

# History of Changes

## Table of contents

1	Introduction and explanation of symbols.....	3
2	Version 0.7.4-dev (Planned release:unknown).....	3
2.1	Changes to the Code Base.....	3
3	Version 0.7.3 (10/12/2006).....	4
3.1	Changes to the Code Base.....	4
3.2	Changes to Documentation.....	6
4	Version 0.7.2 (09/11/2005).....	6
4.1	Changes to the Code Base.....	6
5	Version 0.7.1 (04/10/2005).....	7
5.1	Changes to the Code Base.....	7
6	Version 0.7.0 (1/22/2005).....	9
6.1	Changes to the Code Base.....	9
7	Version 0.6.7 (10/09/2004).....	10
7.1	Changes to the Code Base.....	10
8	Version 0.6.6 (07/20/2004).....	12
8.1	Changes to the Code Base.....	12
9	Version 0.6.5 (03/08/2004).....	13
9.1	Changes to the Code Base.....	13
10	Version 0.6.4 (11/02/2003).....	14
10.1	Changes to the Code Base.....	14
11	Version 0.6.3 (09/13/2003).....	15
11.1	Changes to the Code Base.....	15
12	Version 0.6.2 (4/18/2003).....	16

12.1 Changes to the Code Base.....	16
13 Version 0.6.1 (3/9/2003).....	16
13.1 Changes to the Code Base.....	16
14 Version 0.6.0 (3/5/2003).....	16
14.1 Changes to the Code Base.....	16
15 Version 0.5.6 (11/28/2002).....	17
15.1 Changes to the Code Base.....	17
16 Version 0.5.5 (10/03/2002).....	17
16.1 Changes to the Code Base.....	17
17 Version 0.5.4 (09/17/2002).....	18
17.1 Changes to the Code Base.....	18
18 Version 0.5.3 (09/13/2002).....	18
18.1 Changes to the Code Base.....	18
19 Version 0.5.2 (09/06/2002).....	18
19.1 Changes to the Code Base.....	18
20 Version 0.5.1 (09/04/2002).....	18
20.1 Changes to the Code Base.....	18
21 Version 0.5.0 (08/31/2002).....	19
21.1 Changes to the Code Base.....	19
22 Version 0.4.1 (07/25/2002).....	19
22.1 Changes to the Code Base.....	19
23 Version 0.4.0 (07/23/2002).....	19
23.1 Changes to the Code Base.....	19
24 Version 0.3.0 (07/09/2002).....	19
24.1 Changes to the Code Base.....	19
25 Version 0.2.0 (06/03/2002).....	20
25.1 Changes to the Code Base.....	20
26 Version 0.1.0 (05/25/2002).....	20
26.1 Changes to the Code Base.....	20









## 1 Introduction and explanation of symbols

Changes are sorted by "false" and then chronologically with the most recent at the top. These symbols denote the various action types: =add, =fix, =remove, =update

## 2 Version 0.7.4-dev (Planned release:unknown)

















### 2.1 Changes to the Code Base





















- [ 1648374 ] Unexpected end of ZLIB input stream Committed by BJL.
- [ 1806912 ] Access to stream arrays Committed by BJL.
- [ 1793593 ] PDFStreamEngine doesn't handle rotation of -90. Committed by BJL.
- [ 1792084 ] spot colors' getTintTransform needs quick fix Committed by BJL.
- Added dependency on JAI Committed by BJL.
- [ 1666960 ] Patch for any kind of tiff based pdf. Committed by BJL.
- Fixed issue with "Could not read all parameters to decode image" Committed by BJL.
- [ 1726446 ] Fixing merge w&w/o acroform Committed by BJL.
- [ 1745133 ] OutOfMemoryError exception during text Committed by BJL.
- Upgraded to lucene 2.2.0 Committed by BJL.
- [ 1672991 ] Specify non default printer in silentPrint mode. Committed by BJL.
- [ 1607074 ] Portrait and Landscape PageFormat Fix Committed by BJL.
- \*\*\*API Change:Renamed PDPageContentStream.drawImage to drawXObject\*\*\* Committed by BJL.
- [ 1600672 ] Painting Form XObjects Committed by BJL.
- Added integration with JempBox library Committed by BJL.
- Upgraded to IKVM 0.32.0.0 Committed by BJL.
- [ 1674903 ] InvalidCastException in write2file for images Committed by BJL.
- [ 1057560 ] Overlay incorrectly orders overlay pages Committed by BJL.

