



PDF Compressor Manual

Foxit Europe GmbH



PDF Compressor Manual

Foxit Europe GmbH



Table of Contents

| | |
|--|----|
| 1. Introduction | 1 |
| 2. Installation | 2 |
| System Requirements | 2 |
| Requirements of the Born Digital component | 2 |
| Network Requirements | 4 |
| Location of Temporary Files | 4 |
| 3. Concept of PDF Compressor | 5 |
| Graphical User Interface | 5 |
| Job List | 5 |
| Job Entry and its Properties | 6 |
| Default Properties | 6 |
| PDF Compressor Service | 6 |
| Born Digital Conversion | 6 |
| 4. Getting Started | 8 |
| Starting the GUI | 8 |
| Migration from CVISION Products | 8 |
| Adding Jobs to and Removing Them from the List | 9 |
| Running Jobs | 10 |
| 5. Setting up Job Entries | 11 |
| Configuring Input Data | 11 |
| FTP Input Options | 15 |
| Input Filter Options | 17 |
| Configuring Advanced Input Options | 20 |
| Configuring Data Output | 25 |
| Configuring E-Mail Conversion Options | 29 |
| Configuring Advanced Output Options | 31 |
| Configuring Image Preprocessing | 39 |
| Configuring OCR | 40 |
| Advanced OCR Options | 44 |
| Configuring Barcode Recognition | 47 |
| Barcode Profile Options | 47 |
| Configuring Document Compression | 50 |
| Configuring Advanced Compression Options | 51 |
| Configuring File and Data Embedding | 54 |
| Options for Embedding XMP Metadata | 55 |
| Options for Embedding Bookmarks | 56 |
| Options for Embedding XML OCR results | 57 |
| Options for File Embedding | 58 |
| Configuring Header and Footer | 60 |
| Embedding Watermarks | 62 |
| Configuring Post-Processing | 63 |
| Actions upon Successful Processing | 64 |
| Actions upon Processing Failure | 65 |
| Post-processing <i>command execution</i> | 66 |
| Configuring General Job Properties | 69 |
| Job List Processing Order | 70 |
| Priority Processing Order | 70 |
| List Processing Order | 70 |
| Time Scheduled Processing and Processing Timeout | 71 |
| Configuring Advanced Job Options | 71 |



| | |
|---|-----|
| Audit Log | 72 |
| Setting up the Default Properties | 73 |
| Template String Syntax Description | 73 |
| Escape Sequences | 74 |
| Modifiers | 75 |
| Modifier Examples | 76 |
| Examples | 76 |
| Conditional Substitution | 76 |
| Examples | 77 |
| Regular Expression Substitution | 77 |
| Regular Expression Substitution Examples | 78 |
| 6. Managing the Job List | 80 |
| Adding, Deleting and Copying Entries | 80 |
| Changing the Sequence of Job Entries | 81 |
| Starting and Stopping Job Entries | 81 |
| Individual Entries | 81 |
| All Entries | 81 |
| Monitoring Jobs | 81 |
| Importing and Exporting Job settings | 82 |
| 7. Automatic Job List Processing | 83 |
| Finding and Processing Job Lists | 83 |
| Job List File Order and Priorities | 84 |
| Job List File and Control File Syntax | 84 |
| 8. Submitting Jobs via the PDF Compressor API | 87 |
| API Job Lifecycle | 88 |
| The C/C++ API | 88 |
| The .NET API | 88 |
| API Demo Applications | 89 |
| 9. Administration and Licensing | 90 |
| Online Services and Data Protection | 90 |
| General Configuration Settings | 91 |
| Log Files | 91 |
| Log File Analysis | 92 |
| Online Check for Updates | 93 |
| Job List Processing | 93 |
| Priority Processing | 94 |
| Language | 94 |
| Managing Licenses | 94 |
| Trial Mode | 95 |
| License | 95 |
| Caveats | 97 |
| Updating Licenses | 97 |
| Moving Licenses | 97 |
| License Monitoring | 97 |
| Managing the PDF Compressor Service | 101 |
| Required Privileges | 102 |
| Choosing the Service Account | 102 |
| Setting up the Service Priority | 103 |
| Taskbar Status Icon | 103 |
| CPU Configuration | 103 |
| Born Digital Options | 105 |
| Miscellaneous | 106 |



| | |
|-------------------------------|-----|
| 10. Tips and Tricks | 108 |
| 11. Troubleshooting | 109 |
| Born Digital Conversion | 110 |
| 12. License Agreement | 112 |
| 13. Support and Contact | 119 |



Chapter 1. Introduction

PDF Compressor is an industrial strength solution for automated document conversion. Through its scalability it is suited for processing small quantities of data as well as large quantities.

The PDF Compressor compresses and converts scanned documents to PDF or PDF/A. Using the Born Digital component further digital input formats such as MS Office documents, PDF and e-mails in MSG or EML format including attachments can be converted to PDF/A.

The outstanding document compression provided by the PDF Compressor greatly reduces file sizes while ensuring prime document quality in an ISO standard compliant format suitable for long-term archiving.

The fully integrated text recognition engine (OCR) provides full-text searchable PDF compression results. Created documents can automatically be handed over to downstream systems for various further business processes.

PDF Compressor can generate PDF files for different requirements, e.g. without using PDF features in order to be compatible archivable documents that meet the PDF/A standard (PDF/A-1, PDF/A-2, PDF/A-3).

PDF Compressor offers the following features:

- Scalable high-quality compression of color document images
- Applying OCR to create full-text searchable PDFs from scanned pages
- Conversion from PDF to PDF/A
- Optimization of images or scanned pages inside existing PDF documents
- Merging, splitting and renaming of documents
- 1D-/2D-Barcode recognition: splitting, renaming output files or adding PDF bookmark to output files, triggered by barcodes
- Combine digital documents with scanned pages, e.g. attaching testimony and certificates to an online application
- Unified conversion of digitally created and scanned documents to PDF/A for long term archiving
- Automatic processing of input files in watched folders
- Time-triggered starting and stopping of jobs
- Fully automated processing in Job List Processing mode



Chapter 2. Installation

Run the installer and follow the instructions. The installer is typically named PDFC_<xxx>_v<N.N.NN.NNN>.exe, where xxx corresponds to the system architecture "x64" and N.N.NN.NNN corresponds to the product version number, e.g. 8.0.2.5.

Note: You need Administrator privileges to install PDF Compressor. After an initial installation PDF Compressor runs in evaluation mode. In this mode a Foxit Europe logo is added to each output page.

System Requirements

| | |
|-------------------|--|
| Operating System: | Windows 64 Bit version of Win 7 / 8.1 / 10 / Windows Server 2008 R2 / 2012 / 2016 |
| Processor(CPU): | Intel / AMD or compatible x64 processors, single core and multi core processors supported |
| Processor Speed: | Minimum 1.6 GHz, 2 GHz or more recommended |
| Memory: | 4 GB per core recommended, at least 6 GB per core for conversion of very large documents on 64 bit systems |
| Disk space: | 1 GB required for installation |
| Misc: | Microsoft .NET Framework 4.6.2 is required and will automatically installed if not already available. |

Note: These are minimum installation requirements. Required resources for production environments should be clarified with Foxit Europe.

Requirements of the Born Digital component

In addition to the general system requirements, the following conditions must be met for Born Digital document conversion of the PDF Compressor.

To convert PDF to PDF/A no additional preparations are necessary.

For successful conversion of other born digital file formats, the following conditions must be met:

1. A 64 Bit version of Windows 7 / 8.1 / 10 / Windows Server 2008 R2 / Server 2012 / Server 2016
2. The necessary Office applications have to be installed.
 - For conversion of Microsoft Office documents the appropriate components of Office 2007, 2010, 2013, 2016 or 2019 must be installed such that all Office document types submitted to conversion can be opened.



- Ensure that the Microsoft Office applications can save as PDF. A printer driver must be installed on the system to convert Excel files to PDF. With Office 2007 prior to SP2 the 'Microsoft Save as PDF or XPS' Add-In must be installed.
- For conversion of ODF documents (Open Document Format) the appropriate components of OpenOffice.org 3.2 or newer or LibreOffice must be installed.

Note: We strongly recommend using the Born Digital document conversion on a dedicated server, without users interactively using the Office applications on the same system, since intermittent or simultaneous interactive use is likely to cause errors.

While Born Digital conversion is performed, the account chosen for the PDF Compressor service may not be logged in on the server.

