

# **PDF-Analyzer Pro 4.0 (c)2003-2010 by Ingo Schmoekel**

## **Program - Author :**

Ingo Schmoekel  
- Software-Dev.& Distribution -  
Zedernstr. 30a  
D-28832 Achim-Uesen  
G E R M A N Y

**Email:** [webmaster@pdf-analyzer.com](mailto:webmaster@pdf-analyzer.com)

**Web:** <http://www.pdf-analyzer.com>

**COPYRIGHT:** All in the documentation and in the program used companies, products and trademarks are subject usually to marks, patents or rademarks and belongs to the registered companies and/or individuals in each case!

**NOTE:** In the pro version the program works on your files. For the protection of your original files you should save these before the treatment! In principle you should save your files on a regular base.

## **Program description**

With the PDF-Analyzer you have an informationstool specially for pdf files. An indispensable tool for everyone who's working much with pdf files, it will be appreciate, if one - without starting large pdf-applications - e.g. in the Explorer (context menu) can see directly all worth knowing of a marked pdf file. That reaches from the document informations e.g. titles, topic affiliation and page layout or also pagecount of a file over security-relevant things such as data compression and copying permission or also encryption strength and password use up to the file attributes! Textextraction already included! It's a two-way-tool... You can use it from the explorer contextmenu and "stand alone" as a "PDF Browser", too. If one found a desired file, it can be started immediately over a linked reader (e.g. the Acrobat reader).

## **The main functions are**

The Pro-version offers the possibility of resetting (unknown) main passwords, modification for all relevant document-, security-, layout- and file-properties. There are further functions such as inserting/appending images and txt-files, merging up to 10 pdf-files together, extensive search functionality for the whole directory, fulltext-search, a comfortable watermark-/logo-function, extraction for javascript- and rdf-(xml)content, image conversion to pdf (bmp, jpg, gif, png, tif, wmf, emf), copying/extracting/deleting single or more pages of a document, popup-textboxes while opening a pdf-file, Link-integration (url/email), exporting the shown pdf-attributes as xml- and csv-data for whole directories, IndexCut for really fast web access, making your own notes inside the file...

You can insert additional files. It's a supported acrobat feature - you can transport this files with your pdf-file on a special attachment-page where the files later can be detached by the users.

## **What's possible and what not**

The function "Run" needs an installed PDF Reader. You get it for example (own download) for free from the adobe website or from one of these pc magazine cds. Normally the PDF-Analyzer uses the linked (with pdf-files) pdf-viewer... If you want to call a special viewer from PDF-Analyzer you have to create a file called pdfread.cfg (you can do it with notepad for example). This file contains a complete path with the name of the program-file for your special pdf-reader. If you want to use another pdf-reader you can use this file pdfread.cfg. The cfg-file must be copied in the free user-directory (you can see it on the last screen of the PDF-Analyzer). There's no need to use the Acrobat Reader - take the reader you like the most (i'm preferring the Foxit Reader).

The program works unlimited on pdf-files 'till 24 mbytes. On bigger files (we've tested 'till 262 mbytes it's sometimes a bit limited.

There isn't any guarantee that this application can analyze all documents but normally it should works with all pdf-documents, 40 and 128 bit, RC4 and AES.

If there's a user-password it seems that the file creator had the wish that his pdf file should be "unaffected" for ever - in this case therefore also no text-, rdf- and java-script-extraction is possible. Few of the special functions (for example watermark/logo) isn't possible if there's an unknown password. If it's an unknown main password you should reset it first with the new function of the PDF-Analyzer Pro.

If there's a copy protection on the files sometimes the javascript- and rdf-extraction has a poor quality.

Few functions (like the watermark-function) are working directly on the original pdf-files. If you want to keep an original one you should copy/save it first. Please read the help-text in the special functions!

