

3-Heights® PDF Analysis & Repair Service

Version 6.27.6



Contents

1	Introduction	3
1.1	Description	3
1.2	Functions	3
1.2.1	Features	4
1.2.2	Formats	4
1.2.3	Conformance	4
1.3	Operating systems	5
2	Installation	6
2.1	Overview	6
2.2	Windows	6
2.3	Uninstall	6
2.4	Note about the evaluation license	6
3	License management	7
4	Getting started	8
4.1	Configuration	8
4.1.1	Retrieve information about available options and settings	8
4.2	Managing the service	9
4.2.1	Service state diagram	10
4.3	Using the service	11
4.4	Log files	11
5	Interface reference	13
5.1	Service control commands	13
5.1.1	-a Pause service	13
5.1.2	- c Create service	13
5.1.3	-d Delete service	13
5.1.4	-i List the usage	13
5.1.5	-o Continue service	14
5.1.6	-q Query current status of service	14
5.1.7	-s Start service	14
5.1.8	-t Stop service	14
5.1.9	-x Run as executable	14
5.2	Configuration options	15
5.2.1	PdfRepairSvr.ini configuration file	15
	Autodelete of successfully processed files	
	Job number prefix	
	Logpath	
	Polling interval	
5.2.2	-a Analyze only	17
5.2.3	-b Set the rebuild options	17
5.2.4	-dp Do not recover pages	17
5.2.5	-dx Do not repair cross-reference table	18
5.2.5 5.2.6	-p Read an encrypted PDF file	18
5.2.0 5.2.7	-w Specify the path to the root directory	18
5.2.7 5.2.8	-wd Specify the drop path	19
5.2.6 5.2.9	-wd specify the drop pathwfi lanore files with certain extensions	19
ノ.ム.ツ	-WI 1 IGHULE HIES WITH CELEMINE EXTENSIONS	17

5.2.10	-wfs Process only files with certain extensions	19
6	Troubleshooting	20
6.1	Repair time takes too long	20
6.2	File cannot be repaired	20
7	Version history	21
7.1	Changes in versions 6.19–6.27	21
7.2	Changes in versions 6.13–6.18	21
7.3	Changes in versions 6.1–6.12	21
7.4	Changes in version 5	21
7.5	Changes in version 4.12	21
7.6	Changes in version 4.11	21
7.7	Changes in version 4.10	21
7.8	Changes in version 4.9	21
7.9	Changes in version 4.8	22
8	Licensing, copyright, and contact	23

1 Introduction

The 3-Heights® PDF Analysis & Repair Service is a ready-to-use product that allows to install a Windows NT service process to automatically analyze and repair PDF documents from watched folders.

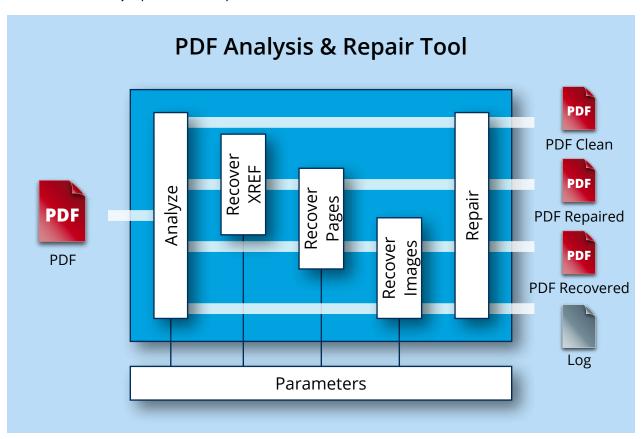
1.1 Description

The 3-Heights® PDF Analysis & Repair Service 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, and defective copying on optical media.

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 Service 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 Service 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.

