3-Heights™ PDF Analysis & Repair Shell

Version 4.5

User Manual

Contact: pdfsupport@pdf-tools.com

Owner: **PDF Tools AG**

                    Kasernenstrasse 1
                    8184 Bachenbülach
                    Switzerland

                    http://www.pdf-tools.com

Table of Contents

1 Introduction

1.1 Description

The 3-Heights™ PDF Analysis & Repair tool is used to analyze, repair and restore the content of corrupt PDF documents.

Unfortunately, the number of corrupt PDF documents is incredibly huge. The cause is usually down to defective generating tools, converters and other influences such as attempts at manual editing, copying via FTP without correct settings, system crashes during PDF creation, network interruptions, defective copying on optical media, etc.

The result leads to an enormous loss of important information and to production downtimes caused by corrupt PDF documents.

The 3-Heights™ PDF Analysis & Repair analyzes PDF documents with regard to PDF specifications. Defective files are automatically repaired as far as possible and unreadable data is restored.
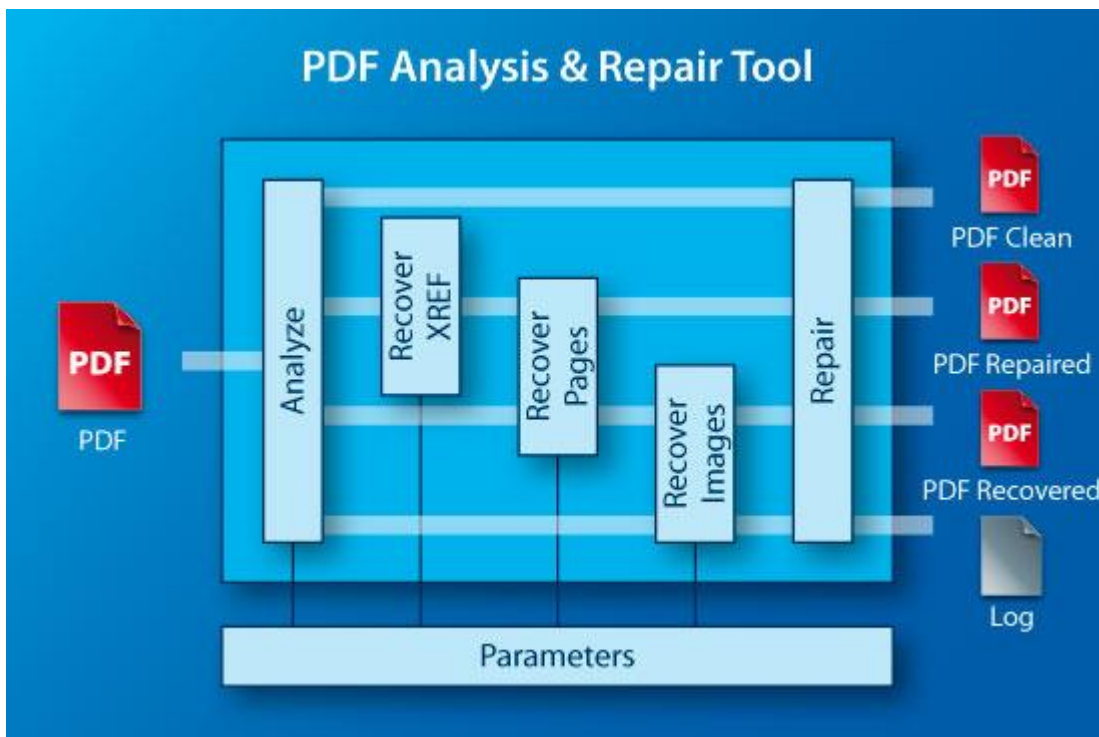
1.2 Functions

PDF Analysis & Repair is used to check and, where indicated, repair PDF documents. Users can determine customized profiles from a broad range of analysis and repair options. An exact and detailed description is issued for each reported error. The tool is also capable of reading and processing encrypted PDF files without any problems.