-  Added ability to get/set document level JavaScript. Committed by BJJ.
-  Fixed issue with PNG encoding when using LZW/Flate filters Committed by BJJ.
-  [ 1562539 ] Implemented "B" operator for PageDrawer Committed by BJJ.
-  [ 1562539 ] PageDrawer now uses the correct color to display text Committed by BJJ.
-  Upgraded to junit4.3 Committed by BJJ.
-  Upgraded to Lucene-2.1.0 Committed by BJJ.

### 3 Version 0.7.3 (10/12/2006)

#### 3.1 Changes to the Code Base

-  Upgraded to Checkstyle 4.2 Committed by BJJ.
-  Upgraded to IKVM 0.30.0.0 Committed by BJJ.
-  [ 1546399 ] Use get/set functions for separators in PDFTextStripper Committed by BJJ.
-  PDDocument.silentPrint() to print without prompting for a printer Committed by BJJ.
-  [ 1544118 ] Bug in PDFont.getCodeFromArray Committed by BJJ.
-  [ 1529835 ] Add COSFloat.setValue() Committed by BJJ.
-  [ 1492555 ] PDChoiceField dead loop Committed by BJJ.
-  [ 1499521 ] NPE PDAppearance.convertToMultiLine Committed by BJJ.
-  [ 1522007 ] Error converting date Committed by BJJ.
-  Upgraded to Lucene 2.0.0 Committed by BJJ.
-  [ 1451164 ] Problems filling combo and radio form fields Committed by BJJ.
-  Upgraded to lucene 1.9.1 Committed by BJJ.
-  [ 1023133 ] Support PDF Functions Committed by BJJ.
-  Added command line org.pdfbox.PDFMerger Committed by BJJ.
-  **\*\*\*API Change\*\*\*** Promoted AppendDoc from example to util package, renamed to PDFMergerUtility. Committed by BJJ.
-  Upgraded to IKVM-0.24.0.1 Committed by BJJ.

-  [ 1391952 ] Problem extracting embedded attachments Committed by BJJL.
-  [ 1249607 ] Fixed issue with broken PDFs that contain multiple endobj Committed by BJJL.
-  [ 1153174 ] Added documentation for PDFHighlighter Committed by BJJL.
-  Removed log4j dependency Committed by BJJL.
-  [ 974661 ] getKids() Null Pointer Exception when parsing pdf Committed by BJJL.
-  Added better support for CJK encoding Committed by BJJL.
-  Changed signature of PDFPageContentStream.drawImage to take float arguments instead of int Committed by BJJL.
-  Fixed issue where form xobjects were not being drawn in the viewer Committed by BJJL.
-  Changed signature to PDDocumentCatalog.OpenAction to be an PDDestinationOrAction instead of just action. Committed by BJJL.
-  Added tolerance to text extraction sorting where text on a line was not at the same exact y coordinate but very close Committed by BJJL.
-  [ 1327133 ] Printing with form data Committed by BJJL.
-  Fixed issue with DateConverter that was trying to parse an empty string Committed by BJJL.
-  [ 1324846 ] appending text to PDPageContentStream messes up fonts Committed by BJJL.
-  Added new example ReplaceURLs to show how to replace a clickable URL in a PDF Committed by BJJL.
-  Implemented annotation drawing Committed by BJJL.
-  Implemented EndPath and StrokeAndClosePath operators Committed by BJJL.
-  Move text extraction permission checking from PDFTextStripper to ExtractText Committed by BJJL.
-  Added support for more annotations, thanks to a contribution from Paul King Committed by BJJL.
-  Created new FontBox project to hold all font library code Committed by BJJL.
-  Fixed issue where only the first page was sent to the printer Committed by BJJL.















-  Now automatically sets the page orientation when printing Committed by BJJ.

















### 3.2 Changes to Documentation

-  Upgraded to Apache Forrest 0.8-dev Committed by BJJ.