3. The service "PDF Compressor" must run under a dedicated user account. During installation of the Born Digital conversion option the *PDF Compressor Configuration Tool* is started to set this user account and associated permissions.
 - The default account "Local System" cannot be used to convert Microsoft Office documents.
 - The account can be a local or a domain user account, but it should have administrator privileges on the computer.
 - The account must have sufficient permissions to access the actual files to convert and to access the office applications via the (D)COM interface.
 - Use the button *Configure system for office conversion* under *File* → *Options* → *Conversion Options* to change the account settings and necessary system configurations after installation.
4. Manual preparation before first automatic conversion:
 - Log in with the dedicated user account (chosen for the service) and start each of the Office applications used for conversion.
 - Verify that all required input file types can be opened properly in Microsoft Office and/or OpenOffice.org. This procedure is required to complete the Office installation and to ensure its integrity.
 - Test the "Save as PDF" feature for each Microsoft Office application.
 - Ensure that all startup dialogs requiring user input are deactivated.
 - We recommend disabling macro execution because of security concerns and because it could stall the processing. In the Office application options, the trust center settings you should disable macros and other active elements (ActiveX) without notification.
 - Automatic updates of references, fields and external content during document opening or printing must be deactivated.
 - For some conversion tasks it is necessary to have a default printer installed.



Network Requirements

PDF Compressor can run as a stand-alone application without any network, but typically it will access net shares by reading and/or writing files. The system administrator must ensure that the account used to run the PDF Compressor service has the corresponding access rights. See *"Managing the PDF Compressor Service"* (p. 101) for more details.

Location of Temporary Files

PDF Compressor may store temporary files during processing. These files will be located in a subfolder "FX_PDF_Compressor" in the temporary folder of the service account. The path of the temporary folder is specified by the following environment variables: TMP, TEMP or USERPROFILE. If none of these variables is specified for the user or system, the Windows directory is used instead. Alternatively the environment variable FX_PDFCOMP_TMP may be used to specify a temporary folder for PDF Compressor.



Chapter 3. Concept of PDF Compressor

The PDF Compressor is designed to convert files automatically without any user interaction. The PDF Compressor, after being configured, will run in the background. You do not need to be logged in on the system any longer. Even after a shutdown and reboot of your system, PDF Compressor will continue to process its jobs.

This functionality results from PDF Compressor architecture with two basic components:

- A graphical user interface (GUI) that allows you to configure and monitor the system
- An underlying Windows service which performs the conversion jobs

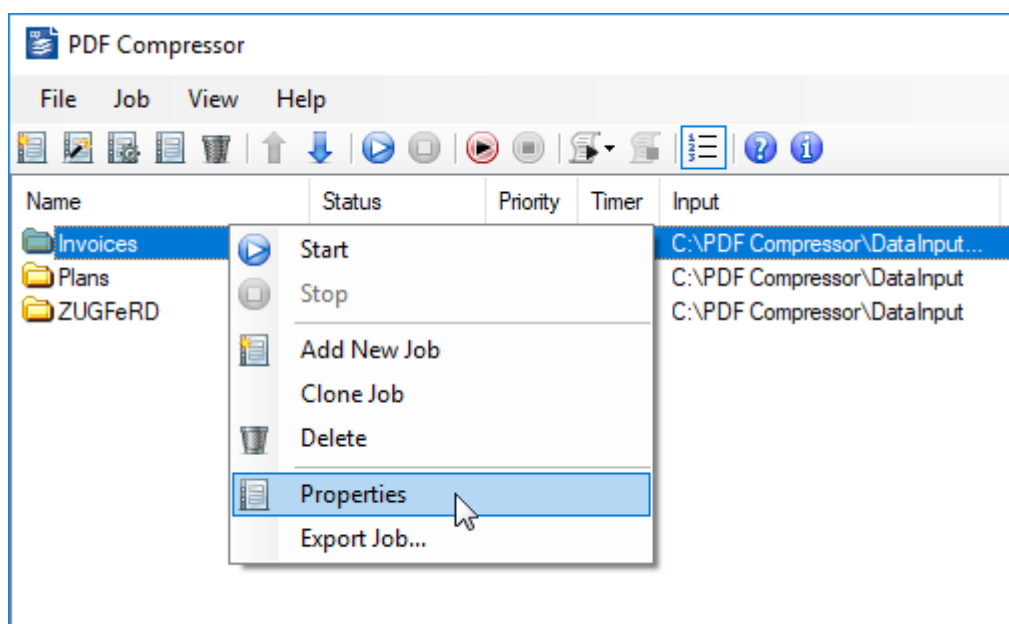
However, the GUI can be used as any ordinary Windows based application without even noticing the underlying Windows service.

Graphical User Interface

Job List

The job list is the main element of the GUI. When you first start the PDF Compressor you will see an empty list. Before you can do anything, you need to add an entry to this list. This can be done by creating a new entry or by simply dragging & dropping files or folders onto the list window (compare *“Adding Jobs to and Removing Them from the List”* (p. 9)).

An entry has many properties. One of them is the input file or folder you want to process, others include all the parameters that can be configured for the file conversion. Since you have a list with a virtually unlimited number of entries, you can setup different types of file conversions, each with its own, independent set of parameters.



You can configure entries to access different net shares within your network, or you might use them to create different versions of PDF files (high vs. low resolution, high quality vs. fast OCR, etc.) from the same input files.



The complete job list can be stored to a file or restored from a file to enable backup functionality or the management of completely different setups.

Job Entry and its Properties

Within each entry you define its properties:

- what kind of input is accessed (a file, all files within a folder, what file types, etc.),
- how it is accessed (once or frequently to make a folder a *hot folder*),
- where to put the output,
- what to do in case of an error,
- which additional data (metadata) to add,
- how to compress the data (quality, resolution, PDF compatibility, etc.).

Each entry can be started and stopped individually allowing you to run individual file conversions to your needs.

Default Properties

To simplify and fast-track creation of new job entries, PDF Compressor maintains a set of *Default Properties*. Whenever you create a new entry the settings of the default entry will be used as a template. Setting up this configuration accordingly you will need to adapt only very few settings for newly created jobs.

With a properly chosen set of default properties, converting files to PDF is achieved by simply dragging & dropping files onto the list and pressing the start button.

PDF Compressor Service

Normally you will not notice the underlying Windows service. It just processes the files you setup within the GUI. Since the GUI is independent from the service, you can configure and start one or more job entries, close the GUI and log off if you want. When you open the GUI again, you will be informed of the progress of your running jobs.

Born Digital Conversion

The PDF Compressor comprises a component for conversion of *Born Digital* files, such as Microsoft Office, OpenOffice documents and e-mail messages. The Born Digital component provides following types of conversions:

1. Microsoft Office, OpenOffice, RTF, HTML and plain text documents are converted to PDF or PDF/A. The distinction between these input formats is specified via the options on the *Input* tab of the *Properties* dialog (see “*Configuring Input Data*” (p. 11)).

The exact list of format types depends on your installation and configuration of the MS Office and OpenOffice/LibreOffice applications. Generally supported applications and formats include Microsoft Word, Excel, PowerPoint and Outlook as well as OpenOffice Writer, Calc and Impress.




2. PDF Compressor also supports the conversion of e-mail messages saved in the EML and MSG formats. Attachments of the supported input file types are extracted and converted, too.
3. PDF documents can be converted to PDF or PDF/A. Again, the distinction between these output formats is specified via the options on the *Output* tab of the *Properties* dialog (see “*Configuring Data Output*” (p. 25)). Note that the conversion of PDF does not use the MS Office or OpenOffice applications.