1.2.1 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 permission flags
- Set error level to identify whether errors, warnings or merely information occur
- Set reporting level to determine the messages to be issued (errors, warnings, information)
- Differentiate between "Repair" (corrects the errors in the document) and "Restore" (recreates the document based on the remaining legible information)

1.2.2 Formats

Input formats

- PDF 1.x (PDF 1.0, ..., PDF 1.7)
- PDF 2.0
- PDF/A-1, PDF/A-2, PDF/A-3

Output formats

- PDF 1.x (PDF 1.0, ..., PDF 1.7)
- PDF 2.0

1.2.3 Conformance

Standards:

- ISO 32000-1 (PDF 1.7)
- ISO 32000-2 (PDF 2.0)

1.3 Operating systems

The 3-Heights® PDF Analysis & Repair Service is available for the following operating systems:

- Windows Client 7+ | x86 and x64
- Windows Server 2008, 2008 R2, 2012, 2012 R2, 2016, 2019, 2022 | x86 and x64

^{&#}x27;+' indicates the minimum supported version.

2 Installation

2.1 Overview

The PDF Analysis & Repair Service is configured by the file PdfRepairSvr.ini, which needs to be located in the same directory as the executable pdfrepairsvr.exe. Before starting the service, the configuration file needs to be adjusted. The procedure to perform this step is described in PdfRepairSvr.ini configuration file.

Once configured, the service can be created, started, paused, continued, stopped, and deleted via the command line. To use the create and delete functions, administrator permissions are required. To start and stop the service, operator permissions are required.

When the service is running, it processes PDF documents that are copied or moved into watched folders.

They are then renamed and moved to the folder Jobs. The renaming gives the PDF a 16 character long timestamp to create unique job tickets. This ensures there are no conflicts with documents that have the same name.

2.2 Windows

The 3-Heights® PDF Analysis & Repair Service comes as an MSI installer.

To install the software, proceed as follows:

- 1. You need administrator rights to install this software.
- 2. Log in to your download account at https://www.pdf-tools.com. Select the product "PDF Analysis & Repair Service". If you have no active downloads available or cannot log in, please contact pdfsales@pdf-tools.com for assistance.

You can find different versions of the product available. Download the version that is selected by default. You can select a different version.

The product comes as an MSI (Microsoft Installer) package that provides an installation routine for installing and uninstalling the 3-Heights® PDF Analysis & Repair Service.

- The package installs the 64-bit version, which runs on 64-bit platforms only.
- 3. Start the MSI package and follow the steps in the installation routine.

2.3 Uninstall

If you have used the MSI for the installation, go to Start \rightarrow 3-Heights $^{\circ}$ PDF Analysis & Repair Service... \rightarrow Uninstall

2.4 Note about the evaluation license

With the evaluation license, the 3-Heights® PDF Analysis & Repair Service automatically adds a watermark to the output files.

3 License management

The 3-Heights® PDF Analysis & Repair Service requires a valid license in order to run correctly. If no license key is set or the license is not valid, then an error message is printed to the service log.

More information about license management is available in the license key technote.

4 Getting started

4.1 Configuration

Before starting the PDF Analysis & Repair Service for the first time, the file PdfRepairSvr.ini needs to be modified. Editing this file while the PDF Analysis & Repair Service is running has no impact. The service first needs to be stopped and restarted after the modification. When opening this file with a text editor, it looks like this:

```
[PdfRepairSvr]
AutoDelete=True
Threads=2
Thread1=-w C:\PdfRepairSvr\Analyze -a
Thread2=-w C:\PdfRepairSvr\Repair
```

The meaning of these keys and values in this example is as follows:

AutoDelete=True This option automatically deletes a PDF file after it is processed successfully. When set to **False**, the processed file is copied to the sub directory Succeeded.

Threads The given value stands for the total number of concurrent threads. Each thread can have its own assigned settings. One thread corresponds to one watched folder.

Threads1= Sets the options such as name of watched folder and settings etc. for thread 1.

- -w C:\pdfrepairsvr\Analyze Creates a watched folder with the given name for this thread. The path must be an absolute path. Network mapped drive letters or relative paths or driver letters mapped via the subst command are not recognized, because the service process per default runs under the "LocalSystem" account. (The user can be changed as described in Managing the service.)
- -a (Option) "Analyze only" option.

This means that any PDF document that is moved or copied to the folder C:\ProductBinary\Analyze or C:\pdfrepairsvr\Repair is processed by the service. One thread only analyzes, the other also repairs or recovers documents.

Note: Any string, such as a file name, that contains spaces must be enclosed in quotation marks. For example, if the watched folder contains spaces in its path, the entire path needs to be quoted: -w "C:\A path\with spaces".