## 4 Version 0.7.2 (09/11/2005)


### 4.1 Changes to the Code Base






















-  Upgraded to IKVM-0.20.0.0 Committed by BJJ.
-  Added support to get annotations from a page and to create a RubberStamp annotation Committed by BJJ.
-  Added PDDocument.print() to send the PDF to a printer. Committed by BJJ.
-  [ 1276623 ] NullPointerException in PageDrawer:241 when extractin images Committed by BJJ.
-  Allow creation of PDJpeg from a BufferedImage, thanks to contribution from Paul King Committed by BJJ.
-  Removed PDTiff in favor of PDCcitt Committed by BJJ.
-  PDFBox no longer requires log4j!! Committed by BJJ.
-  New class to allow you to specify 'named' regions where text is to be extracted. Committed by BJJ.
-  [ 1261555 ] Unexpected end of ZLIB input stream when stream has a zero length Committed by BJJ.
-  [ 1226665 ] ImportXFDF giving NPE error Committed by BJJ.
-  renamed COSDictionary.setItem( String, boolean ) to COSDictionary.setBoolean( String, boolean ) Committed by BJJ.
-  Added sorting parameter to PDFTextStripper Committed by BJJ.
-  Fixed issues with PDF encryption Committed by BJJ.
-  Better date support, added support for PDFs that use non standard dates, support for time zone offsets Committed by BJJ.

-  FlateFilter-class now supports PNG-Predictors for decoding the imagedata, thanks to a contribution from Marcel Kammer Committed by BJJ.
-  Added support for extracting tiff images, thanks to a contribution from Marcel Kammer Committed by BJJ.
-  Added PDDocument.removePage to remove PDF pages Committed by BJJ.
-  Fixed issue when creating a COSString with a UTF 16 string Committed by BJJ.
-  Committed patch for type 1 PFB font parser(special thanks to Michael Niedermair) Committed by BJJ.
-  Committed patch for PNG predictors (special thanks Erik Martino) Committed by BJJ.
-  [ 1227428 ] failure of getMediaBox Committed by BJJ.
-  [ 1227426 ] null pointer in PDFToImage(ColorModel is null) Committed by BJJ.
-  [ 1207113 ] Enhancement: runtime accessible version Committed by BJJ.
-  [ 1213320 ] setFfFlag() of PDFField not working correctly Committed by BJJ.
-  [ 1215945 ] Error in COSString.writePDF() - fixed escape sequences Committed by BJJ.
-  [ 1198912 ] COSName with escaped characters not parsed correctly Committed by BJJ.
-  Fixed issue where resources were not being cleared in PDFStreamEngine Committed by BJJ.
-  [ 1165686 ] Expected int type parse error Committed by BJJ.
-  [ 1182825 ] Wrong handling of signed/unsigned byte/int in TTF parsing Committed by BJJ.
-  [ 1182892 ] PDFHighlight.setHighlightColor was removed because it is not implemented by adobe Committed by BJJ.




## 5 Version 0.7.1 (04/10/2005)

### 5.1 Changes to the Code Base

-  [ 1170068 ] text field is not found Committed by BJJ.















-  fixed NPE issue where an image did not have any applied filters Committed by BJJ.
-  Fixed issue where extra spaces were being added during text extraction for type3 fonts Committed by BJJ.
-  [ 1119420 ] Extract and Update the Meta-Information as XML Committed by BJJ.
-  [ 1119410 ] Extract text in/between bookmarks Committed by BJJ.
-  [ 1164476 ] XFDFImport should fail with non XFDF document Committed by BJJ.
-  [ 1119408 ] Support named target for Bookmark extraction. Committed by BJJ.
-  Created Resources/PDFBox\_External\_Fonts.properties to create a mapping for non-embedded fonts Committed by BJJ.
-  **\*\*API Change\*\*** Renamed PDFField.getName() to PDFField.getPartialName(), added method getFullyQualifiedName() Committed by BJJ.
-  **\*\*API Change\*\*** Renamed PDWidget to PDAnnotationWidget for naming consistency Committed by BJJ.
-  Text is now extracted from embedded form xobjects. Committed by BJJ.
-  Deployed site to new hosting vendor. Committed by BJJ.
-  committed code for PDFHighlighter to highlight words in a PDF document. Committed by BJJ.
-  Added command line application org.pdfbox.PDFToImage Committed by BJJ.
-  Implemented runlength decoding Committed by BJJ.
-  Added patch from Jorge Hernández Sellés to append content streams to existing page. Committed by BJJ.
-  **\*\*API Change\*\***renamed package from pdmodel.graphics.image to pdmodel.graphics.xobject Committed by BJJ.
-  **\*\*API Change\*\***Removed PDRadioButton, should use PDCheckbox instead Committed by BJJ.
-  **\*\*API Change\*\***COSStream now extends COSDictionary instead of containing a dictionary Committed by BJJ.
-  [ 1021241 ] Text extraction should follow PDF article divisions Committed by BJJ.
-  Added implementation for PDF page articles Committed by BJJ.
-  Created TextToPDF command line application Committed by BJJ.