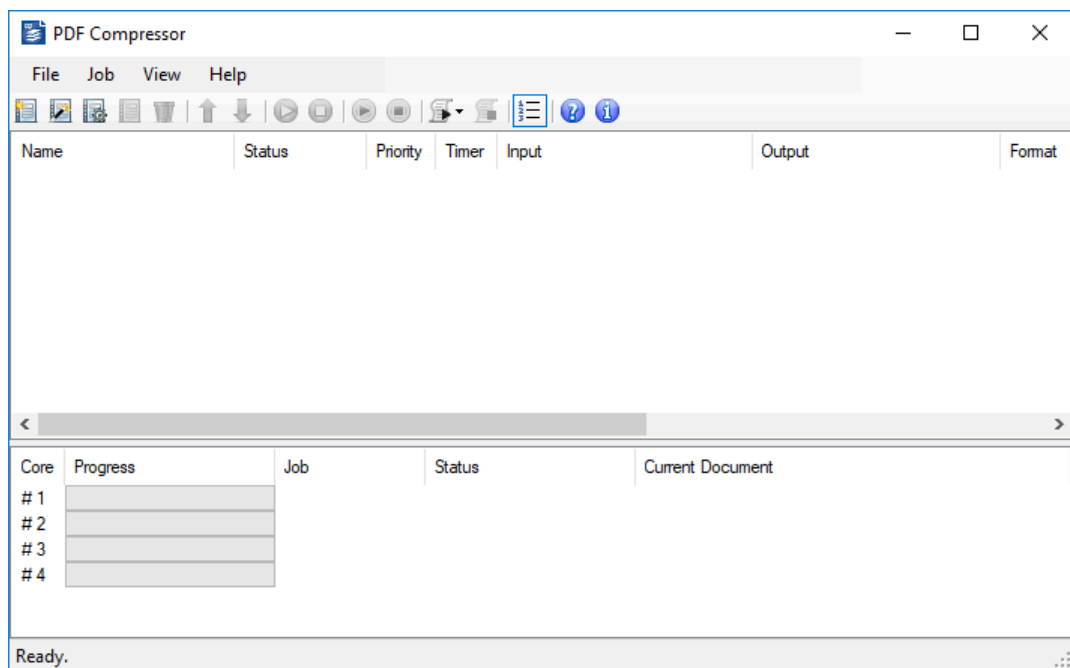


Chapter 4. Getting Started

This section describes basic steps to convert scanned document files to PDF files. More details can be found in Chapter 5, *Setting up Job Entries* (p. 11).

Starting the GUI

To start the GUI, either double click the PDF Compressor taskbar icon  or use the Windows start menu entry *Foxit → PDF Compressor → PDF Compressor*. When you start the GUI for the first time you will see an empty list:



Migration from CVISION Products

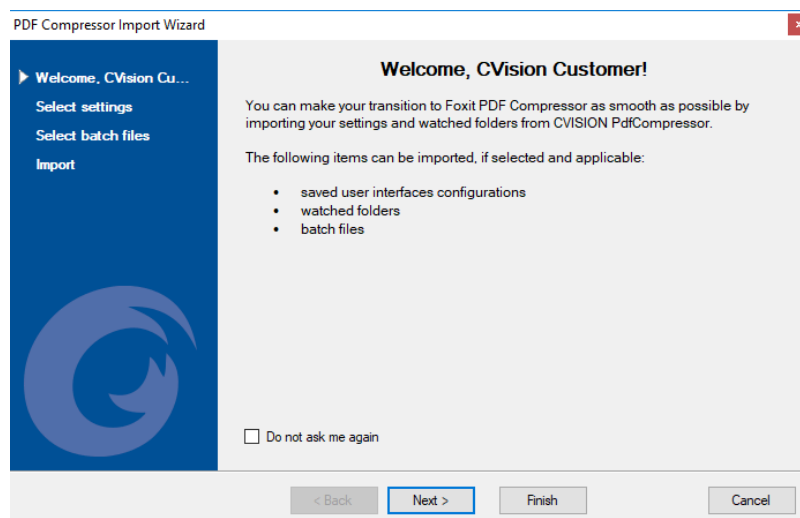
Foxit PDF Compressor allows you to import baseline settings, watch folder settings and batch files from the following CVISION products:

- PdfCompressor
- Maestro Recognition Server


If such settings are detected, the import wizard will automatically be opened each time the Foxit PDF Compressor GUI is started. After you have performed an import once or select "Do not ask me again", it will not open automatically again.

You can start the wizard at any time from *File → Import CVISION Settings*.

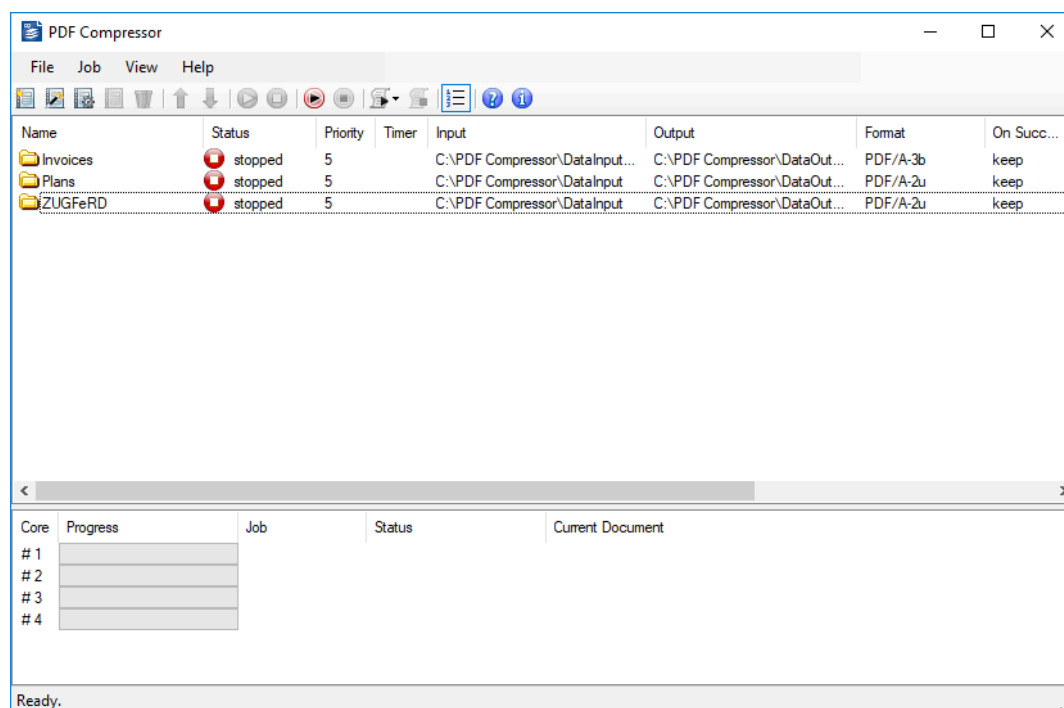
The import of the existing settings is also possible without having these products still installed.




Adding Jobs to and Removing Them from the List

There are several ways to add new jobs: By using the tool bar, the menu bar, or the context menu. Use the *Job Wizard*  to get a step-by-step introduction through the most basic job settings.

A very quick way to add a new job to the list is by opening the Windows Explorer and dragging & dropping a file or a folder onto the list window. The new entry will be set up to convert the given file (or all files within the given folder) to PDF/A. The output will be placed next to the corresponding input file ¹.





Your job list is preserved without the need to save it manually. Thus, even if you close and re-open the GUI, you find your job list unaltered. To remove an entry from the list, select it and press the delete button  on the toolbar.

¹Changing the default properties can change this behavior. The explanations given here are related to the *Factory Default* settings.



Running Jobs

There are several ways to start jobs represented by the list entries. For example, you may select the entry you want to start and press the start button  on the toolbar. The status of the entry will change from *stopped* via *starting over running* to *working*, and back to *stopped* when the job is finished.

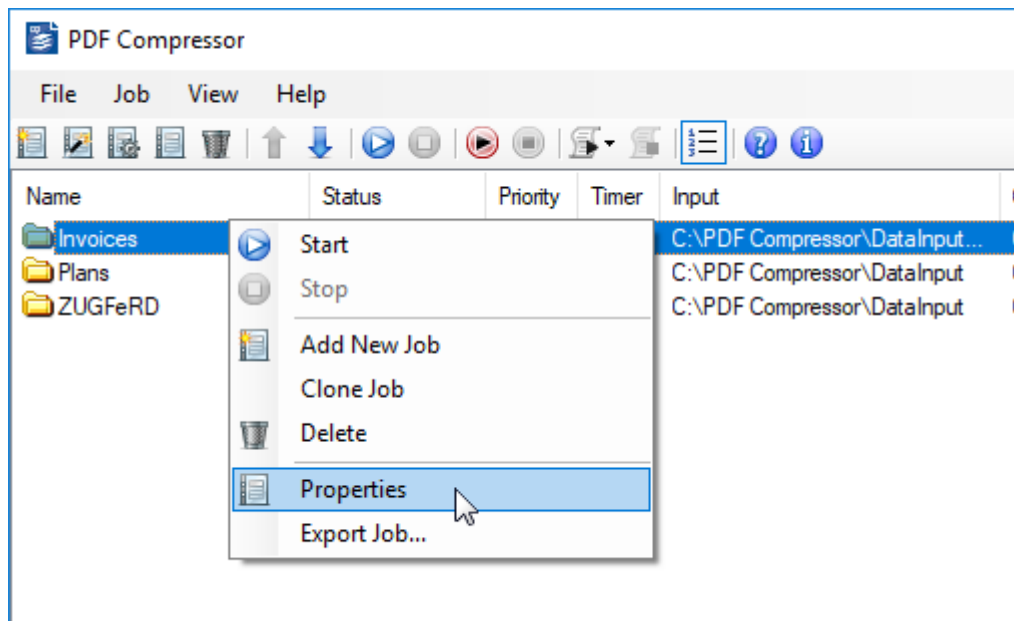
While the job is running, you can abort it by pressing the stop button  on the toolbar. It may take some seconds until the conversion has been aborted. Any incomplete PDF output file will be removed.




