

Apache Lucene - Contributions

Peter Carlson

Table of contents

1 Overview.....	2
2 Lucene Tools.....	2
2.1 Luke.....	2
2.2 LIMO (Lucene Index Monitor).....	2
3 Lucene Document Converters.....	2
3.1 XML Document #1.....	2
3.2 XML Document #2.....	2
3.3 PDF Box.....	3
3.4 XPDF - PDF Document Conversion.....	3
3.5 PDFTextStream -- PDF text and metadata extraction.....	3
3.6 PJ Classic & PJ Professional - PDF Document Conversion.....	3
4 Miscellaneous.....	3
4.1 Arabic Analyzer for Java.....	3
4.2 Phonetix.....	3
4.3 ejIndex - JBoss MBean for Lucene.....	3
4.4 JavaCC.....	4
4.5 LuSQL - Index databases with Lucene.....	4

1. Overview

This page lists external Lucene resources. If you have written something that should be included, please post all relevant information to one of the mailing lists. Nothing listed here is directly supported by the Lucene developers, so if you encounter any problems with any of this software, please use the author's contact information to get help.

If you are looking for information on contributing patches or other improvements to Lucene, see How To Contribute on the Lucene Wiki.

2. Lucene Tools

Software that works with Lucene indices.

2.1. Luke

URL	http://www.getopt.org/luke/
author	Andrzej Bialecki

2.2. LIMO (Lucene Index Monitor)

URL	http://limo.sf.net/
author	Julien Nioche

3. Lucene Document Converters

Lucene requires information you want to index to be converted into a Document class. Here are contributions for various solutions that convert different content types to Lucene's Document classes.

3.1. XML Document #1

URL	http://marc.theaimsgroup.com/?l=luce-dev&m=1007233335062
author	Philip Ogren - ogren@mayo.edu

3.2. XML Document #2

URL	http://www.mail-archive.com/lucene-user@jakarta.apache.org/msg
author	Peter Carlson - carlson@bookandhammer.com

3.3. PDF Box

URL	http://www.pdfbox.org/
author	Ben Litchfield - ben@csh.rit.edu

3.4. XPDF - PDF Document Conversion

URL	http://www.foolabs.com/xpdf
author	N/A

3.5. PDFTextStream -- PDF text and metadata extraction

URL	http://snowtide.com
author	N/A

3.6. PJ Classic & PJ Professional - PDF Document Conversion

URL	http://www.etymon.com/
author	N/A

4. Miscellaneous

4.1. Arabic Analyzer for Java

URL	http://savannah.nongnu.org/projects/aramorph
author	Pierrick Brihaye

4.2. Phonetix

URL	http://www.companywebstore.de/tangentum/mirror/en/products/ph
author	tangentum technologies

4.3. ejIndex - JBoss MBean for Lucene

URL	http://ejindex.sourceforge.net/
author	Andy Scholz

4.4. JavaCC

URL	https://javacc.dev.java.net/
author	Sun Microsystems (java.net)

4.5. LuSQL - Index databases with Lucene

URL	http://lab.cisti-icist.nrc-cnrc.gc.ca/cistilabswiki/index.php/LuSql
author	Glen Newton