-  Created ImageToPDF example Committed by BJJ.
-  fixed parsing of header where a trailing % exists Committed by BJJ.
-  [ 1110029 ] Character ">" not quoted in COSName::writePDF Committed by BJJ.

## 6 Version 0.7.0 (1/22/2005)








### 6.1 Changes to the Code Base





















-  committed [ 1097913 ] Enhance LucenePDFDocument streams(thanks to Olivier Parent) Committed by BJJ.
-  Added implementation for PDF Bookmarks Committed by BJJ.
-  Added implementation for PDF Destinations Committed by BJJ.
-  Updated website for better format for documentation Committed by BJJ.
-  Now ExportFDF and ExportXFDF will default output files to pdfname.fdf and pdfname.xfdf Committed by BJJ.
-  [ 1046278 ] ClassCastException when doing FDF/XFDF Committed by BJJ.
-  ExtractText now allows you to extract text if you decrypt with the owner password Committed by BJJ.
-  Added PDF 1.5 Object Stream support Committed by BJJ.
-  Added pdmodel.common.PDStream to represent COSStream Committed by BJJ.
-  changed PDPPage.getContents to use PDStream instead of COSStream Committed by BJJ.
-  Updated LucenePDFDocument Javadoc to tell which Lucene fields it populates Committed by BJJ.
-  moved HelloWorld example from persistence to pdmodel and updated to use new PD Model features Committed by BJJ.
-  Refactored PDFStreamEngine based on contributions from Christophe Huault Committed by BJJ.
-  This class no longer uses a gigantic if/else statement for all of the operators they are defined as properties when instantiating the class Committed by BJJ.








-  Updated AFM resources to be ones released on Adobe's site, include AFM license as well Committed by BJJL.
-  Added ability to embed TTF fonts, only WinAnsiEncoding is supported at this time Committed by BJJL.
-  Added ability to extract images, thanks to contributions by Brigitte Mathiak Committed by BJJL.
-  COSWriter now generates the document id if it does not already exist Committed by BJJL.
-  improved performance for text extraction Committed by BJJL.
-  [ 1058693 ] TextPosition does not take account of tz operator Committed by BJJL.
-  upgraded to log4j-1.2.9 Committed by BJJL.
-  include package-list for javadocs Committed by BJJL.
-  [ 1037145 ] Infinite loop in PDFParser.parseObject Committed by BJJL.
-  fixed error where spaces before integers was causing parse errors Committed by BJJL.

## 7 Version 0.6.7 (10/09/2004)

### 7.1 Changes to the Code Base










-  Revamped the way character spacing and font information is obtained Committed by BJJL.
-  Improved location information about a character drawn on the screen. Committed by BJJL.
-  Changed the PDFStreamEngine.showString to showCharacter to support the newly improved location information. This will now only show one character at a time. Committed by BJJL.
-  Fixed bug in PDDocument.isOwnerPassword and isUserPassword that was using the wrong length for the encryption key Committed by BJJL.
-  Upgraded to ant 1.6.2 Committed by BJJL.
-  Upgraded to checkstyle-3.4 Committed by BJJL.
-  Upgraded to JUnit-3.8.1 Committed by BJJL.