## **Informations when adding text-files**

You can use all ascii-files with carriage returns / line feed (i mean fixed lines). The files need the suffix ".txt". The text-files will be automatic linked to the selected pdf-files. The automatic procedure works in the manner i'll explain now:

The text will be read in line by line and will be inserted line by line to the pdf-file. The font is courier (this font is well readable and good when using tables). The first text-line begins 50 pixels from the top (5 lines). 48 lines on each page will be used. Till 100 characters per line the pageformat is "A4 up". More characters means "A4 landscape".

Till 60 characters per line means fontsize 14. Till 80 means fontsize 12 and 100 Characters is size 10. If it's landscape then from 101 to 132 characters it's fontsize 14. More then 132 means size 10.

## **What's modifiable...**

### **D o c u m e n t d a t a :**

Titel, Theme, Author, made by, created with, keywords, pdf-version, personal notes.

### **S e c u r i t y d a t a :**

Allowed or not: Printing, copying, changing, formfields, comments, content-access, assembling documents.  
Main- and/or user-password, encryptionlevel.

### **L a y o u t d a t e n :**

Page-Layout, pagemode.

### **F i l e d a t a :**

Changed at (date/time), fileattributes.

### **O w n n o t e s . . . :**

In unprotected pdf-files you can insert and change your own notes. There's enough place for a few lines of text. Please keep in mind to hit again the enter-key after the last text-line. This function is available via the button [-Change-]...

## **How to use ...**

### **G e n e r a l l y :**

Apart from the mouse operation also keyboard operation is possible. The individual selection and input fields can be started successively with the Tab key. So also drives and directories can be selected and opened by the use of character keys with the enter key. The file listbox (down, on the left) indicates in principle only pdf files. Either a file can to be selected by a mouse-click or one can browse with the arrow-keys (up and down) through the files.

For safety and data security reasons no pdf files can be worked on, which are protected with a user password... Which also should be in your interest as a respectable user ;-)  
Working on files with a main password is often possible. If not you should reset the main password.

For a comfortable working the fields in the security section are automatically filled with the made settings from the last run. It's the same in the special functions! These suggestion values are overwriteable. They'll be saved in single files (\*.cfg) in the free user directory which will be created after the first start (...pdfanalyzer\...) of the PDF-Analyzer.

## **Context menu :**

Important switching surface functions and more can be released also over the context menu (right mouse button). Also the installation and uninstallation of the program into the context menu of the Windows Explorer can be initiated here. In the future then only a pdf file needs to be marked in the Explorer, the right mouse button ... "PDF-Analyzer Pro" ... and then you can see it all ...

Sometimes there can be few troubles if there's already the "normal" Acrobat Reader inserted. In this case please use the possibility to insert the PDF-Analyzer in the explorer context-menu under "send to ...".

## **In the context menu there're special functions...**

### **[-As a new starting file for a pdf-merge-].**

You want to merge few pdf-files. With this menuitem the actual selected pdf-file will be the first file when merging few files together... With the menuitem "Cancel the pdf-merge" you can stop the merge-function.

### **[-Add to the starting pdf-merge-file-].**

You want to merge few pdf-files. After selecting a starting file this function will be active. With this function you can append up to 10 selected pdf-files (file by file). With the menuitem "Cancel the pdf-merge" you can stop the merge-function.

### **[-Terminate the pdf-merge-].**

You want to merge few pdf-files. After all pdf-files are selected for the big merge with this function the real merge starts. The new file is named like the starting file and a "...-new" at the end. With the menuitem "Cancel the pdf-merge" you can stop the merge-function.

### **[-Cancel the pdf-merge-].**

With this function you can stop the merge-function. Perhaps when the wrong files were selected... or...

### **[-Removing layers-].**

A pdf-document consists of several layers. Layers are for better organization of graphical and text data in a document. You can use single layers for special images for example. Removing the image isn't possible but removing this layer with the image is possible with this function. The last inserted layer has number 0 and the oldest one has a higher one.

This function is a bit an experimental function... so the results should be saved with a new name.

