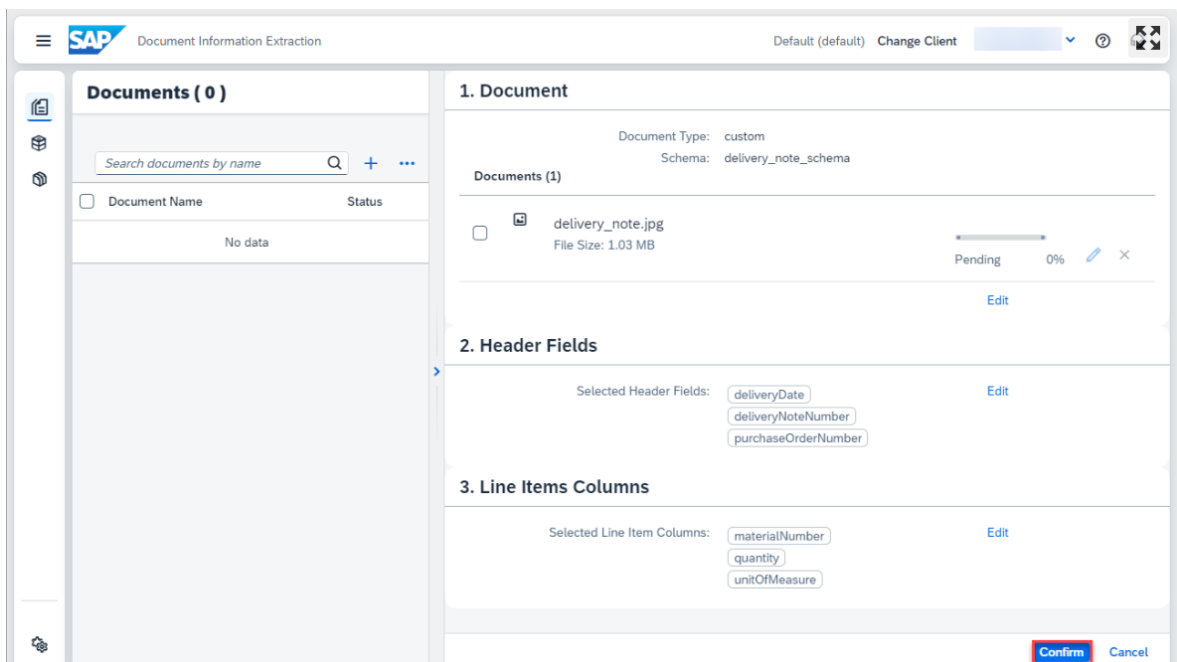
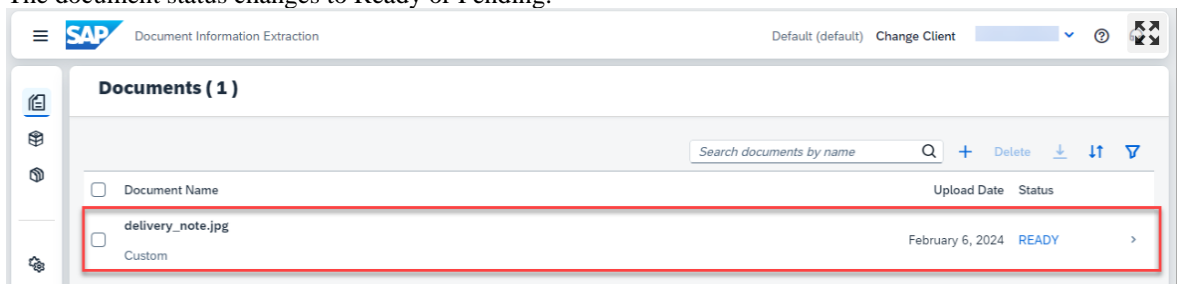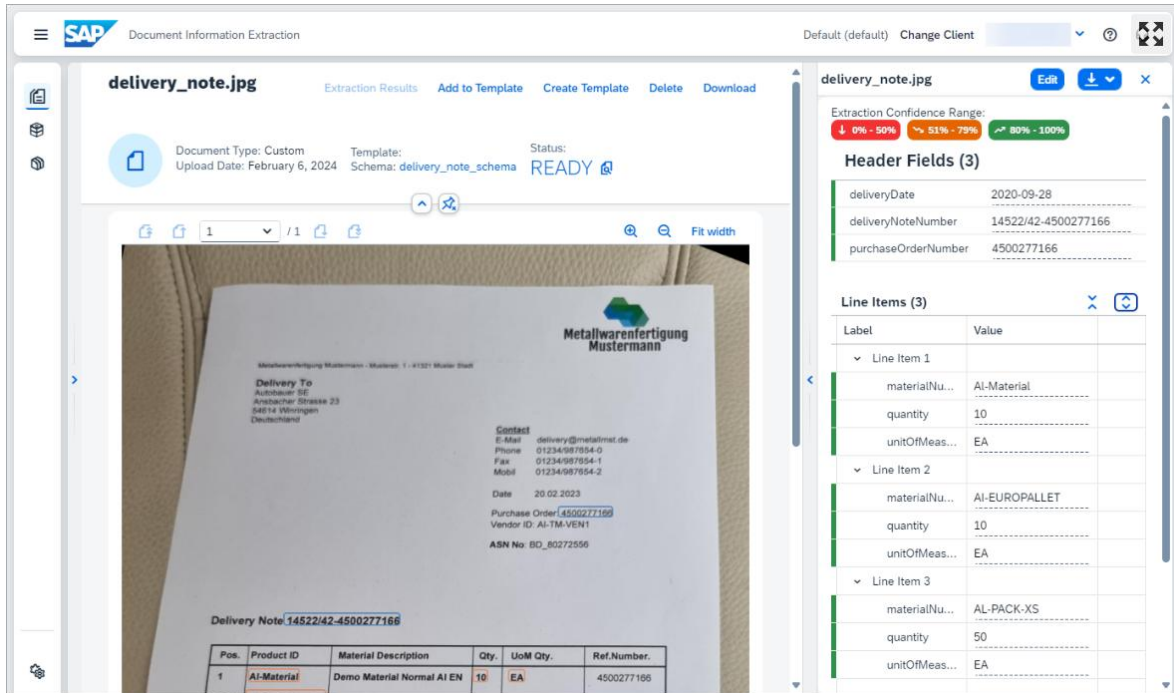The document status changes to Ready or Pending.