Features

Analyze and/or repair one or more PDF documents

Set analysis options, including:

- Objects
- Page tree
- Content stream

Set repair options, including:

- Restore data streams
- Restore fonts
- Restore XRef table
- Restore pages
- Restore images if pages cannot be restored

Display error description for every message, including:

- Type (errors, warnings, information)
- Error code
- Text-based description
- Page number
- Number of events

Write error messages to log file

Read encrypted PDF files

Encrypt restored file and set user authorizations

Differentiate between Repair (corrects the errors in the document) and Restore (recreates the document based on the remaining legible information)

Formats

Input Formats

- PDF 1.x (e.g. PDF 1.4, PDF 1.5, etc.), PDF/A

Target Formats

- PDF 1.x (e.g. PDF 1.4, PDF 1.5, etc.)

Compliance

Standards: ISO 32000 (PDF 1.7)

1.3   Operating Systems

- Windows XP, Vista, 7, 8, 8.1 - 32 and 64 bit

- Windows Server 2003, 2008, 2008 R2, 2012, 2012 R2 - 32 and 64 bit
- HP-UX 11 and later PA-RISC2.0 32 bit or HP-UX 11i and later ia64 (Itanium) 64 bit
- IBM AIX 5.1 and later (64 bit)
- Linux (32 and 64 bit)
- Mac OS X 10.4 and later (32 and 64 bit)
- Sun Solaris 2.8 and later, SPARC and Intel
- FreeBSD 4.7 and later 32 bit or FreeBSD 9.3 and later 64 bit (on request)

2   Installation

2.1   Installing the 3-Heights™ PDF Analysis & Repair Shell

1. Download the ZIP archive of the product from your download account at www.pdf-tools.com.

2. Open the ZIP archive.

3. Check the appropriate option to preserve file paths (folder names) and unzip the archive to a local folder (e.g. *C:\program files\pdf-tools*).

4. The unzip process now creates the following subdirectories:

   - *Bin*: Contains the runtime executable binary code

   - *Doc*: Contains documentation files

5. To start the 3-Heights™ PDF Analysis & Repair Tool from a shell, the directory needs to be included in the "Path" environment variable.

How to set the Environment Variable "Path"

To set the environment variable "Path" on Windows 2000: Go to **Start** -> **Settings** -> **Control Panel** -> **System** -> **Advanced** -> **Environment Variables**

Windows XP: Go to **Start** -> **Control Panel** (classic view) -> **System** -> **Advanced** -> **Environment Variables**.

Select "Path" and **Edit**, then add the directory where *pdfrepair.exe* is located to the "Path". If the environment variable "Path" does not exist, create it.
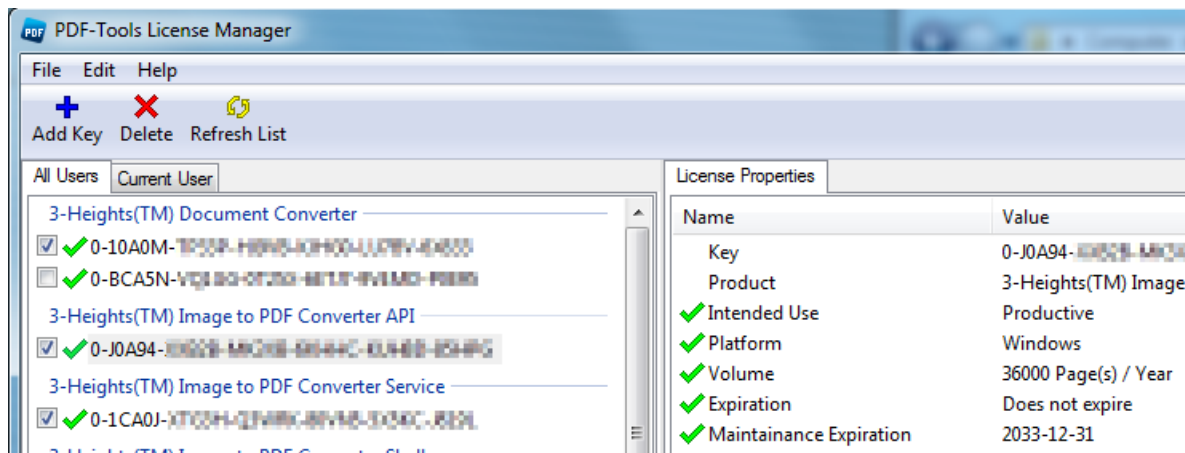
3   License Management