### **[-Beginning with a new text- or imagefile-].**

When a pdf-file is selected you can insert in an automatic way a new first page with an image or the content of a textfile.

### **[-Appending a text- or imagefile-].**

When a pdf-file is selected you can insert in an automatic way a new last page with an image or the content of a textfile.

### **[-Working with document page(s)-]**

This functions are working directly on the selected pdf file. For your information there is the pagecount (top/right) and the page width in pixel (bottom). If the original file will be

changed you can use a new filename in the save dialog... So you can keep the original file! You can use only one function from this page at a time! In the first (top) textfield you see the selected pdf file. Aside there is the pagecount.

**>> Function:** Delete from page ... to ...

I think you know what this mean.

**>> Function:** Separate after page ...

In the same directory you can generate two new files (parts of the original file) with the suffix ...-t1.pdf and ...-t2.pdf.

**>> Function:** Copying this page(s) in a new file...

Here you can insert nearly all you want. For example ... 1-3, 4, 8, 15-20, 34 ... or only ... 5 ... is possible.

**>> Function:** Complete separation into single pages ...

In the same directory a sub directory with the name of the pdf-file will be created automatically and each separated page will be inserted as a single ...-pagenumber.pdf.

**>> Function:** Overwrite page footer / page header ...

With this function you can overwrite the original page ranges at the top/bottom with a white area. In a second step you can use a new pagenummering ... or perhaps you want to insert new (your own) text there ...

**>> Function:** New page numbers top ... / down ...

All following settings impinge on the top or on the bottom of the pages. The font height for the pagenumbers should be in the range 6-12. You can adjust from which page the numbers shall begin - If your document has a title page it could be better to begin with page nummber -1- on the real second page. If beside "after the page number follows" there is only a "from" or "/" then later on your document pages you will find something like "Page x from xxx" or "Page x / xxx". You can work very accurate with the horizontal declaration. All made settings will be suggestion values if you use this function the next time/again!

### **[-Linkintegration-]**

This function is working directly on the selected pdf file. You can integrate an external link. This can be an url like (<http://www...>) or an email-address like ([webmaster@ForExample.com](mailto:webmaster@ForExample.com)).

For your orientation you can see the selected pdf-file in the text-field at the top.

In the next line you can insert the link:

If you want to use a website-link the correct string is like...

<http://www.forexample.com>

If you want to use an email-address the correct string is like...

<mailto:webmaster@forexample.com>

If the user should not see the real link you can use an alternative text instead of the link. For example not "mailto:web..." but "An email for the webmaster"...

In the next line you can choose the link-position on the document pages.

Beneath you can select the pages for link integration.

Use the next line if you want to work with special single pages.

The pages must be separated by a comma. You can use this line or the line before - Not both.

You can distinguish the link using a frame around it.

A coloured link-text is feasible, too. Keep in mind that an activated link frame will stay in black anymore - So using coloured link texts you should better work without frames.

The settings will work directly on the selected, original file.  
The original file will be changed! No backups will be made! All settings will be suggestion values if you use this function the next time/again!

### **[-Integration of popup-notes-]**

This function is working directly on the selected pdf file. Here you can create an additional popup-window which opens when the pdf-file is open...

The made settings will work directly on the selected, original file. The original file will be changed! No backups will be made! You cannot remove the popup-text later!  
All made settings will be suggestion values if you use this function the next time/again!

### **[-Viewer properties for the document-]**

This function is working directly on the selected pdf file. Here you can set special options in which way the pdf-file should be displayed. Perhaps without any menus?!

Please keep in mind: If you have documents with more than one page it is better NOT to deactivate the windows-user-interface because you do not have a better protection but a poor handling! If you want to undo settings remove the hooks in the relevant boxes. This works for the most functions...

The made settings will work directly on the selected, original file. The original Not all options work together - have a try first with example files! Files will be changed! No backups will be made! All made settings will be suggestion values if you use this function the next time/again!

