

Apache PDFBox - Java PDF Library

Java PDF Library, pdfbox, pdftotext, print pdf

Table of contents

1 Introduction.....	2
2 Features.....	2

1 Introduction

Apache PDFBox is an open source Java PDF library 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](#).

Disclaimer

Apache PDFBox is an effort undergoing incubation at The Apache Software Foundation (ASF), sponsored by the Apache Incubator PMC. Incubation is required of all newly accepted projects until a further review indicates that the infrastructure, communications, and decision making process have stabilized in a manner consistent with other successful ASF projects. While incubation status is not necessarily a reflection of the completeness or stability of the code, it does indicate that the project has yet to be fully endorsed by the ASF.

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.

2 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