Chapter 5. Setting up Job Entries

You can edit properties of a job entry once it has been added to the list (see “*Adding Jobs to and Removing Them from the List*” (p. 9)). The properties dialog window can be opened by any of the following actions:

- Double-click the corresponding entry within the list,
- Use the right-click context menu associated with the list entry,

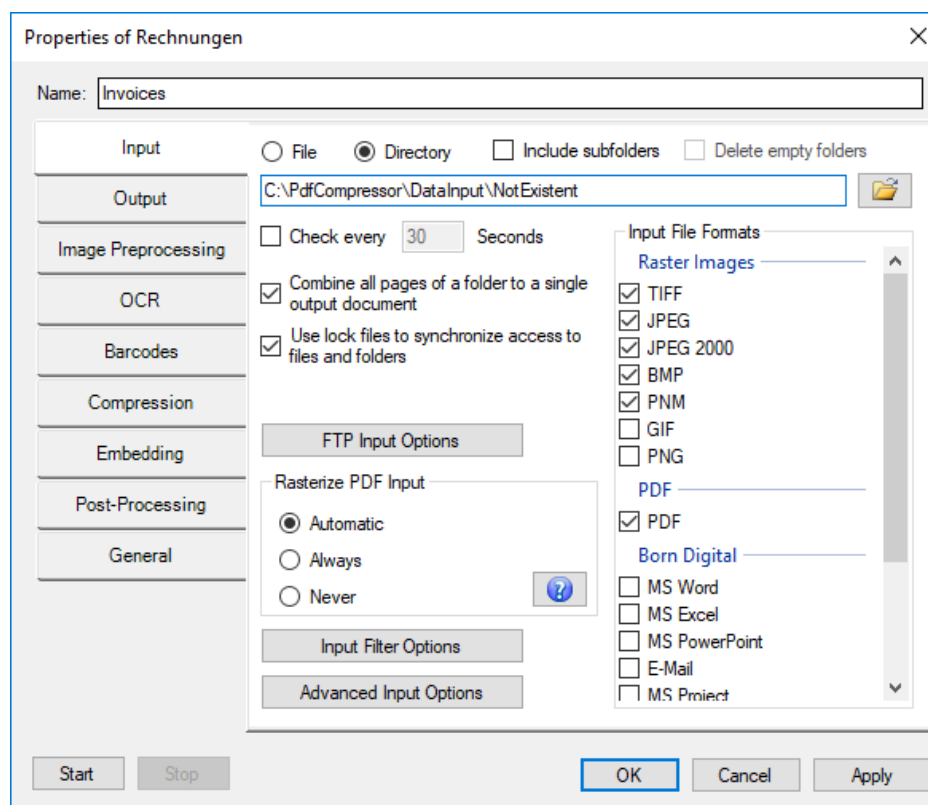


- Select the entry and use the properties button in the toolbar ,
- Select the entry and use the menu *Entry* → *Properties*.

Within the *properties* dialog you can change the name of the entry to make it easier for you to identify the corresponding job. The entry name is also used within the log file (see “*Log Files*” (p. 91)) to relate information, warnings and errors to the respective job.

Configuring Input Data


The *Input* tab lets you configure which input documents should be converted to PDF format.



Currently TIFF, JPEG, JP2 (JPEG 2000), BMP, PNM, GIF, PNG, PDF

formats are accepted as input files. With the *Born Digital component*, additional input formats such as Microsoft Office and saved e-mail messages can be converted.

Multipage TIFF and PDF files are recognized and converted to multipage PDF documents. Moreover you can set up input resampling to adjust the resolution of your data.

1. Select *File* if you want to convert a single file or *Directory* if you want to convert all files within the given folder. You may enter the file or directory, use the browse button  or do drag&drop.

Note: You cannot use mapped network drives to specify the input file or directory. Mapped network drives are not exposed to Windows services such as the underlying PDF Compressor Service. Instead, you can specify the network share using the UNC-path notation `\\host\share\dir`. See “*Choosing the Service Account*” (p. 102), for setting access rights in this case.

Regular paths are supported up to a length of 255 characters. Processing of longer paths is possible by specifying the prefix `\\?\` for local paths or `\\?\UNC\` for paths on network shares. Corresponding paths can only be entered manually and not via the file or folder selection. When processing Born Digital files, the option to *Read local copy of each input file* must be enabled under *Advanced Input Options*.

1. When *Directory* is selected, you have the following options:
 - a. Enable *Include subfolders* if the whole directory sub tree should be converted. Otherwise only files directly within the given folder will be processed.



- b. Enable *Delete empty folders* if you have selected *Include subfolders* and want to clean up the input sub tree. This can be useful if you configure a setup that moves or deletes all processed input files. (See also “*Configuring Post-Processing*” (p. 63).) Once all files of the given job have been processed, the output folder will be examined and all empty subfolders will be deleted. If the input folder is a hot folder (see next item), the cleanup process will be performed whenever the hot folder becomes idle, i.e. when no more files are found to be processed at the moment.
- c. Enable *Check every <number> seconds* to make the given folder a hot folder, i.e. to let the job check that folder frequently for newly generated files (e.g. by an external scanner or capture software). A job with this option enabled will never stop automatically but it keeps on monitoring the given folder. You have to stop it manually.

Note: This option only makes sense, if you delete successfully converted input files, or move them into another folder. See also “*Configuring Post-Processing*” (p. 63).

- a. Enable *Merge all pages* if all pages from all processed files of a folder should be converted to a single PDF output document. The output file will be named <name_of_the_input_folder>.pdf. The input files of a folder will be processed in alphabetical, case-sensitive order. This defines the page order of the output PDF.

This option can be used to merge pages from scanners that put numbered single page files (such as page000.tif, page001.tif, etc.) into a single folder.

- b. *Input file formats* can be selected from the given list. The format is derived from the file extension.
 - *TIFF* enables the conversion of *.tif and *.tiff files. Currently PDF Compressor supports bilevel, grayscale and RGB images. CMYK and CIELab color as well as palletized images are not supported yet.
 - *JPEG* enables the conversion of *.jpg files.
 - *JPEG 2000* enables the conversion of JPEG 2000 files. The extensions *.jp2, *.jpf, *.jpx, *.j2k, *.jpc and *.j2c are recognized.
 - *PDF* enables the conversion of PDF files. If *Always* is chosen under *Rasterize PDF Input*, the input will be rasterized to the resolution specified on the *Advanced* tab of the *Advanced Input Options* dialog (see “*Configuring Advanced Input Options*” (p. 20)) before applying the MRC compression
 - 1. If *Automatic* has been chosen for *Rasterize PDF Input*, PDF files may be rasterized, if for example conversion to PDF/A is not possible otherwise.
 - *BMP* enables the conversion of *.bmp files. Palletized images as well as some exotic variants are not supported yet.
 - *PNM* enables the conversion of *.ppm and *.pgm files.
 - *GIF* enables the conversion of *.gif files. For animated GIF files only the first frame is converted and a warning is logged.
 - *PNG* enables the conversion of *.png files.