There are three possibilities to pass the license key to the application:

1. The license key is installed using the GUI tool (Graphical user interface). This is the easiest way if the licenses are managed manually. It is only available on Windows.

2. The license key is installed using the shell tool. This is the preferred solution for all non-Windows systems and for automated license management.

3. The license key is passed to the application at runtime via the command line switch `-lk` property. This is the preferred solution for OEM scenarios.

3.1   Graphical License Manager Tool

The GUI tool *LicenseManager.exe* is located in the *bin* directory of the product kit.

List all installed license keys

The license manager always shows a list of all installed license keys on the left pane of the window. This includes licenses of other PDF Tools products.

The user can choose between:

- Licenses available for all users. Administrator rights are needed for modifications.
- Licenses available for the current user only.

Add and delete license keys

License keys can be added or deleted with the "Add Key" and "Delete" buttons in the toolbar.

- The "Add key" button installs the license key into the currently selected list.
- The "Delete" button deletes the currently selected license keys.

Display the properties of a license

If a license is selected in the license list, its properties are displayed in the right pane of the window.

Select between different license keys for a single product

More than one license key can be installed for a specific product. The checkbox on the left side in the license list marks the currently active license key.

3.2 Command Line License Manager Tool

The command line license manager tool *licmgr* is available in the *bin* directory for all platforms except Windows.

A complete description of all commands and options can be obtained by running the program without parameters:

```
licmgr
```

List all installed license keys

```
licmgr list
```

The currently active license for a specific product is marked with a star '*' on the left side.

Add and delete license keys

Install new license key

```
licmgr store X-XXXXX-XXXXX-XXXXX-XXXXX-XXXXX-XXXXX
```

Delete old license key

```
licmgr delete X-XXXXX-XXXXX-XXXXX-XXXXX-XXXXX-XXXXX
```

Both commands have the optional argument `-s` that defines the scope of the action:

- `g`: For all users
- `u`: Current user

Select between different license keys for a single product

```
licmgr select X-XXXXX-XXXXX-XXXXX-XXXXX-XXXXX-XXXXX
```

3.3 License Key Storage

Depending on the platform the license management system uses different stores for the license keys.

Windows

The license keys are stored in the registry:

- HKLM\Software\PDF Tools AG   (for all users)
- HKCU\Software\PDF Tools AG   (for the current user)

Mac OS X

The license keys are stored in the file system:

- /Library/Application Support/PDF Tools AG   (for all users)
- ~/Library/Application Support/PDF Tools AG   (for the current user)

Unix / Linux

The license keys are stored in the file system:

- /etc/opt/pdf-tools   (for all users)
- ~/.pdf-tools   (for the current user)

Note: The user, group and permissions of those directories are set explicitly by the license manager tool.

