

Classify and Split PDF Documents in Salesforce

Split PDF bundles into one file per document, classify each one and route it, inside Salesforce

How do you split PDF bundles and classify documents in Salesforce?

CloudFiles is Document Automation for Salesforce, and CloudFiles Document AI splits, classifies and routes documents inside your Salesforce org. It takes the packets and bundles your team receives, turns them into one file per document, works out what each document is, and sends each one to the right record, the right folder and the right next step.

The manual version is familiar. A case packet, a master invoice or a scanned vendor packet arrives as one long PDF. Someone opens it, finds where each document starts, splits the PDF by hand, renames every piece, decides what each one is and drags it to the right place. Document AI replaces all of that with one process that runs in Salesforce Flow.

How does Split Document using AI split a PDF?

CloudFiles: Split Document using AI takes a multi-document PDF bundle and splits it into one file per document. It names each output file from its content and saves the pieces as Salesforce Files. The split comes from what Document AI reads in the bundle.

When the PDF split is done, a document-split event comes back, and your Flow reads it with CloudFiles: Get Event Details. From there the Flow can take each piece to the next step: classify it, extract its fields, attach it to a record or hold it for review.

When does a team need a PDF split?

- **Case packets.** For example, a financial services team can split multi-document case packets into one file per document, then classify each one, rename it from its content and attach it to the right record.
- **Master invoices.** A telecom or utility vendor bill can arrive as one long master invoice. It can be split by account number into one invoice per account, and each invoice's header fields and line items written to Invoice and Invoice Line Item records.
- **Vendor packets.** Accounts payable teams receive scanned vendor packets that bundle several documents together.
- **Application packets.** Intake teams receive applications with supporting documents in the same file.

How does document classification work in Document AI?

Document classification tells the process what each document is, so the right steps run on it. Document AI lets you classify documents in three places:

Where you classify	How it works	Where it fits
Doc AI Workflow	A Classify step identifies the document type on the visual canvas	Processes with branches, matching and several writes
Salesforce Flow	CloudFiles: Query Document/Dataset asks the document what it is and returns a structured answer	Flows you already run, where the type drives the next action
Document Mapper	A Document Type field is read with the other fields	One document becoming one record, with its type stored on it

In every case you describe your document types in plain language, such as quarterly report, valuation, audit, board pack, budget or investment committee memo. There are no layout templates and no training sets to build.

How do you classify and organize file types automatically?

To classify and organize file types automatically, pair classification with routing. In a Doc AI Workflow, a Decision step routes by document type, so each type follows its own branch: an invoice goes to invoice processing, a certificate goes to a supplier record, and a document that fits no type goes to a person.

Routing can also pick which mapper runs. A pattern we see in food and consumer goods manufacturing: third-party audit reports and certificates are attached to a supplier location record, singly or through a batch-upload screen Flow, and a filename rule picks the right mapper. Audit date, auditor, scheme, type, score and finding counts are written back to the record, and results that need a check go to the review queue.

Organizing also means putting the file where it belongs. Documents live in Salesforce Files, and with CloudFiles Document Management also in SharePoint, OneDrive, Google Drive, AWS S3, Azure Blob, Box and Dropbox. A file dropped into a connected folder can trigger processing, and classified documents can be moved into the right folder with Document Management's Move Resource action.

For example, a private equity firm can give portfolio companies a SharePoint "Drop Off" folder. Each file dropped there is classified against the firm's document types and matched to the portfolio company. After a confirmation reply it is moved into that company's subfolder, and a custom object logs every route.

How are split and classified documents renamed?

Split Document using AI names each output file from its content, so the pieces arrive with names that say what they are. The same pattern extends to whole packets: for example, an accounts payable team

can have a scanned vendor packet read, match the voucher or invoice reference to the check-run export, and rename and file the packet to the vendor's folder, or to an Unmatched folder when there is no match.

What does document parsing look like after classification?

Once the type is known, document parsing can run the steps that type needs. A Doc AI Workflow uses Extract fields for typed fields and lists of items, Ask the document for plain-language questions, and find or search records to match what was extracted against Salesforce. A check can be one Ask the document step with a Yes or No answer, followed by a Decision step, for example a consent form checked against a funding request, with failed checks listed for a person.

Document Mapper parses each type into a Salesforce record plus child records, with lookups resolved per field by Get Existing, Create New or Call Flow Action. Writes go through standard Salesforce DML, so validation rules and field-level security apply. In Mapper, Test Run History shows past test runs, so your admin can see how a configuration handled each sample document before it runs on live intake.

Can someone check a classification before anything is written?

Yes. In Document Mapper, the review queue setting holds finished runs as Pending Review, and a reviewer can edit cells, add rows, then choose "Approve & Write to Salesforce" or Reject. A lookup with no match, or with several matches, forces review. In a Doc AI Workflow, the review policy can run straight through, place a hold on the first write, or ask for approval of each write, and held runs show "Awaiting review".

Each run is recorded in Salesforce as a CloudFiles DocAI Run record, with a status of Processing, Completed, Pending Review, Failed or Rejected. Workflows keep a step-by-step history of every run, and every pause, approval and rejection is recorded.

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. A free sandbox proof of concept on your own bundles and packets is also offered.