- The file formats underneath *Born Digital* can be used in any combination to convert to PDF or PDF/A. Refer to “*Born Digital Conversion*” (p. 6) and “*Requirements of the Born Digital component*” (p. 2) for details on Born Digital Conversion. The following formats can be handled:
 - *MS Word* : *.rtf, *.doc, *.docx, *.docm
 - *MS Excel* : *.xls, *.xlsx, *.xlsm
 - *MS PowerPoint* : *.ppt, *.pps, *.pptx, *.pptm, *.ppsx
 - *MS Project* : *.mpp
 - *MS Visio* : *.vsd, *.vdx, *.vsw
 - *E-Mail* : *.msg, *.eml, *.emlx
 - *OpenOffice Writer* : *.odt, *.sxw
 - *OpenOffice Calc* : *.ods, *.sxc, *.csv
 - *OpenOffice Impress* : *.odp, *.sxi
 - *Text Files* : *.txt
 - *HTML* : *.htm, *.html
- 1. The *Advanced Input Options* button opens a dialog to configure more input options. See “*Configuring Advanced Input Options*” (p. 20) for details.
- 2. If you configure multiple PDF Compressor installations to work with the same input directory (typically a net share), you need to enable *Use lock files for PDF Compressor synchronization*. This ensures that input files currently processed by one PDF Compressor are not accessed by another PDF Compressor. This is implemented using *lock files* *.lock and *.dlock which the PDF Compressor uses to lock certain input files or all files in a directory (when *Merge all pages* is enabled).

Note: The PDF Compressor Service must have write access to input directories when using lock files.

Important note: Files ending with *.lock and *.dlock will be overwritten without notice when using lock files. You should not use these extensions for your own files.

Note: A file PDF_Compressor.ulong can be used as a user defined lock for all files of a directory. In contrast to the lock files *.lock and *.dlock, this file can be managed by the user. If *Use lock files for PDF Compressor synchronization* is enabled, PDF Compressor will never process any file it finds within a directory that contains a file PDF_Compressor.ulong.



This can be used to configure a hot folder with *Merge all pages* enabled: The external process creates a new directory underneath the PDF Compressor hot folder containing a file `PDF_Compressor.unlock` and copies all files that need to be merged afterwards. When the last file copy is done, the external process removes `PDF_Compressor.unlock`. Next time PDF Compressor inspects the hot folder, the processing of the given files begins.

1. The *Rasterize PDF Input* setting offers different modes for handling PDF input documents:
 - a. *Automatic* is the default setting. In this mode PDF Compressor will automatically decide whether or not typeset or vector contents of PDF input files will be rastered. It will try to avoid such conversions but will resort to them as a fallback option for documents that cannot be converted otherwise.
 - b. *Always* advises the PDF Compressor to convert all pages of all PDF input documents to raster graphics before compressing or converting them. For PDF input this mode is equivalent to the former *MRC Compression mode*.
 - c. With the *Never* setting the PDF Compressor will not try to rasterize any non-raster input. In this way you avoid unintentional conversion of typeset or vector content into raster graphics. This may result in failure to convert certain kinds of PDF input files to PDF/A.

Note: With the *Automatic* or *Never* setting chosen, raster graphics already present in the input documents may still be resampled or compressed, if *Color image optimization* or *Bitonal image optimization* are selected on the *Compression* tab. (Such graphics may e.g. be scanned pages or large images inside a PDF document.)

FTP Input Options

The *Read input from FTP server* option allows processing of input files which are directly downloaded from an FTP server. The path to the input folder or file in this case must be configured relative to either root or if given the *Default folder*. Please note that file and folder names are case sensitive on many FTP servers and use slashes '/' instead backslashes '\ ' to separate folders.



FTP Options

☒ Read input from FTP server

FTP Properties

Host name or IP address: Port:

☐ Use SSH FTP (SFTP) ☐ Use secure FTP (FTPS)

Default folder (optional):

☐ Use passive transfers

Login

☐ Use anonymous login

Username: Password:

Connection timeout interval: seconds

Disconnect if idle for at least: seconds

The connection settings for the FTP server can be configured by clicking on the *FTP Input Options* button which opens a separate dialog. The following must be considered:

1. To establish a connection to the FTP server *Host name or IP address* as well as the *Port* fields must be specified. The default port number for FTP is 21 and for SSH FTP is 22.
2. The remaining settings can be specified/activated depending on the server configuration.
3. PDF Compressor supports SSH FTP and FTP unsecured as well as secured over SSL/TLS.
4. The definition of the *Default folder* is optional and serves as basis for specifying a target folder or file.
5. *Use passive transfers* switch enables an automatic port selection by the server.
6. *Use anonymous login* allows to login without username and password.
7. The value of the *Connection timeout interval* specifies the maximum timespan in seconds used to establish a connection. The trial will be aborted with an error message as soon as the limit is reached.
8. *Disconnect if idle for at least* option causes an automatic disconnect to an established connection after a specified number of seconds.



9. Using the *Test connection* button establishes a test connection to the specified FTP server considering the given parameters and disconnects afterwards. The result is shown as a pop-up with either a success or an error message.

Input Filter Options

File Extensions

On the *File Extensions* tab of the *Input Filter Options* dialog you can specify additional input file extensions, such that files with non-standard extensions are processed by PDF Compressor, too.

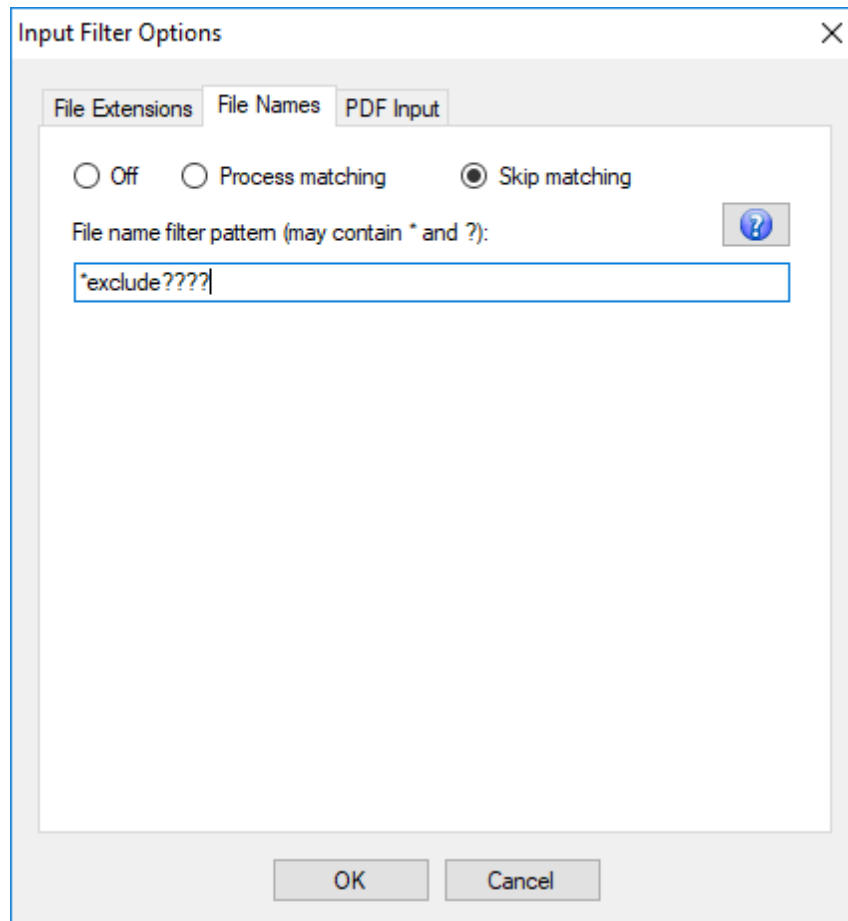
| File Type | Additional Extensions |
|----------------------|-----------------------|
| Raster Images | |
| TIFF | tifx |
| JPEG | |
| JPEG 2000 | |
| BMP | |
| GIF | |
| PNG | ping |
| PDF/JPM | |
| PDF | |
| JPM | |
| Born Digital | |
| E-Mail | mail |
| Born Digital | wpf |

OK Cancel

Each line of the list widget on the *File Extensions* tab shows one of the input file types of PDF Compressor. You can find the file extensions configured by default for each of these types under “*Configuring Input Data*” (p. 11). You can enter additional file extensions for each file type in the list. For one job each additional extension may only be associated with one input file type.

File Names

On the *File Names* tab of the *Input Filter Options* dialog you can specify a file name pattern against which potential input file names are compared. You can also define whether to *Process matching* files only or to *Skip matching* ones.