It may be necessary to change permissions to make the licenses readable for all users. Example:

```
chmod -R go+rx /etc/opt/pdf-tools
```

4   Getting Started and User's Guide

4.1   General Settings

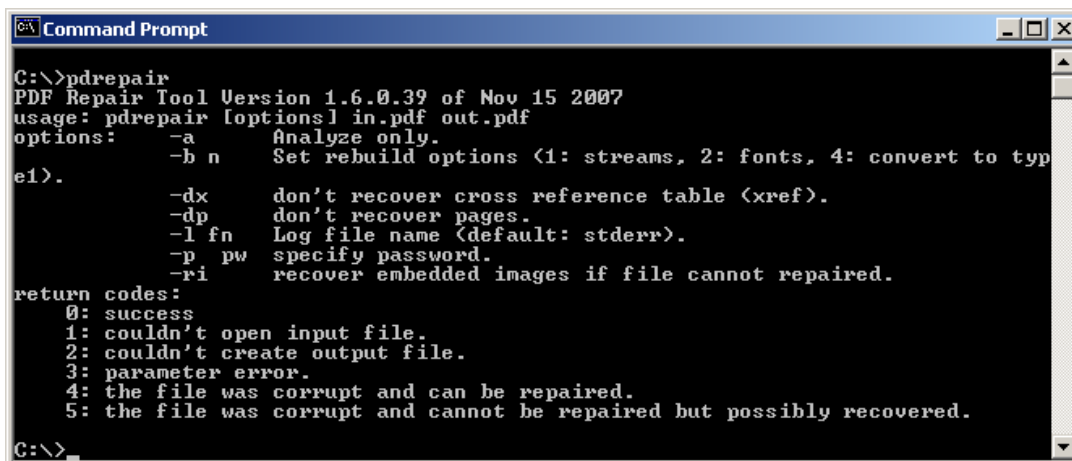Pass a license key to the application at runtime instead of installing it on the system.

```
pdrepair -lk X-XXXXX-XXXXX-XXXXX-XXXXX-XXXXX-XXXXX myfile.pdf
myfile_rep.pdf
```

This is only required in an OEM scenario.

4.2   Usage

Type pdrepair in the shell, and the usage is returned.

```
Command Prompt                                                    _ □ ×

C:\>pdrepair
PDF Repair Tool Version 1.6.0.39 of Nov 15 2007
usage: pdrepair [options] in.pdf out.pdf
options:    -a      Analyze only.
            -b n    Set rebuild options (1: streams, 2: fonts, 4: convert to typ
e1).
            -dx     don't recover cross reference table (xref).
            -dp     don't recover pages.
            -l fn   Log file name (default: stderr).
            -p  pw  specify password.
            -ri     recover embedded images if file cannot repaired.
return codes:
    0: success
    1: couldn't open input file.
    2: couldn't create output file.
    3: parameter error.
    4: the file was corrupt and can be repaired.
    5: the file was corrupt and cannot be repaired but possibly recovered.

C:\>_
```

4.3　Repair a File

The 3-Heights™ PDF Analysis & Repair Tool requires at least two parameters: A name of an existing PDF input file and the desired name for the repaired PDF output file

Example: *Read the damaged input file "myfile.pdf". Repair the file and save the result in a new file called "myfile_rep.pdf".*

```
pdrepair myfile.pdf myfile_rep.pdf
```

4.4　Specify the Folder of the Output File

The output folder can simply be added in front of the output file name

```
pdrepair myfile.pdf myfolder\myfile_rep.pdf
```

or absolute (Windows):

```
pdrepair myfile.pdf C:\myfolder\myfile_rep.pdf
```

4.5　Repairing all Files in a Directory

The 3-Heights™ PDF Repair Shell Tool reads the input file while it already writes on the output file. For this reason, it is not possible to directly overwrite the input file.

If you would like to repair all PDF documents in a directory, it is required to use a variable to name the output files. Here is an example using the FOR command of the CMD shell:

Example: *Process and repair all files in the current directory.*

```
for %i in (*.pdf) do pdrepair %i %~ni_rep.pdf
```

Of course, one can adjust the paths, or use a different output name.

If you would like the repaired file to have the same name as the original file, it is suggested to use the following process:

1. Repair the file and create a repaired copy of the original file (like in the for loop above).

2. Ensure the repaired documents are generated correctly. For example check the return code of the repair tool and require it to be 0 or 4, or ensure the files are not empty (i.e. just a few bytes in size).