-  Upgraded to lucene-1.4.2 Committed by BJL.
-  Integrated patch(1016603) for issue 943319 to fix parsing of open office documents Committed by BJL.
-  Patch:985347 No longer throw exception for "No 'ToUnicode' and no 'Encoding' for Font" Committed by BJL.
-  Patch:996191 Fixed case statement with missing break Committed by BJL.
-  Patch:996781 Fixed null pointer exception in acroform fields Committed by BJL.
-  Renamed DecryptDocument to DocumentEncryption to support encryption and decryption Committed by BJL.
-  Added load/save/encrypt/decrypt convenience methods on the PDDocument class Committed by BJL.
-  COSWriter now attempts to keep object numbers from parsed documents and writes 'free' entries in the xref if necessary Committed by BJL.
-  Added the ability to set the word separator on the PDFTextStripper Committed by BJL.
-  Fixed issue where PDFBox would throw an IOException if a PDF was incorrectly missing an endobj tag Committed by BJL.
-  Fixed 918220 where PDFBox would freeze when parsing certain cmap files Committed by BJL.
-  Added initial colorspace support Committed by BJL.
-  Fixed issue where AppendDoc was throwing ClassCastException Committed by BJL.
-  Fixed 1013163 Can't parse filters that use filter abbreviation Committed by BJL.
-  Fixed 1011244 Where encrypting then decrypting was causing a problem Committed by BJL.
-  renamed TextPosition.getWidth to TextPosition.getCombinedHorizontalDisplacement to better reflect its actual value Committed by BJL.
-  Fixed 919215 PDFBox now support stream replacement Committed by BJL.
-  Fixed 955043 Added support for 'ETenms-B5-H' encoding Committed by BJL.
-  Fixed 996050 Class Cast exception when importing Committed by BJL.
-  Added support for Font descriptors Committed by BJL.

-  Fixed spacing issues when doing textfield FDF import Committed by BJJ.
-  Fixed 1017175 Large number converted when re-written Committed by BJJ.
-  Fixed 1029873 PDFBox now allows for multiple xref sections Committed by BJJ.
-  Added support for document Viewer Preferences Committed by BJJ.
-  Made currentDocument and pdfDocument protected in util.Splitter to allow easier subclassing Committed by BJJ.
-  Fixed 1034427 After Splitting page orientation is lost Committed by BJJ.
-  Added the following command line applications Committed by BJJ.

















## 8 Version 0.6.6 (07/20/2004)











### 8.1 Changes to the Code Base

-  Improved support for setting of checkbox fields(FDF import) Committed by BJJ.
-  Added the org.pdfbox.PDFSplit utility to split a single document into many documents Committed by BJJ.
-  PDFBox now ignore the Length field that is associated with a stream, it has been found to be wrong in some documents Committed by BJJ.
-  Fixed bug when writing out PDF documents and the document contained an non alphabetic character such as ( or ) Committed by BJJ.
-  Fixed bug in PDFont where dictionary encodings where not being processed correctly Committed by BJJ.
-  Fixed bug in COSDocument.isEncrypted which was comparing COSNull to the wrong object Committed by BJJ.
-  Integrated patch for supporting multiple lines in the appearance stream Committed by BJJ.
-  Upgraded to lucene-1.4-final Committed by BJJ.
-  org.pdfbox.ExtractText now uses the system encoding as the default encoding instead of ISO-8859-1 Committed by BJJ.

## 9 Version 0.6.5 (03/08/2004)






### 9.1 Changes to the Code Base







-  Fixed bug in revision 3 encryption algorithm Committed by BJJ.
-  added support for CIDFontType0 glyph widths, which fixed issue with spaces being during text extraction Committed by BJJ.
-  Fixed infinite loop when parsing a corrupt content stream Committed by BJJ.
-  Add characterspacing + wordspacing when determining the width of a space character Committed by BJJ.
-  Added support for more font types Committed by BJJ.
-  refactored the pdmodel.interactive package, form fields use object delegation instead of inheritance for the widget, see PDFField.getWidget and PDFField.getKids Committed by BJJ.
-  Fixed bug where an inheritable cropbox would cause stackoverflow exception Committed by BJJ.
-  Changed usage of PDFField/PDWidget to look like object delegation instead of inheritance by adding a PDFField.getWidget instead of extending PDWidget Committed by BJJ.
-  refactored interactive package, this will break any existing code that uses the PDFField/PDAnnotation classes. You will need to adjust your package names!! Committed by BJJ.
-  Now uses StandardEncoding as the default encoding Committed by BJJ.
-  Bug in AppendDoc example that did not take into account groups of pages Committed by BJJ.
-  PDFont now also tries the bootstrap classloader when loading AFM resources Committed by BJJ.
-  added -startPage and -endPage command line options to org.pdfbox.ExtractText Committed by BJJ.
-  Added support for corrupt PDFs with garbage before the header Committed by BJJ.
-  Fixed bug where there was whitespace instead of garbage characters in front of the first object Committed by BJJ.
-  performance improvements for the Matrix implementation Committed by BJJ.