4.1.1 Retrieve information about available options and settings

A quick overview of all configuration options and service control commands that the 3-Heights® PDF Analysis & Repair Service supports can be output in the form of a usage message in the command line.

To display this information, open a Windows command line (cmd.exe) and then type:

pdfrepairsvr

(See also Service control commands)

A short overview of all the options that can be configured in the PdfRepairSvr.ini is displayed when typing the following in a Windows command line:

pdfrepairsvr -i

(See also Configuration options)

4.2 Managing the service

Once the configuration is done, the service can be started and controlled by executing pdfrepairsvr.exe on the command line. The path can be omitted if the pdfrepairsvr.exe is included in the %PATH% environment variable.

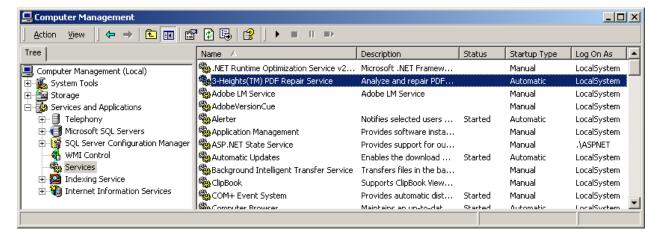
Note: It is essential that the executable pdfrepairsvr.exe and the configuration file PdfRepairSvr.ini be on a non-mapped drive.

Note: To create or delete the service, administrator permissions are required.

1. To create the service, use the -c option.

pdfrepairsvr -c

After executing this command, the service is created. It is now visible in the "Computer Management" window under "Services". To open the "Computer Management" window, go to Start \rightarrow Control Panel \rightarrow Administrative Tools \rightarrow Computer Management or simply right-click the icon "My Computer" on the desktop and select "manage". If the services is created correctly, it appears as "3-Heights® PDF Analysis & Repair Service" as shown in the image below.



2. By default, the 3-Heights® PDF Analysis & Repair Service runs in the "LocalSystem" account. After the service has been created, the user can be changed.

This is required in a situation where a network share is used as a watched folder and the process needs to run under a user with the appropriate access permission rights, since the account "LocalSystem" does not have any permissions on remote systems.

To change the user, right-click the service in the Services window and select "Properties". Then change the user in the tab "Log On".

3. Once created, the service can be started with the option <u>-s</u>.

pdfrepairsvr -s

- 4. Now the 3-Heights® PDF Analysis & Repair Service is up and running, and files can be moved, copied or dragand-dropped into the watched folder.
- 5. To stop the service, use the option <u>-t</u>.

pdfrepairsvr -t

To restart, use <u>-s</u> again.

6. To delete the service, use the option <u>-d</u>.

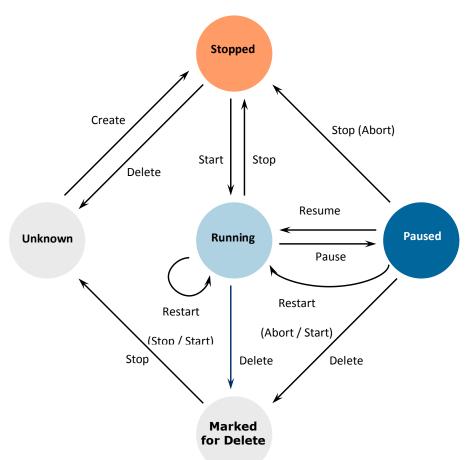
pdfrepairsvr -d

4.2.1 Service state diagram

The 3-Heights® PDF Analysis & Repair Service behaves as described in the state diagram below.

If "Stop" is called when the service is in the "Paused" state, the current job is aborted. This means the current page is finished processing, then the job is terminated.

If "Stop" is called when the service is the "Running" state, the current job (all pages) is finished. Then the service is stopped.