3. When you are sure the repaired file is okay and you do not need the original file anymore, delete the original file and rename the recovered file.

4.6   Difference between Repairable and Recoverable

Repairable: The input PDF file contains errors which are repairable. These are for example PDF syntax errors. In the repaired PDF output file, these errors are fixed.

Recoverable: The input PDF file is missing data. For example data of an embedded image, which got lost/overwritten when copying the file from one location to another failed partially. When relevant data is lost, the file is irreparable, however the file can be recovered, such that the output PDF is valid according to the PDF Specification.

Example: Assuming an original PDF (file O) is valid.

- Somehow the file gets corrupted (-> file C) and file O gets lost.

- If file O can be rebuilt based on file C, then file C is repairable.

- If a new file N can be built based on file C, and file N is a valid PDF and contains (part of) the content of file O, then file C is recoverable.

5   Reference Manual

5.1   Switches

-a   Analyze Only

By default, documents will be analyzed and invalid PDF files will be repaired or recovered. With the switch **–a**, only the analysis process will be executed. The results of the analysis will be displayed on screen.

Example: *Analyze the file "myfile.pdf".*

```
pdrepair –a myfile.pdf
```

-b   Set the Rebuild Options

This options controls what parts of the PDF are to be repaired. Available options are:

1: Rebuild streams
2: Rebuild fonts
4: Convert CFF fonts to Type1 fonts.

If 4 is applied, the compressed fonts are decompressed, this can potentially lead to an increase of the file size.

If multiple options are to be selected, add the values.

Example: *Repair file and rebuild all.*

```
pdrepair –b 7 myfile.pdf myfile_rep.pdf
```

-dp   Do not Recover Pages

By default, if pages are not part of the page tree (loose pages), they will be recovered and added at the end of the document. With the switch **–dp**, these pages are recovered, but removed from the document.

-dx   Do Not Repair Cross-Reference Table

This option repairs the PDF document except for the cross-reference table. This option is useful if it is taking too long to process a document, since repairing the cross-reference table is very time consuming.

-l   Set the Log File

Using the switch **–l** followed by a file name, log messages during the repair process are written into the specified log file. If no log file is specified, the log messages are written to standard error (stderr).

Example: Write messages to the log file "log.txt" instead of standard error.

```
pdrepair –l log.txt myfile.pdf myfile_rep.pdf
```

-p   Specify Password

If the input file is encrypted with a user password, a password needs to be provided to read the input PDF document file. This can be either the user or owner password.

Example: Process a PDF file which is encrypted. Either the owner or the user password of the encrypted PDF file is "password".

```
pdprepair –p password myfile.pdf myfile_rep.pdf
```

-ri   Recover Embedded Images

This option allows you to recover imbedded images if the file cannot be repaired.

Example: Images are extracted and stored as TIFF or JPEG files. The following command recovers all images that it can find in the resources of "myfile.pdf".

```
pdrepair –ri myfile.pdf myfile_rep.pdf
```

It extracts the images and names them as:

`img{objno}.tif` *or*

`img{objno}.jpg.`

where objno is the PDF object number that originally contained the image.

-lk   Set License Key

Pass a license key to the application at runtime instead of installing it on the system.

5.2   Return Codes

All return codes other than "0" indicate a warning or error in the processing.

Table: Return Codes	
Value	*Description*
0	Success
1	PDF Input File could not be opened or invalid parameters
2	PDF Output File could not be created
3	Invalid option or option values were entered
4	PDF input file is corrupt and can be repaired
5	PDF input file is corrupt and cannot be repaired but possibly be recovered

6 Troubleshooting

6.1 The Repair Time Takes Too Long

Try excluding the cross-reference table (switch **–dx**). This will speed up the repair time.

6.2 The File Cannot Be Repaired

If the switch **–dx** has been selected and the cross-reference table is corrupt, the file cannot be repaired. Ensure the switch **–dx** is not set.