1. Depending on the selection of either *Process matching* or *Skip matching* potential input files that match the filter criteria are either included in the processing or skipped.
2. In the *File name filter pattern (may contain * and ?)*: entry you specify a filter pattern, against which the names of all potential input files are matched. This pattern may contain the special characters '*' and '?'.

These characters have the same significance as with the Windows file search. The question mark '?' matches an arbitrary single character and the asterisk '*' matches an arbitrary sequence of characters, which may be empty. The pattern is matched only against the base file name, i.e. without the file extension and all alphabetical characters are matched in a case-insensitive fashion.

Examples:

The filter pattern 'Client-????-??' matches the file names 'Client-0000-12.jpeg', 'Client-9900-00.jpeg' and 'client-ABCD-XY.jpg'. It will not match the names 'Client-0000-1.jpg' or 'Client-ABC-XY.jpeg'.

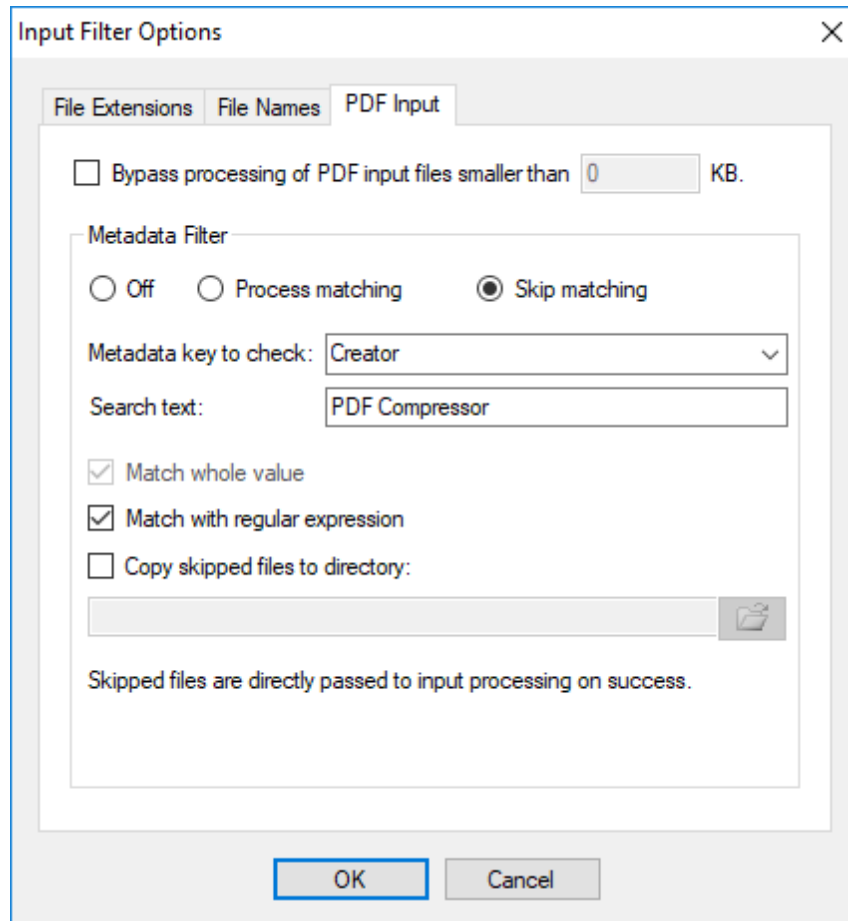
The filter pattern 'File*' matches the file names 'file.jpg', 'File_0123.jpg', 'FILE-A-B-C-D.jpg' and 'File-2016-08.jpg'. It will not match the names 'Fille.jp2', 'Fil.jpeg' or '2016_File.jpg'.

The filter pattern '??_scan*' matches the file names '00_SCAN.TIF', 'ab_Scan.jpg', '09_scan_2014.jpg' and 'XY_Scan_2014.jpeg'. It will not match the names 'scan2014.jpg' or '00_SCA.jpeg'.



PDF Input Options

On the *Input Filter* tab of the *Input Filter Options* dialog you can configure options that exclude certain input files from being processed by PDF Compressor.



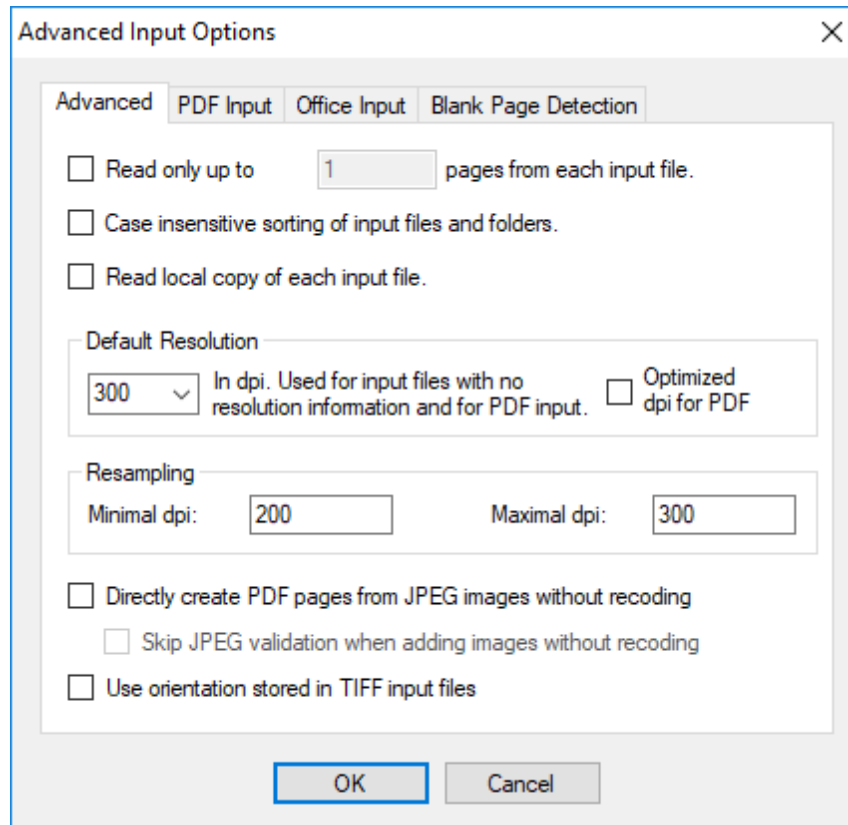
1. The *Bypass processing of PDF input files smaller than ... kB* option serves to skip processing of input files smaller than the given file size. This can be used to suppress the processing of already compressed PDF files.
2. The *Metadata Filter* options allow you to skip processing of PDF input files depending on the PDF metadata. Skipped files will not be processed to output files. Instead they are passed directly to input file post-processing. See also “*Configuring Post-Processing*” (p. 63).
 - a. Depending on the selection of either *Process matching* or *Skip matching* PDF files that match the filter criteria are either omitted from or included in the processing.
 - b. The filter criterion is matched in case the *Metadata key to check:* is found in the PDF Metadata and its value matches the *Search text* below.
 - c. Depending on the option *Match whole value* the *Search text* is either compared to the complete value or searched as part of the metadata value.
 - d. If you need to compare against more than one argument (e.g. *Producer1* and *Producer2*) you can use regular expression by checking *Match with regular expression*.



- e. Skipped input files can be copied to a directory by enabling *Copy skipped files to directory*.

Configuring Advanced Input Options

Clicking the *Advanced Input Options* button brings up a dialog with further settings concerning input options.



Advanced Input Options

1. Enable *Read only up to <number> pages from input file* to limit the number of pages being read from each input file to the given limit. When you merge all files from an input directory (see “*Configuring Input Data*” (p. 11)) to a single PDF output file, setting the maximum number of input pages to 1 can be used to create documents that only contain the cover pages of these files.
2. Select *Case insensitive sorting of files and folders* if you want the input file processing to be in case insensitive alphanumeric order. Otherwise, the ordering will be case sensitive with capital letters lexically smaller than lower case letters and therefore “Ac” coming before “ab”.
3. Enable *Read local copy of each input file* to solve problems related to the location of input files on a network share. The files will be temporarily copied to the local temporary folder where they will be opened and read. This is especially useful when the performance for transmitting small parts of data in the network is low.
4. *Default Resolution* define a resolution in dpi, which is used if the input image file does not contain any resolution information. In addition, this value is used when an input PDF file is rasterized (PDF files do not have specific resolution information in general).