-  upgraded to lucene 1.3 Committed by BJL.
-  fixed bug in cmap parser for cmap files that all ended in 'def' Committed by BJL.
-  Removed createObject method from COSDocument, COSWriter will handle all object references for you Committed by BJL.
-  Updated AppendDoc to use PDDocument instead of COSDocument and a couple bug fixes Committed by BJL.
-  PDFParser now closes the document if there were parse errors Committed by BJL.
-  TextPosition now has the PDFFont that is associated with the piece of text Committed by BJL.
-  Added initial version of org.pdfbox.PDFViewer, a GUI application to view the internal structure of a PDF document. This can be used for debugging purposes at this time but may end up being a Adobe Reader like application if there is enough interest Committed by BJL.
-  Changed COSNumber/COSInteger/COSFloat interface to have both intValue and longValue Committed by BJL.
-  Added methods isUserPassword & isOwnerPassword to PDDocument Committed by BJL.
-  Added cmap files for CJK languages, please give me some feedback Committed by BJL.

## 10 Version 0.6.4 (11/02/2003)










### 10.1 Changes to the Code Base

-  Fixed bug which caused infinite loop Committed by BJL.
-  Fixed bug in encoding where DictionaryEncoding kept a reference instead of making a copy leading to encoding problems Committed by BJL.
-  Added PDFTextStripper.(get|set)PageSeparator, which will allow the user to output a string after every page Committed by BJL.
-  refactored text stripping code to separate the logic processing of PDF operators and the logic of extracting text Committed by BJL.
-  ran findbugs on source code and fixed a couple minor issues Committed by BJL.

-  Refactored font functionality to PDFont, some API methods are no longer available in COSObject Committed by BJL.
-  changed name of org.pdfbox.Main to org.pdfbox.ExtractText Committed by BJL.
-  added contribution of org.pdfbox.Overlay from Mario Ivankovits Committed by BJL.
-  added log.isDebugEnabled checks to log4j calls Committed by BJL.
-  added better escaping when writing COSNames Committed by BJL.
-  fixed bug where encryption dictionary is sometimes set to COSNull instead of not being present Committed by BJL.












## 11 Version 0.6.3 (09/13/2003)

### 11.1 Changes to the Code Base

-  Now contains the ability to import/set FDF data thanks to a contribution from Stefan Uldum Grinsted Committed by BJL.
-  No longer throw an error when stream is not followed by 0A or 0D0A to allow more PDFs to be parsed Committed by BJL.
-  Added -encoding argument to org.pdfbox.Main to control the encoding of the output Committed by BJL.
-  Remove Prev entry from trailer if it exists because PDFBox automatically clears all old entries, only an issue when modifying/saving an existing PDF document Committed by BJL.
-  Fixed bug in master password encryption algorithm for Revision 3 encrypted documents Committed by BJL.
-  COSString no longer uses UTF-8 when encoding the byte array Committed by BJL.
-  Added PDDocument.getPageCount() Committed by BJL.
-  Fixed bug in PDFEncryption where Committed by BJL.
-  Now enforces text extraction permissions Committed by BJL.



## 12 Version 0.6.2 (4/18/2003)

### 12.1 Changes to the Code Base

-  Modified build so that build.properties settings are no longer required Committed by BJL.
-  Added required libraries to CVS Committed by BJL.
-  Added log4j logging Committed by BJL.
-  Significant text extraction work Committed by BJL.
-  Added automatic handling of files encrypted with the empty password Committed by BJL.
-  Added automated tests and test data for text extraction Committed by BJL.
-  Removed unimplemented decoders from filters test Committed by BJL.
-  Fixed several LZW decode bugs introduced after 0.5.6 Committed by BJL.
-  Fixed bugs relating to processing out of spec PDF's with bad # escaping in the name ("java.io.IOException: Error: expected hex number" bug) Committed by BJL.
-  Fixed Lucene UID generation bug Committed by BJL.
-  Fixed GetFontWidths null pointer exception bug Committed by BJL.

## 13 Version 0.6.1 (3/9/2003)

### 13.1 Changes to the Code Base







-  Fixed bug in parsing stream objects which led to "Unexpected end of ZLIB input stream" Committed by BJL.
-  Changed license from LGPL to BSD to allow pdfbox to be used easily in Apache projects Committed by BJL.