4.3 Using the service

Once the service is created and started, the watched folders configured in PdfRepairSvr.ini are created automatically. In each watched folder, the following subfolders are created:

- Jobs
- InProgress
- Succeeded
- Failed
- Recovered
- Repaired
- Logs

When a file is moved, copied, or drag-and-dropped into the configured watched folder, the service performs these tasks:

- 1. Each file is moved to the Jobs subfolder. While moving, the file is renamed by adding a 16 character long jobnumber prefix. This ensures a well-defined processing order and unique file names.
- 2. A worker-thread takes the file from the Jobs folder and moves it to InProgress. The file is then processed.
- 3. Depending on the outcome of the processing, the following is done:

The file was processed successfully

- The input file is moved to the Succeeded folder or it is deleted, depending on whether AutoDelete or AutoDeleteAll is set to true or false in the configuration file PdfRepairSvr.ini.
- Repaired documents are stored in Repaired, recovered documents in Recovered.

The file was not processed successfully

- The input file is moved to the Failed folder or it is deleted, depending on whether AutoDeleteALL is set to true or false in the configuration file PdfRepairSvr.ini.
- 4. In any case, an entry in the log file of this thread is created.

4.4 Log files

There are two types of log files.

The log file per thread Each thread (watched folder) has a log file. The log file resides in the same directory as the pdfrepairsvr.exe executable and the PdfRepairSvr.ini configuration file. It is named PdfRepairSvr-log-<n>.txt, where the number of the log file <n> is increased whenever the service is re-started. The log file is locked by the service as long as the service is running.

- The log file contains general messages (including a timestamp that is not shown here) such as:
 - [1] Worker thread for directory C:\pdfrepairsvr\Folder started.
- Error messages such as:
 - * Error 0 while opening file C:\pdfrepairsvr\Folder\InProgress\Job-...

There are three types of messages: information, warnings, and errors. They are labeled with the corresponding letters "I", "W", and "E".

Here is an example of how a log file looks:

Open file.

Analyze Objects.

0x00418018 - W - The generation 0 of reference doesn't match with the generation 1 of the object. 0x80410028 - E - The "endobj" keyword is missing.

Analyze Pages.

0x80410016 - E - The file is corrupt and cannot be repaired. The file can possibly be recovered.

Save output file.

Close file.

5 Interface reference

5.1 Service control commands

These options are used to control the service. The create and delete functions require administrator rights. The start and stop functions require operator rights.

5.1.1 -a Pause service

Pause service -a

This option pauses the service.

pdfrepairsvr -a

5.1.2 - C Create service

Create service -c

The 3-Heights® PDF Analysis & Repair Service is created using the -c option.

pdfrepairsvr -c

Important: drive.

It is essential that pdfrepairsvr.exe is on a non-mapped

5.1.3 -d Delete service

Delete service -d

The 3-Heights® PDF Analysis & Repair Service can be deleted with the -d option. It is best used after the service has already been stopped.

pdfrepairsvr -d

5.1.4 -i List the usage

List the usage -i

The -i option lists the current version and date of the service along with all available settings.

pdfrepairsvr -i

5.1.5 - O Continue service

Continue service - o

This option resumes the service.

pdfrepairsvr -o

5.1.6 -q Query current status of service

Query current status of service -q

This option returns the current status of the service.

pdfrepairsvr -q
The service starts automatically during system startup.
The service is stopped.
[pdfrepairsvr] QueryService: The operation completed successfully.

5.1.7 -s Start service

Start service -s

Once created, the 3-Heights® PDF Analysis & Repair Service can be started with the -s option.

pdfrepairsvr -s

5.1.8 -t Stop service

Stop service -t

To stop the service, use the -t option.

pdfrepairsvr -t

If "stop" is called while the service is "running", the current job (all pages) will be finished, after that the service is stopped.

If the service was "paused" before calling "stop", the current page will be finished processing. After that page, the job is aborted.

5.1.9 -x Run as executable

Run as executable -x

With this option, the PDF Analysis & Repair Service runs as an executable instead of as a Windows Service. It provides the same functionality as long as the executable is "running".

pdfrepairsvr -x

5.2 Configuration options

Analyzing and repairing or recovering PDF documents is a complex task. To increase to processing speed, certain analyze and recovery options can be disabled.

5.2.1 PdfRepairSvr.ini configuration file

The PdfRepairSvr.ini configuration file defines the setting for the watched folders. It is read on starting the service.

