

APACHE PDFBOX

[Overview \(/index.html\)](/index.html)

[Downloads \(/download.cgi\)](/download.cgi)

COMMUNITY

[Support \(/support.html\)](/support.html)

[Mailing Lists \(/mailinglists.html\)](/mailinglists.html)

[Project Team \(/team.html\)](/team.html)

DOCUMENTATION**▼ 2.0**

[Migration Guide \(/2.0/migration.html\)](/2.0/migration.html)

[Getting Started \(/2.0/getting-started.html\)](/2.0/getting-started.html)

[Examples \(/2.0/examples.html\)](/2.0/examples.html)

[Dependencies \(/2.0/dependencies.html\)](/2.0/dependencies.html)

▶ Cookbook

[Command Line Tools
\(/2.0/commandline.html\)](/2.0/commandline.html)

[API Docs \(/docs/2.0.2/javadocs/\)](/docs/2.0.2/javadocs/)

▶ 1.8**DEVELOPMENT**

[Building from Source \(/building.html\)](/building.html)

[Coding Conventions \(/codingconventions.html\)](/codingconventions.html)

[Update the Website \(/siteupdate.html\)](/siteupdate.html)

Jenkins

[\(https://builds.apache.org/view/All/job/PDFBox-trunk/\)](https://builds.apache.org/view/All/job/PDFBox-trunk/)

Travis CI (<https://travis-ci.org/apache/pdfbox>)

SonarQube

<https://analysis.apache.org/dashboard/index/org.apache.pdfbox:pdfbox-reactor>

JIRA

<https://issues.apache.org/jira/browse/PDFBOX>

[Ideas \(/ideas.html\)](/ideas.html)

[External Links \(/references.html\)](/references.html)

APACHE SOFTWARE FOUNDATION

Apache Software Foundation
(<http://www.apache.org/>)

ASF Sponsors
(<http://www.apache.org/foundation/thanks.html>)

Security (<http://www.apache.org/security/>)

Apache PDFBox[®] - A 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.

Apache PDFBox 2.0.2 released (2016-06-09)

The Apache PDFBox community is pleased to announce the release of Apache PDFBox version 2.0.2. It is available for download at:

<http://pdfbox.apache.org/download.cgi> (<http://pdfbox.apache.org/download.cgi>)

See the full release notes (<https://issues.apache.org/jira/secure/ReleaseNote.jspa?projectId=12310760&version=12335585>) for details about this release.

Getting Help

To get help on using PDFBox, please Subscribe to the Users Mailing List (<mailto:users-subscribe@pdfbox.apache.org>) and post your questions there. We're happy to help.

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 Mailing Lists (</mailinglists.html>) and find out how you can help.

Features



Extract Text

Extract Unicode text from PDF files.



Split & Merge

Split a single PDF into many files or merge multiple PDF files.



Fill Forms

Extract data from PDF forms or fill a PDF form.



Preflight

Validate PDF files against the PDF/A-1b standard.



Print

Print a PDF file using the standard Java printing API.



Save as Image

Save PDFs as image files, such as PNG or JPEG.



Create PDFs

Create a PDF from scratch, with embedded fonts and images.



Signing

Digitally sign PDF files.

News

CVE-2016-2175 XML External Entity vulnerability (2016-05-27)

Due to a XML External Entity vulnerability we strongly recommend to update to the most recent version of Apache PDFBox.

Versions Affected: Apache PDFBox 1.8.0 to 1.8.11 and 2.0.0. Earlier, unsupported versions may be affected as well.

Mitigation: Upgrade to Apache PDFBox 1.8.12 respectively 2.0.1

Apache PDFBox 1.8.12 and 2.0.1 released (2016-04-26)

The Apache PDFBox community is pleased to announce the release of Apache PDFBox version 1.8.12 and 2.0.1. They are available for download at:

<http://pdfbox.apache.org/download.cgi> (<http://pdfbox.apache.org/download.cgi>)

See the full release notes 1.8.12 (<https://issues.apache.org/jira/secure/ReleaseNote.jspa?projectId=12310760&version=12334652>) and 2.0.1 (<https://issues.apache.org/jira/secure/ReleaseNote.jspa?projectId=12310760&version=12335342>) for details about this release.

Apache PDFBox 2.0.0 released (2016-03-18)

After more than 3 years of development the Apache PDFBox community is pleased to announce the release of Apache PDFBox version 2.0.0. It is available for download at:

<http://pdfbox.apache.org/download.cgi> (<http://pdfbox.apache.org/download.cgi>)

The Migration Guide (<http://pdfbox.apache.org/2.0/migration.html>) shall give users coming from PDFBox 1.8 or earlier an overview about things to look at when switching over. More details to come.

See the full release notes (<https://issues.apache.org/jira/secure/ReleaseNote.jspa?projectId=12310760&version=12319281>) for details about this release.

Apache PDFBox 1.8.11 released (2016-01-18)

The Apache PDFBox community is pleased to announce the release of Apache PDFBox version 1.8.11.

The release is available for download at: <http://pdfbox.apache.org/download.cgi>
(<http://pdfbox.apache.org/download.cgi>)

See the full release notes (<https://issues.apache.org/jira/secure/ReleaseNote.jspa?projectId=12310760&version=12333061>) for details about this release.

Apache PDFBox 2.0.0 RC3 released (2016-01-15)

The Apache PDFBox community is pleased to announce the release of Apache PDFBox version 2.0.0 RC3. The release is available for download at:

<http://pdfbox.apache.org/download.cgi> (<http://pdfbox.apache.org/download.cgi>)

The numerous feedback on our second release candidate helps us to make this release candidate better again, e. g. optimized font cache, improved text extraction. A lot of bug fixes are included as well. We'd like to thank everybody who helps us to get a step forward. Please have a look at the new release candidate as well, so that the next release hopefully could be the final one.

See the full release notes (<https://issues.apache.org/jira/secure/ReleaseNote.jspa?projectId=12310760&version=12319281>) for details about this release.

Copyright © 2009–2016 The Apache Software Foundation (<http://www.apache.org/>), Licensed under the Apache License, Version 2.0 (<http://www.apache.org/licenses/LICENSE-2.0>).

Apache PDFBox, PDFBox, Apache, the Apache feather logo and the Apache PDFBox project logos are trademarks of The Apache Software Foundation.