## 14 Version 0.6.0 (3/5/2003)

### 14.1 Changes to the Code Base








-  Massive improvements to memory footprint Committed by BJL.



-  Must call close() on the COSDocument(LucenePDFDocument does this for you) Committed by BJJL.
-  Really fixed the bug where small documents were not being indexed Committed by BJJL.
-  Fixed bug where no whitespace existed between obj and start of object. Exception in thread "main" java.io.IOException: expected='obj' actual='obj<</Pro Committed by BJJL.
-  Fixed issue with spacing where textLineMatrix was not being copied properly Committed by BJJL.
-  Fixed 'bug' where parsing would fail with some pdfs with double endobj definitions Committed by BJJL.
-  Added PDF document summary fields to the lucene document Committed by BJJL.



## 15 Version 0.5.6 (11/28/2002)



### 15.1 Changes to the Code Base

-  Fixed bug in LucenePDFDocument where stream was not being closed and small documents were not being indexed Committed by BJJL.
-  Fixed a spacing issue for some PDF documents Committed by BJJL.
-  Fixed error while parsing the version number Committed by BJJL.
-  Fixed NullPointerException in persistence example Committed by BJJL.
-  Create example lucene IndexFiles class which models the demo from lucene Committed by BJJL.
-  Fixed bug where garbage at the end of file caused an infinite loop Committed by BJJL.
-  Fixed bug in parsing boolean values with stuff at the end like "true>>" Committed by BJJL.

## 16 Version 0.5.5 (10/03/2002)



### 16.1 Changes to the Code Base

-  Added example of printing document signature Committed by BJJL.
-  Added example to print out form fields values Committed by BJJL.

-  Fixed bug when appending documents Committed by BJL.
-  Various other bug fixes Committed by BJL.


## 17 Version 0.5.4 (09/17/2002)

### 17.1 Changes to the Code Base

-  Fixed bug in text output where '?' instead of the proper character Committed by BJL.
-  Fixed bug where sections of text were not being output at all Committed by BJL.



## 18 Version 0.5.3 (09/13/2002)

### 18.1 Changes to the Code Base

-  Fixed bug in 128 bit encryption Committed by BJL.





## 19 Version 0.5.2 (09/06/2002)

### 19.1 Changes to the Code Base

-  Fixed bug where FDF documents could not be appended to PDF Documents Committed by BJL.
-  Catch all NumberFormatExceptions and wrap them with IOExceptions Committed by BJL.





## 20 Version 0.5.1 (09/04/2002)

### 20.1 Changes to the Code Base

-  Now supports unicode for the document summary Committed by BJL.
-  Better support for Type0 fonts Committed by BJL.
-  Fixed bug with an empty LZW stream Committed by BJL.
-  Fixed parsing error for ID operator Committed by BJL.


## 21 Version 0.5.0 (08/31/2002)

### 21.1 Changes to the Code Base

-  Now supports unicode for the document summary Committed by BJL.
-  Better support for Type0 fonts Committed by BJL.
-  Fixed bug with an empty LZW stream Committed by BJL.
-  Fixed parsing error for ID operator Committed by BJL.





## 22 Version 0.4.1 (07/25/2002)

### 22.1 Changes to the Code Base

-  Fixed bug where .notdef was being output as document text Committed by BJL.

## 23 Version 0.4.0 (07/23/2002)

### 23.1 Changes to the Code Base

-  Added extract text ant task Committed by BJL.
-  Implemented AFM(Adobe Font Metrics) resource loading Committed by BJL.
-  Fixed numerous bugs submitted by users Committed by BJL.
-  Changed project from pdfparser to pdfbox to better reflect future needs Committed by BJL.



## 24 Version 0.3.0 (07/09/2002)

### 24.1 Changes to the Code Base

-  Added indexer for the lucene project Committed by BJL.
-  Initial implementation of PDF encryption(not working yet) Committed by BJL.

## 25 Version 0.2.0 (06/03/2002)

### 25.1 Changes to the Code Base

-  Added support for the various encodings Committed by BJL.
-  Improved the accuracy of the text output Committed by BJL.

## 26 Version 0.1.0 (05/25/2002)

### 26.1 Changes to the Code Base

-  Initial Version Committed by BJL.