### **[-Image to pdf-]**

Automatically pdf-generations after selecting an image-file. You can work with the types jpg, tif, png, gif, bmp, wmf and emf.

### **[-Add a file as attachment-]**

Here you can select different files in a file dialogue which can be inserted in the selected pdf file. Later you can see and detach these files on the attachment page of the pdf file. You can attach (for example) images, exe-files, text files, and so on. It's an acrobat feature we'll use here.

### **[-IndexCut-]**

This function is working directly on the selected pdf file. Here you can split large pdf-files into an index- and a content-file. First you have to set the number of pages for the index-file. The new filename will be the original name with an "-idx" at the end. Together with the content-file you can upload it to your webspace. You can now load the index-file very fast and you can get first impressions about the content of the main document - So there is no need for large and time-consuming downloads. Now the user knows immediately (without large downloads) if a large document will feed their needs or not! IndexCut is the best when you want to work on ebooks or catalogues! On choosed positions on the pages of the index-file you can click on links to open the large main-document...

[ The selected pdf-file ]

For your orientation you can see the selected pdf-file in the text-field at the top.

[ This link will be used ]

This link will be inserted in the index-file to open the main document after a mouse-click.

[ The displayed text ]

If you have a very cryptical link-name it is better to use an alternative text for example like "A click and you will learn more ..." or "One click to the whole content...".

[ The name of the new index-file ]

Here you can see the new generated index-filename. If you want to use this solution "IndexCut" it is necessary to load this file together with the original file up to the same weblocation or the same directory.

[ Position of the content-links on the index-pages ]

Where you want the links to be displayed? At the top (left, centered, right) or at the bottom (left, centered, right)? Here you can determine this.

[ Frame ]

Do you want a frame around the link-text?

[ Colour of the link-text ]

You can use colours for the link-text! Keep in mind that a used frame will stay in black...

[ Separation after ... pages ]

Here you can determine the number of pages for the index-file. Perhaps only one page with the title picture of the main document? ...Or 2, 3 pages with the index? ...

All made settings will be suggestion values if you use this function the next time/again!

### **[-CSV-directory export-]**

You can export the pdf-attributes/-properties of all pdf-files in the actual directory as one csv-file in the automatically generated subdirectory under ...\`xml`export.

As an option you can select to check all subdirectories, too.

An example: The attributes from pdf-files in C:\PDFFILES should be exported as csv-file into c:\pdfanalyzer\csvexp\pdfa.csv

With this function you can create complete csv-archives from your pdf-fileattributes.

### **[-XML-directory export-]**

You can export the pdf-attributes/-properties of the files as xml-file from the selected directory in the automatically generated subdirectory under ...\`xml`export.

As an option you can select to check all subdirectories, too.

Example: The attributes from all pdf-files in C:\PDFFILES should be exported as xml-file into c:\pdfanalyzer\xmlexport\c\_pdffiles.xml.

With this function you can create complete xml-archives from your pdf-files. If you want a single export please use "XMLsingle" from the main program window.

### **[-Watermark/Logo-].**

This function is working directly on the selected pdf file.

Additional settings like image-file-selection, force of the transparency in percent, image position on the page and which pages shall get an image are possible.

The selected pdf file

Here you see the selected pdf-file. For your orientation there is the document-width and -height (in pixels) on the right side. It make no sense to work with a big image- and a small pdf-file!

Selecting the image file/text for the watermark/logo With the button -SELECT- you can select an image-file for the watermark/logo. On the right side you can see the height and width (in pixels). You can use file-types like bmp, jpg, png, tif and gif. If you do not want to use an image you can work with own text in the line beneath. Using a textstring the alignment fullpage is not possible!

Transparency . . .

If you want to use the image or text as a logo, you should work with a percent value next to zeroe. For watermarks you should use values starting at 50 percent.

Insert on which pages . . .

I think you know what this mean. If you want to work on special single pages of the document you can insert them in the next line (on which pages...). The pages must be separated by a comma. You can use this line or the line above. Not both!

