

PDFBox - References

Table of contents

1 PDFBox References.....	2
2 Projects.....	2
3 Articles/Books.....	3

1. PDFBox References

This page lists projects that utilize PDFBox and articles that have been written about PDFBox. Send [me](mailto:ben@benlitchfield.com) (mailto:ben@benlitchfield.com) an e-mail if your article or project is missing.

2. Projects

Project Name	License	Project Description
Centric CRM (http://www.centriccrm.com)	Free To Use But Restricted/Commercial	The Most Advanced Open Source CRM Software.
Canoo Webtest (http://webtest.canoo.com/webtest/)	BSD Like (http://www.bsdlicense.com/)	Free OpenSource tool for Fast and Accurate testing of Java-based Web applications.
Jomic (http://jomic.sourceforge.net/)	GPL	Jomic is a viewer for comic book archives.
LIUS (http://www.bibl.ulaval.ca/lius/index.en.html)	GPL	LIUS is an indexing Java framework based on the Jakarta Lucene project. The LIUS framework adds to Lucene many files format indexing functionalities as: Ms World, Ms Excel, Ms PowerPoint, RTF, PDF, XML, HTML, TXT, Open Office suite and JavaBeans.
LuceGene (http://www.gmod.org/lucegene/)	Artistic License	LuceGene is an open-source document/object search and retrieval system specially tuned for bioinformatics text databases and documents.
Nutch (http://incubator.apache.org/nutch/)	ASL	Nutch is open source web-search software. It builds on Lucene, adding web-specifics, such as a crawler, a link-graph database, parsers for HTML and other document formats, etc.
OpenCms	Custom	OpenCms is a professional

PDFBox - References

(http://www.opencms.org/)	(http://www.opencms.org/opencms-3.0.0-RC1/)	Open Source Website Content Management System.
PodReader (http://homepage.mac.com/ptwobrussell/podreader.html)	GPL	PodReader is an application that facilitates making electronic documents like eBooks readable on your iPod.
SearchBlox (http://www.searchblox.com)	Commercial	SearchBlox is a high-performance corporate search software designed for the Java 2 Enterprise Edition (J2EE) platform.
Terrier (http://ir.dcs.gla.ac.uk/terrier/)	MPL	Terrier is software for the rapid development of Web, intranet and desktop search engines.

3. Articles/Books

Article Name	Article Abstract
Build an eDoc Reader for your iPod Part 1 (http://www.oreillynet.com/pub/a/mac/2004/12/14/pod_reader.html) - User Interface Part 2 (http://www.oreillynet.com/pub/a/mac/2004/12/17/pod_reader.html) - Document Reading Engine Part 3 (http://www.oreillynet.com/pub/a/mac/2005/01/07/pod_reader.html) - *Integration with PDFBox*	A three part article that discusses the implementation of the PodReader application. PodReader is a Cocoa application written in Objective-C and article discusses how to use the Cocoa-Java bridge to integrate with the Java port of PDFBox.
Lucene In Action (http://www.lucenebook.com/)	A book that discusses integrating with the lucene search engine. One chapter discusses how to index various file formats and highlights PDFBox for indexing PDF documents.
Java Developers Journal - March 2005 (http://www.sys-con.com/story/?storyid=48543&DE)	An article written by the lead developer of PDFBox discussing text extraction and AcroForm integration using PDFBox functionality.