[PdfRepairSvr]	Required	
AutoDelete=	Optional	true or false
AutoDeleteAll=	Optional	true or false
LogPath=	Optional	Either a path like C:\mypath\log or the keyword EventLog
PollingInterval=	Optional	Value in milliseconds, default 1000
JobPrefix=	Optional	true or false
LogLevel=	Optional	0 or 1
Threads=n	Required	The number of threads
Thread1=-w	Required	Options for the first thread
Thread2=-w		Options for the second thread
Threadn=		There must be exactly as many threads as defined in Threads=n.

Example:

```
[PdfRepairSvr]
AutoDelete=true
LogPath=EventLog
JobPrefix=false
Threads=2
Thread1=-w C:\PdfRepairSvr\Analyze -a
Thread2=-w C:\PdfRepairSvr\Repair
```

Autodelete of successfully processed files

When a repair job succeeded, the PDF document is moved from the Jobs folder to the Repaired folder or Recovered folder. To automatically delete the input file after it has be successfully repaired or recovered, the value AutoDelete can be set to true in the control file PdfRepairSvr.ini. The documents which cannot be repaired nor recovered are still copied to the Failed subfolder.

[PdfRepairSvr]
AutoDelete=true

To delete failed documents as well, use the following setting:

[PdfRepairSvr]
AutoDeleteAll=true

Job number prefix

Every time a document is copied from the watched folder to the Jobs subfolder, it is renamed by adding a 21 character prefix containing a timestamp of the form Job- <8 digits>- <8 digits>. For example

Note: When JobPrefix=false then the incoming documents must have unique file names.

Job-01C61DD4-E72E1BCE

The job number prefix ensures that several documents with the same name can correctly be processed. Adding the prefix can be prevented with the following line in the configuration file:

[PdfRepairSvr]
JobPrefix=false

Logpath

Log messages created by the service are by default written to the \log subdirectory. To alter the directory, add a line similar as shown below to the configuration file:

[PdfRepairSvr]
LogPath=C:\path\log

Messages created by the service can be added to the system's application event log instead of written to a log file. This is achieved by adding the following line to the configuration file:

[PdfRepairSvr]
LogPath=EventLog

The system's application log event then logs messages similar as shown below:

- CreateService: The operation completed successfully.
- StartService: The operation completed successfully.

Note: The messages are only fully accessible while the service is created.

Otherwise, a message as shown below is displayed:

The description for Event ID (1) in Source (pdfrepairsvr) cannot be found. The local computer may not have the necessary registry information or message DLL files to display messages from a remote computer. The following information is part of the event: DeleteService: The operation completed successfully.

Polling interval

The polling interval defines the time in milliseconds that the polling thread pauses between two polls. The time passing until the same watched folder is polled again (maximum pick-up time) is: The value of PollingInterval plus the actual time it takes to poll all watched folders. The higher the polling interval, the lower the network traffic, and the longer it takes until documents are picked up.

Suggested values for the polling intervals are 1000 to 10000 milliseconds.

```
[PdfRepairSvr]
PollingInterval=5000
```

5.2.2 -a Analyze only

```
Analyze only -a
```

When using this option, the processed input files are only analyzed and a log file is generated. There is no output created.

5.2.3 -b Set the rebuild options

```
Set the rebuild options -b <n>
```

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

- 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. CFF fonts can be converted to Type1 fonts only if fonts are rebuilt, i.e. 4 can be used together with 2 only.

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

Example: Repair the file and rebuild all

-b 7

5.2.4 -dp Do not recover pages

Do not recover pages -dp

If pages are not part of the page tree (loose pages), they are recovered and added to the end of the document. If the option -dp is selected, loose pages are not recovered and are left out of the recovered document.

5.2.5 -dx Do not repair cross-reference table

Do not repair cross-reference table -dx

With this option, a corrupt XREF table is not repaired or recovered. This option is useful if it is taking too long to process a document, since repairing the cross-reference table is very time-consuming.

5.2.6 -p Read an encrypted PDF file

Read an encrypted PDF file -p <password>

A PDF document that has a user password (the password to open the document) can only be processed when either the user or the owner password is provided. The password can be provided using the option -p followed by the password.

Example: The input PDF document is encrypted with a user password. Either the user or the owner password of the input PDF is "mypassword". The command to process such an encrypted file is:

-p mypassword

When a PDF is encrypted with a user password and the password is not provided or is incorrect, the 3-Heights® PDF Analysis & Repair Service cannot read and process the file. Instead it generates the following error message:

Password wasn't correct.