The alignment . . .

I think you know what this mean.

All made settings will be suggestion values if you use this function the next time/again!

#### **[-Show RDF-content-].**

With this function you can see all embedded rdf-content. Rdf-content is for example xml-code... It will be extracted from the selected pdf-file and you can see it in your editor-window.

#### **[-Show JavaScript-content-].**

With this function you can see all embedded javascript-content. It will be extracted from the selected pdf-file and you can see it in your editor-window.

#### **[-Refresh the filelist-window-].**

When you've initiated a searching there will be only few pdf-files (the result of the searching) in the filelist-window... If you want to see all pdf-files again you can use this function.

## **T h e b u t t o n s ...**

#### **[-Show/Run-].**

Hereby the selected file could be seen with a linked pdf reader (in general that will be probably the Acrobat Reader). If you want to use another special pdf-reader you have to insert a file named pdfread.cfg in the free, local user directory "...pdfanalyzer\..." which will be created with the first start of the PDF-Analyzer. This cfg-file contains the new pdf-reader-file with its complete path and drive. An example for this text-line could be:  
"C:\Program Files\Adobe\Acrobat 7.0\Reader\AcroRd32.exe" or perhaps  
"c:\program files\Foxit Software\Foxit Reader\Foxit Reader.exe".

#### **[-Print-].**

Hereby the selected file could be print out immediately in an automatic and optimized way without any pdf reader installed.

#### **[-Search-].**

You can examine all pdf-files in the actual selected directory relating to nearly all the data you can see on the program-screen. With a klick on "Search" the button notation change to "Start".

After entering all the search-data you can click on "Start" and the Analyzer begins searching the whole directory... After this in the filelist-window you can see only the matching files. If you want to see all pdf-files again - a right mouseclick and "Refresh..." - That's it!

You can search for...:

Title, theme(subject), author, made by, created with, keywords, pdf-version, created at, modified at, number of pages, images (or not), fast web access, colourmodel, number of layers, page layout, page mode, changed at (the real file), fileattributes, encryption level, javascript- and/or rdf-content (or not), main- and/or user-password, is copying/modifying/inserting comments /using formfields allowed (or not).

You can search in your "own notes", too!

If there're string fields you can work/search with substrings, too! For example if there's the subject "animals in the forest" you can search with "animal"...

If you've the smaller resulting list you can go on searching with new keys...

### **[-Change-].**

With this function you can change nearly all file-attributes (document-, security- and file-data) seeing on the program-screen.

You don't need additional special pdf-software to do this! You can even enter main- or user-passwords to lock or unlock the file!

If there is a password on a file you can't change this file. First you must give in the correct password to unlock the file. If you don't want to change anything - click on another file and the change-mode stops and you are again in the browse-mode. Wherever it's possible you can use values from fixed selections. A new date you can get in a comfortable manner from a calendar-module.

If you want to work in the same way on a lot of files (one after the other)... you don't have to insert everytime the same data for securitydata, encryption-level or passwords. The program keeps the last settings in mind using them the next time as suggestion values. If you want to protect your files from copying (content) and printing using a main password the settings have to be inserted only one time!

For the next files you have to click only three times:

Selecting the file -> button "Change" -> button "Save" -> READY!

You can modify the following attributes/properties:

Title, theme/subject, author, made with, created by, keywords, pdf-version, pagelayout, pagemode.

Allowed or not: Printing, copying, inserting comments, using formfields. Encryption level, main- and/or user-password. The file-properties "Modified at" (date/time) and the file-attribute (archive, read only, ...). Your own notes can be changed, too.

### **[-Save-].**

After using the button [-Change-] the program-screen switches to the change-mode and as a logical consequence the formerly change-button has now a save-functionality. Pushing this button will save all actual made settings in the selected pdf-file. After this the save-button changes again to the change-function.

### **[-Preview-].**

Not only a preview but a small included pdf-viewer with browse- (page back, page forward, to the first page, to the last page) and

zoom-functionality. The last settings here are still known for the next files as long as the application is running.

