How to Convert PDF Files Into Structured Data

Need to move data from PDFs to a spreadsheet, or another system? Save this free checklist or share it with your IT consultant.



4 ways to extract data from PDFs

- [ ] Manually rekey or copy and paste

  Yes, this an option, but it certainly is not the fastest or most effective approach unless you only have a couple of PDF documents.

- [ ] Try a free tool like Tabula

  Tabula will return a spreadsheet file which you probably need to post-process manually. It's definitely a good starting point if you deal with native PDF files (not scans).

- [ ] Outsource manual data entry

  There are literally thousands of data entry providers out there you can hire. However it comes with a lot of overhead, and it takes a lot of time/effort to find the right provider, agree on terms, and explain your specific use-case.

- [ ] Use a fully automated PDF data extraction software

  Automated PDF data extraction solutions come in different flavours, ranging from simple OCR tools to enterprise ready document processing and workflow automation platforms.

How to use Docparser to automatically convert PDF documents into structured data

- [ ] Sign up for a trial at Docparser.com — it's free, and there's no credit card required
- [ ] Start creating your first document parser by selecting a category that's the closest match to the document you want to extract data from

  Pick from Invoices, Purchase Order, Application Forms, and more. If you can't find a close match, choose the Miscellaneous option.

- [ ] Import your business documents

  You can upload documents by dragging and dropping them from a folder on your computer into Docparser, emailing them, or connecting to a document storage drive like Dropbox, Box, or Google Drive.

- [ ] Train our PDF and Document Parser for the type of document you want to process

  Typically you set up a rule for each field you want to extract data from. For example - on a PDF form with Name and email address fields you want to extract, you'd need to set up two rules.

- [ ] Set optional filtering and formatting criteria you'd like for fields like dates, dollar amounts, etc.
- [ ] Export your parsed data to an Excel spreadsheet, Google Sheets
- [ ] Work with your IT department or an IT consultant to have your parsed data populate business systems like an ERM or CRM

Make and Share Free Checklists

checkli.com