5.2.7 -w Specify the path to the root directory

Specify the path to the root directory -w <dir>

This option sets the path to the root directory.

Note: This parameter must always be the first parameter of a thread.

Parameter:

<dir> The given path should not contain mapped drives, since other users (such as LocalSystem) do not recognize them.

Example:

-w C:\pdfrepairsvr\Root

The service supports path lengths including file name of up to 258 characters. This includes the 21 characters of the job ticket.

If a file name exceeds this value, its file name is truncated at the end of the file name and before the file extension. Therefore, watched folder names should be kept reasonably short.

5.2.8 -wd Specify the drop path

```
Specify the drop path -wd <dir>
```

This option sets the path of the drop directory. If this option is not set, then the drop directory equals the root directory -w. The drop directory is the watched folder, where input files are picked up and processed by the service.

Parameter:

dir> The drop directory can be at any existing (network-) location with the following conditions:

- The directory exists (it is not automatically created, unlike the root directory).
- The user under which the service runs has access permissions to this directory.

Example:

-wd C:\Path\DropFilesInHere

5.2.9 -wfi Ignore files with certain extensions

```
Ignore files with certain extensions -wfi <exts>
```

By default, the service tries to process all files dropped into the drop-in folder, regardless of the extension. With this option, files with certain file extensions can be ignored.

Example: Ignore temporary files.

-wfi .temp.tmp

5.2.10 -wfs Process only files with certain extensions

```
Process only files with certain extensions -wfs <exts>
```

By default, the service tries to process all files dropped into the drop directory, regardless of the extension. With this option, the processing can be restricted to a set of known file extensions.

Example: Restrict the processing to PDF and FDF files.

-wfs .pdf.fdf

6 Troubleshooting

6.1 Repair time takes too long

Try excluding the cross-reference table (switch $\underline{-dx}$). This speeds up the repair time.

6.2 File cannot be repaired

If the $\underline{-dx}$ switch has been selected and the cross-reference table is corrupt, the file cannot be repaired. Ensure the $\underline{-dx}$ switch is not set.

7 Version history

7.1 Changes in versions 6.19–6.27

Update license agreement to version 2.9

7.2 Changes in versions 6.13–6.18

No functional changes.

7.3 Changes in versions 6.1-6.12

No functional changes.

7.4 Changes in version 5

- Changed error reporting behavior: Errors in the XMP metadata are no longer reported when saving a recovered document.
- **New** additional supported operating system: Windows Server 2019.

7.5 Changes in version 4.12

New HTTP proxy setting in the GUI license manager.

7.6 Changes in version 4.11

- **New** support for reading and writing PDF 2.0 documents.
- **New** support for the creation of output files larger than 10GB (not PDF/A-1).
- New treatment of the DocumentID. In contrast to the InstanceID the DocumentID of the output document is inherited from the input document.

7.7 Changes in version 4.10

- Improved robustness against corrupt embedded font files.
- Improved robustness against corrupt input PDF documents.

7.8 Changes in version 4.9

- Improved support for recovering certain corruption types.
- Improved support for and robustness against corrupt input PDF documents.
- Improved repair of embedded font programs that are corrupt.
- New support for OpenType font collections in installed font collection.
- Improved metadata generation for standard PDF properties.

7.9 Changes in version 4.8

Added repair functionality for TrueType font programs whose glyphs are not ordered correctly.

8 Licensing, copyright, and contact

Pdftools (PDF Tools AG) is a world leader in PDF software, delivering reliable PDF products to international customers in all market segments.

Pdftools provides server-based software products designed specifically for developers, integrators, consultants, customizing specialists, and IT departments. Thousands of companies worldwide use our products directly and hundreds of thousands of users benefit from the technology indirectly via a global network of OEM partners. The tools can be easily embedded into application programs and are available for a multitude of operating system platforms.

Licensing and copyright The 3-Heights® PDF Analysis & Repair Service is copyrighted. This user manual is also copyright protected; It may be copied and distributed provided that it remains unchanged including the copyright notice.

Contact

PDF Tools AG
Brown-Boveri-Strasse 5
8050 Zürich
Switzerland
https://www.pdf-tools.com
pdfsales@pdf-tools.com