

Invoice Processing in Salesforce with Document AI

Read, split, match and write every invoice and its line items to Salesforce records, with review before anything is written

Why is invoice processing still manual in so many Salesforce orgs?

Invoices arrive by email, portal upload, scan and fax, and someone keys them in by hand. Manual data entry from paper invoices is slow work that scales with every new vendor, and re-keying line items is the part nobody wants. Template-based invoice OCR was supposed to fix it, but a template breaks the moment a vendor moves a field or a value is not where the template expects it.

CloudFiles is Document Automation for Salesforce, and Document AI turns invoice processing into a process that runs inside Salesforce. An invoice arrives, is read, is split if it holds several invoices, is matched to the right vendor and account, has its header fields and line items written to Salesforce records, and is held for review wherever your team wants a sign-off. Accounts payable documents stop being a keying job and become a Salesforce process you can see and change.

How do invoices get into the process?

Invoices enter from wherever they already land:

- **Email.** Email-to-Case attachments are a common entry point, and a CloudFiles event for a Salesforce file attached starts processing.
- **Folder drop.** With CloudFiles Document Management, a file dropped into a connected folder in AWS S3, SharePoint, OneDrive, Google Drive, Azure Blob, Box or Dropbox can trigger processing.
- **Upload on a record.** A file uploaded to a vendor, account or invoice record starts a record-triggered Flow.
- **Batch upload.** A screen Flow takes a stack of invoices at once.
- **Buttons.** Document Mapper runs from List View and Detail Page buttons.

Photographed invoices in JPEG or PNG work with Process Document using AI and Document Mapper. Doc AI Workflows take PDFs and Office documents.

How does invoice OCR read a scanned invoice?

OCR is the reading step. Scanned and mixed pages are sent to OCR, and native PDF text is read directly, so a clean PDF from a billing system and a crooked fax both go through the same process. A page-range setting limits a run to the pages that matter, such as skipping a cover sheet.

Large language models then read the fields you described in plain language: "the invoice number", "the total amount due", "the service period", "each line item with its description, quantity and amount". There are no layout templates to draw or maintain, so a new vendor or a redesigned invoice needs no new setup.

Can Document AI split a PDF that holds many invoices?

Yes. CloudFiles: Split Document using AI splits a multi-invoice PDF into one file per invoice, names each output file from its content, and saves the pieces as Salesforce Files. Each piece then runs through invoice processing on its own.

A pattern we see with telecom and utility vendor bills: accounts and vendors get AWS S3 folders automatically through Document Management, an invoice dropped in a vendor's Invoices folder triggers processing, and a long master invoice is split by account number into one invoice per account. Header fields and line items are then written to Invoice and Invoice Line Item records, so nobody re-keys line items.

How do header fields and line items reach Salesforce records?

CloudFiles: Document Mapper runs a saved configuration that reads the invoice and creates or updates an Invoice record plus its Invoice Line Item child records in one run.

Invoice header fields	Line item fields
Invoice number, invoice date, due date	Description
Vendor and account number	Quantity and unit price
Service period	Line amount
Subtotal, tax and total due	Account or service reference

Writes go through standard Salesforce DML, so validation rules and field-level security apply. The per-field option "Do not update if a value is already present" protects values your team has already confirmed, such as a corrected vendor code.

For invoice processing with more logic, a Doc AI Workflow extracts typed fields and lists of items, loops through lines with For each, and writes them with "Create Salesforce records (bulk)". A blank value never wipes a field, so a missing value on one invoice does not erase what is already on the record.

How are vendors and accounts matched?

In Document Mapper, lookups are resolved per field by Get Existing, Create New or Call Flow Action. Match the vendor by name or tax ID, match the account by account number, and create a vendor record when a new one appears. A lookup with no match or several matches always goes to review, so a missing or ambiguous vendor or account always reaches a person.

In workflows, find and search steps match on a field value, by fuzzy match, or by an AI-picked match, and a Decision step routes invoices that match nothing.

Can someone review an invoice before it is written?

Yes, and you choose where. In Document Mapper, "When automation runs finish" is set to "Auto approve" or to the review queue setting. Held runs show Pending Review, and a reviewer can edit cells, add a line that was missed, then choose "Approve & Write to Salesforce" or Reject.

In workflows, the review policy is "Straight through", a hold on the first write, or approval on every write, and a single step can be set to "Require approval for this write", for example only the step that creates the payable. Runs waiting on a person show "Awaiting review".

What about vendor packets and other accounts payable documents?

Document AI supports a vendor packet pattern end to end. A scanned vendor packet is read, the voucher or invoice reference is matched to the check-run export, and the packet is renamed and filed to the vendor's folder, or to an Unmatched folder when there is no match. Filing uses Document Management's Move Resource action, so the right document ends up in the right SharePoint or S3 folder with a name your team can search.

Classification sorts mixed accounts payable documents before any of this: a Classify step in a workflow, a Query prompt in Flow, or a Document Type field in Mapper tells invoices, credit memos and statements apart, and a Decision step routes each type.

How is each invoice run recorded?

Each run is recorded in Salesforce as a CloudFiles DocAI Run record with its status: Processing, Completed, Pending Review, Failed or Rejected. Reprocessing creates a new run and keeps the old one. Workflows keep a step-by-step history of every run, and every pause, approval and rejection is recorded.

What does Document AI replace?

Teams move to Document AI when replacing template-based OCR tools, a MuleSoft IDP project or a custom build. Everything runs as Salesforce Flow actions from one managed package, and Document Generation can create documents from the extracted data in the same process, such as a remittance summary.

Get started

Install CloudFiles Document AI from AgentExchange and start a free 14-day trial or a Test Drive, or book a demo at cloudfiles.io.

CloudFiles, Document Automation for Salesforce · cloudfiles.io · Install CloudFiles Document AI from AgentExchange