- a. Enable *Optimized dpi for PDF input* to automatically choose an optimal dpi value for the rendering of each PDF input page. The optimal value is calculated from the resolutions of all images embedded in the PDF input page. If there are no images on the page, the setting of the *Default Resolution* will be used. Otherwise the highest resolution of the detected images is calculated and clipped to the resolution range given by the *Resampling* option of the *Input* tab. This option should only be enabled if the input PDF documents originate from a scan process.
 - b. Born Digital PDF documents may contain images of small physical size (e.g. less than an inch in either dimension) but containing many pixels (e.g. 1000 x 1000 pixels) and thus yielding a non-realistic resolution (more than 1000 dpi in the example). For Born dDgital documents you should either use a fixed *Default Resolution* value without enabling *Optimized dpi for PDF input* or specify a narrow resolution range in the *Resampling* settings.
5. *Resampling* generates images in a resolution different from the resolution of the input. That might be useful in the following cases:
- a. If the input resolution is too low (below 150 dpi), the PDF Compressor output might have a quality that is too low for your needs (although the file size will be extremely small). Changing the resolution e.g. to 300 dpi will produce better results.
 - b. If your original scans are high resolution (e.g. more than 600 dpi), you might want to produce a lower resolution PDF output.

The *Minimal* and *Maximal dpi settings* address these requirements. Whenever the resolution of the input files is outside the given [min, max] range, the resolution will be changed by resampling to the minimum value or the maximum value.

Note: If the resolution of the input files is below 150 dpi extremely small output files will be created. But quality and readability are limited. It is recommended to have a minimum resolution of at least 200 dpi.

Please keep in mind that the *Resampling* setting only affects grayscale and color images. The resolution of bilevel or B/W images will not be changed.

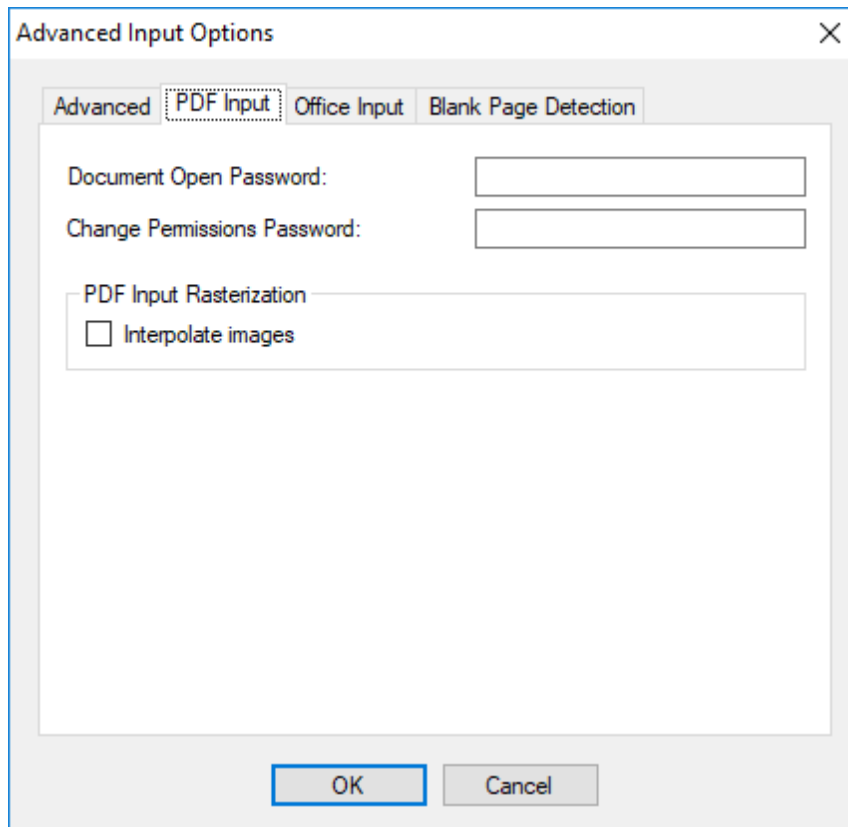
' In case of JPEG 2000 input, the ability to decode a JPEG 2000 file to a lower resolution will be used. This enables fast access to huge JPEG 2000 images, when only a lower resolution is needed.

6. When *Directly create PDF pages from JPEG input without recoding* is enabled, JPEG input images are not submitted to MRC or JPEG 2000 compression, but are directly embedded as pages in the PDF output document. That way repeated compression of input documents is avoided. Thus, these pages are not reduced in size.
7. *Skip JPEG validation when adding images without recoding* omits all integrity checks for JPEG images directly to embed into PDF pages. This involves the risk of creating corrupt PDF output by incorporating corrupt JPEG input images.
8. With the *Use orientation stored in TIFF input files* option the orientation information optionally contained in TIFF input files determines the orientation of the corresponding pages in the PDF output document.



PDF Input Options

The *PDF Input* tab of the *Advanced Input Options* dialog lets you configure options that are relevant for opening PDF input files.



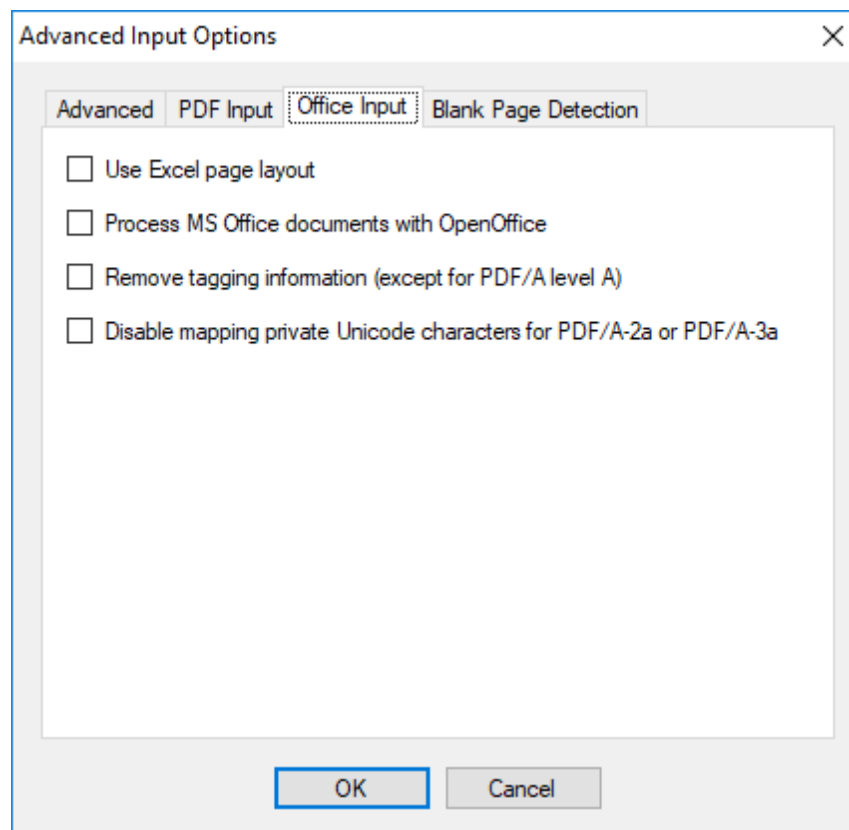
1. The *Document Open Password* is used to open password protected PDF files. The given password is stored within PDF Compressor in an encrypted way. Note that this encryption does not meet the highest security demands. The given password is ignored, when the input PDF file is not password protected.
2. In addition to a password needed to open the PDF document, it can be protected by a *Change Permissions Password*. This password is used whenever the PDF access permissions need to be changed.
3. *PDF input rasterization* options are available unless *Rasterize PDF Input* has been set to *Never*:
 - a. The *Interpolate images* switch influences the rasterization of PDF input files. PDF input files may contain embedded images that need to be scaled for rasterization. *Interpolate images* should be turned on to obtain best image quality. However, if an input PDF file contains only bitonal image content, and this PDF is rasterized to the same resolution as the embedded images' resolution, the image interpolation can be turned off to speed up an OCR process.
4. The option *Restrict PDF Conversion - Use only PDF Optimization for converting PDF input files* is only enabled when the Born Digital component is installed. If selected, it serves to suppress the attempt to use Born Digital Conversion functionality when converting PDF input files. Instead these files are only processed using PDF Optimization features.



