


Apache PDFBox - Java PDF Library

The Apache PDFBox™ library is an open source Java tool for working with PDF documents. This project allows creation of new PDF documents, manipulation of existing documents and the ability to extract content from documents. Apache PDFBox also includes several command line utilities. Apache PDFBox is published under the [Apache License v2.0](#) .

The project is a volunteer effort and we're always looking for interested people to help us improve PDFBox. There are a multitude of ways that you can help us depending on your skills. Subscribe to the [developer mailing list](#) and find out how you can help.

To get help on using PDFBox, please subscribe to the [users mailing list](#) and post your questions there. We're happy to help.

Features

- PDF to text extraction
- Merge PDF Documents
- PDF Document Encryption/Decryption
- Lucene Search Engine Integration
- Fill in form data FDF and XFDF
- Create a PDF from a text file
- Create images from PDF pages
- Print a PDF
- PDF/A validation