Access the document by clicking on it. Click Extraction Results to see the information extracted from the delivery note using LLMs and the schema created in the above steps.

The information is successfully extracted from a delivery note document using the schema configuration feature from Document Information Extraction and LLMs.

4. Conclusion

Automated document information extraction using SAP BTP works seamlessly to extract the required information from a pdf, tiff, jpeg, jpg, tif and tiff files. It supports over 20 languages across 27 countries. Along with information extraction, it provides data enrichment and multitenancy support which allows the service to run on a shared compute unit that can be used by multiple consumers (tenants). With that said, Document Information Extraction regularly undergoes audits and reviews of its policies and controls. It follows SAP's global data protection and privacy guidelines.

Document Information Extraction has immense potential to streamline and expedite the overall supply chain, specifically by reducing the processing times of inbound and outbound deliveries, shipment processing, freight verification, automatically processing payables, invoices, or payment notes while making sure that invoices and payables match. This is one of the strong use cases of Generative Artificial Intelligence.

References

[1] Use trial to extract information from custom documents with Generative AI and document information extraction | SAP tutorials. (n.d.-a). https://developers.sap.com/tutorials/cp-aibus-dox-ui-gen-ai.html

[2] Generative AI with SAP. (n.d.). https://www.sap.com/products/artificial-intelligence/generative-ai.html

[3] What is Generative Ai?. NVIDIA. (n.d.). https://www.nvidia.com/en-us/glossary/generative-ai/

[4] SAP help portal. (n.d.). https://help.sap.com/docs/intelligent-robotic-process-automation/cloud-studio-user-guide/document-information-extraction

[5] Cockpit.hanatrial.ondemand.com. (n.d.). https://cockpit.hanatrial.ondemand.com/trial/#/globalaccount/cc3b7fa0-5cab-4315-b002-efbd9cbaf9ce/boosters

[6] SAP help portal. (n.d.). https://help.sap.com/docs/sap-ai-core/sap-ai-core-service-guide/generative-ai-hub-in-sap-ai-core

[7] SAP discovery center. (n.d.). https://discovery-center.cloud.sap/protected/index.html#/mymissiondetail/98151/

[8] SAP. (2024, February 5). Feature Scope Description for Document Information Extraction. https://help.sap.com/doc/a08fc1abaa294506afafb3f7c890cf87/SHIP/en-US/Feature_Scope_Description_EN.pdf