Genpact AP Capture Overview

Last Modified on 05/02/2025 6:28 am EDT

Genpact AP Capture Overview

Genpact AP Capture is one of the solutions from the Genpact AP Suite, which integrates with AP Flow and other systems using the webAPI. It provides an enhanced capability to extract data from documents like invoices with minimum human intervention. The Genpact AP Capture extracts not only the header data, but also the line level data from documents efficiently and automates data extraction, classification and validation. Automating the extraction and validation reduces manual processing and error occurrence.

With Genpact AP Capture, the data capture process becomes faster, more accurate, and more secure, allowing the customer to focus on strategic decision-making.

Key features and benefits

Genpact AP Capture offers the following key features and benefits to customers:

- Improved extraction of accurate header and line level data from the documents.
- Document classification and intelligent queueing, to fast-track high priority invoices.
- Upto 40% optimized and expedited data extraction, based on the criticality, with minimum human intervention.
- An intuitive UI to validate and update incorrect/missing data, and capture human feedback for continuous learning.

Technical specifications

- APFlow Adaptor for integrating the Cora AP Flow with Genpact AP Capture.
- DocIntel for extraction of data and as a knowledge base.
- Cora AP Flow case management system for getting the documents in email, opening a case for the document, and sending the document to Genpact AP Capture for processing.

How Genpact AP Capture works

Genpact AP Capture at present supports extraction of data from invoices. However, we can build a custom model to extract data from other kinds of documents.

Following are the components of Genpact AP Capture, and the steps in the process of how Genpact AP Capture works:

Components	Description
AP Flow Adaptor	The system receives documents along with metadata from the AP Flow or the web API.
Saga	Based on the defined process, the system handles the lines in queue and redirects to the next step.
Gateway	The gateway takes the document and translates it to the system language, JSON. The extraction starts here.

Components	Description
Classification	The system sends the document and metadata to the DocIntel for classification. The DocIntel returns the extraction model to be used.
Extraction	With the chosen extraction model, the system extracts the data from the DocIntel and translates it to the system language i.e. JSON.
Enrichment	Based on the enrichment rules defined in a yaml file, the system adds new additional data, above the extracted data, from an external source. For example, if you want to know the vendor ID, then the system will take the vendor's name from the extraction results and look for the corresponding vendor ID value in the master data DB and add it.
Validation	Based on the validation rules defined in a yaml file, the system validates the extracted and enriched data. In case of failed validation, the system throws an error. For example, the system can validate and remove some specific symbols or numbers or letters, defined in the yaml, from the data.
Eyeball	Based on the eyeball rules defined in a yaml file, the system displays the extraction results in the UI where an end user validates these results side by side with the original document (image or PDF), and makes correction (add or remove data), if required, before the system processes the data. By default, the eyeball process is set to true. So, the eyeball configuration needs to define the parameters for false i.e. when the eyeball should *not* be performed.
XML generator	The system translates the extraction results (from JSON) to an XML file and shares the file with Cora AP Flow (in case of API) or through web API.