**[-Textcontent-].**

Text contents of the current file will be extracted and shown in (with txt files connected) an editor.

**[-Textanalysis-].**

The text-content of the selected pdf-file will be extracted and examined. All words, characters, letters, punctuations and key-hits will be counted.

**[-TextSearch-].**

The text-content of all pages of all pdf-files in the selected directory will be scanned for up to three searchkeys.

**[-XMLsingle-].**

You can export the attributes/properties of the selected pdf-file in the automatically generated directory ...\\pdfanalyzer\\xmlexport (below the free windows userdirectory) as a xml-file.

## **Good to know while working on pdf files**

There are few rules, which must be considered during processing of pdf files (not only with this program)...

- 1.** If it's possible the 128-bit-encryptionlevel should be avoided, since this is only supported starting from the Acrobat reader 5.0 and users of older reader versions couldn't use this files.
- 2.** If you put pdf files on a website or on a cd rom, there should be also a download link to Adobe to download and install the free acrobat reader if it is already missing.
- 3.** Some extended pdf properties requires 18-bit-encryption. It's up to the user to decide if the advantage of the extended pdf properties are better then the disadvantage of the necessary newer Acrobat reader version.
- 4.** The assignment of a main password protects the made pdf properties securely from changes.
- 5.** The assignment of a user password requires the input of a user-password while opening a pdf file - with the main password the case is not. If there's only a user password set, the user value stands in it as a main password, too.
- 6.** The input of main password is sufficient to unlock a pdf - you don't need the user-password!)
- 7.** If the file should be able to be used unchangeable (without the PDF-Analyzer Pro) by everyone, then you should always use a main password!

## **In the directory of PDF-Analyzer Pro you can find ...**

**- pdfanalyzer.exe**

The program file.

**- pdfaverz.cfg**

This file contains the last directory where PDF-Analyzer Pro was working.

**- pdfread.cfg**

This file contains a complete path with the name of the program-file for a special pdf-reader. If you want to use another pdf-reader than

your standard reader you can use this file pdfread.cfg. This cfg-file should be in the free windows-user-directory (if you don't know it please look on the forth screen of PDF-Analyzer).

**- pdfapi.cfg**

This file makes sure that the program starts with the same data-section, which was open last time.

**- howtofaq.pdf**

A small how-to and faq-file with main and important questions and answers while and for using PDF-Analyzer.

**- getting\_started.pdf**

For fast success... Description of single steps from main functionalities.

**- pdfanalyzer.pdf**

The file you read just now in pdf mode.

**- pdfanalyzer.txt**

The file you read just now in text mode.

**- proversion.pdf**

Here you can read about the additional functionality in the pro-version and how to order it.

While working with the PDF-Analyzer you'll get additional cfg-files in your app-directory. These cfg-files contain additional preset-data for working with the PDF-Analyzer.

## Conditions for your system

CPU	Starting from Pentium upward
OS	Win95/98/Me/2000/XP/Vista/Win7
RAM	Starting from 32 mb (Win95)/ 64 mb (WinNT)/ 128 mb (XP)/ 512 mb (Vista)/ 128 mb (Win7)
HD	Starting from 3,0 MBytes
Program size	About 2,5 MBytes and additional files.

## The author

Ingo Schmoekel  
- Software-Dev.& Distribution -  
Zedernstr.30a  
D-28832 Achim-Uesen  
G E R M A N Y

**eMail:** [webmaster@pdf-analyzer.com](mailto:webmaster@pdf-analyzer.com)

**URL:** <http://www.pdf-analyzer.com>

- >> Member of SAve (S hardware A utoren VE reinigung)  
<http://www.s-a-ve.com>
- >> Member of OISV (Organization of Independent Software Vendors)  
<http://www.oisv.com>
- >> Member of XING/openBC (XING/Open Business Club)  
<http://www.xing.de>