Note: This option primarily serves to guarantee backward compatibility with jobs set up with pre-7.0 versions of PDF Compressor, which used the PDF Optimization mode for converting PDF input. In general, it is preferable to use the conversion provided by the Born Digital component.

Office Input Options

On the Office Input tab, you can configure options for the Born Digital conversion of Microsoft Office input files.



1. If the *Use Excel page layout* option is selected, the print page layout as specified inside the Excel input files is used when converting this file to PDF. This option is disabled by default, such that each non-empty sheet inside an Excel input file is converted to one PDF output page.
2. When *Process MS Office documents with OpenOffice* is chosen, all office input files are converted using OpenOffice. Obviously, an OpenOffice installation must be present in this case. This option is switched off by default and each input document is processed using its native application.
3. The option *Remove tagging information* ensures that structure information generated by office applications during PDF export is disposed. Structure information is important for accessible documents and should generally be kept. However, if the priority is on file sizes, this option can help to create smaller PDF output.
4. Option *Disable mapping private Unicode characters for PDF/A-2a or PDF/A-3a*:

Office and PDF documents that contain Unicode characters from the Private Use Area (PUA), cannot be converted to PDF/A-2a and PDF/A-3a without further actions.



To allow conversion of such documents to these standards, the PDF Compressor maps some typical PUA codes used in Office documents to corresponding regular Unicode characters. This does only affect the PDF text content, accessible for example using the clipboard, but not the representation. These characters are for example bullet points.

Using this option you can disable the mapping.

Blank Page Detection Options

On the *Blank Page Detection* tab, you can configure options for the blank page detection of input files. Pages that are detected as blank are always exempt from text recognition (OCR).

The image shows a screenshot of the 'Advanced Input Options' dialog box, specifically the 'Blank Page Detection' tab. The dialog has four tabs: 'Advanced', 'PDF Input', 'Office Input', and 'Blank Page Detection'. The 'Blank Page Detection' tab is active. It contains the following options:

- ☒ Detect blank pages: Detection applies to images and rasterized PDF input only. Blank pages are excluded from OCR.
- Excluded margins:**
 - Top: 0 mm
 - Left: 0
 - Right: 0
 - Bottom: 0
- Tile size: 142 pt (1 pt = 1/72 inch)
- Block size: 1 pt
- Sensitivity: 80
- Threshold: 9 ‰
- ☐ Exclude blank pages from output

At the bottom are 'OK' and 'Cancel' buttons.

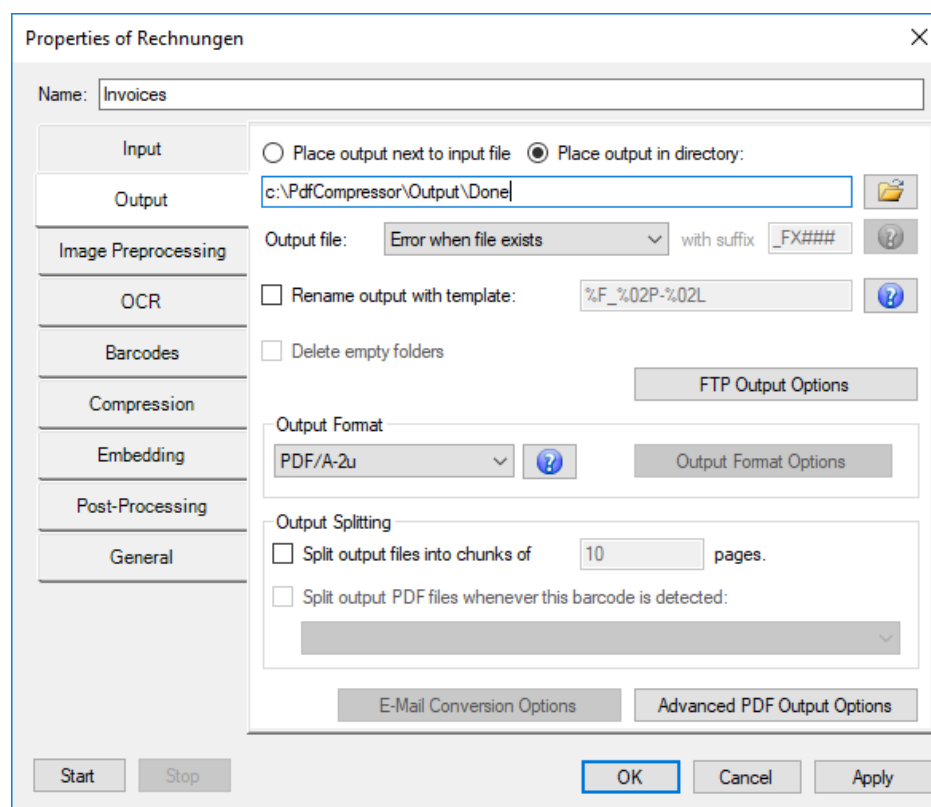
1. *Detect blank pages* enables or disables blank page detection.
2. In the *Excluded margins* box, you can choose how much of the margins of all pages should not be part of the blank page detection. This way, it is possible to exclude invariable page headers or footers or punched holes, for example. You can also choose between the units millimeter (mm) and hundredths of an inch (1/100 inch).
3. The blank page detection decomposes the area of the page to be worked on into *tiles* of configurable size. The detection is done tile per tile. If in the end all tiles are evaluated as blank the whole page is considered blank. The unit for the Tile size value is 'pt', which is equal to 1/72 inch.
4. The analysis of each tile content is block-based. The size of a block should be set to the size of the smallest meaningful symbol, e.g. the size of a punctuation dot. Depending on the page image resolution a block spans multiple pixel. The unit for the Block size is 'pt', which is equal to 1/72 inch.




- For each block a value is calculated that represents how much the block color intensity differs from other blocks in the tile. Each block that has a deviance that exceeds a certain value is considered non-blank. This value is influenced by the *Sensitivity* property. The sensitivity is a value in the range of 1 through 100. A low sensitivity causes only strong colored structures to be counted as being non-blank. In contrast, a high sensitivity also takes pale textures into account.
- A tile is considered non-blank if the number of non-blank blocks exceeds a certain threshold. The *Threshold* is given in per mill (one tenth of a percent) defining the maximum fraction of non-blank blocks in relation to blank blocks for a full-size tile to still be evaluated as blank.
- Checking *Exclude blank pages from output* instructs the PDF Compressor to exclude detected blank pages from the output.

Configuring Data Output

The *Output* tab lets you configure PDF output options including the exact version of PDF to generate and the location where generated PDF files are placed.



- Selecting *Place output next to input file* writes the generated PDF files into the same directory as the input file. Select *Place output in directory* to put them somewhere else. You can either type a path, use the browse button  or drag&drop a folder using Windows Explorer into this window. See “*Managing the PDF Compressor Service*” (p. 101) for the use of network shares.
- There are three ways how PDF Compressor can handle the situation when it attempts to write an output file that is already present at the respective location:



- a. If *Error when file exists* is selected the output file will not be written and you will find an error message in the log file. See “*Log Files*” (p. 91).
- b. If *Overwrite existing file* is selected the existing file will be overwritten.
- c. If *Rename when file exists* is selected PDF Compressor automatically determines a file name that is not yet in use. For this a dynamic suffix will be appended to the base output file name. The format of the suffix can be customized and must meet the following syntax:
 <prefix><counter><postfix>
 - i. <prefix> and <postfix> are optional. If present they may contain lower or upper case latin letters, digits, parentheses, dash, underscore, blank space, or dot, only.
 - ii. <counter> is a sequence of hash characters (#). The number of hash characters denotes the size of the counter, e.g. ## defines a counter that runs from 01 to 99.

Note: If *Output File Renaming* is also enabled, that renaming operation is performed first. The rule specified for *Rename when file exists* is then applied to the new name.

- 3. Select *Delete empty folders* if you want to clean up the output directory sub tree. This option is only available if you select *Include subfolders* for the input processing. (See also “*Configuring Input Data*” (p. 11).) PDF Compressor always creates required output folders in advance, i.e. before any file is processed whose output has to be placed in that folder. On error conditions certain output files may be deleted resulting in empty folders. *Delete empty folders* helps to remove unneeded folders after a job has been finished (or after a hot folder has become idle).
- 4. The *Write output to FTP server* option allows the output files to be uploaded automatically to an FTP server. The path to the output folder in this case must be configured relative to either root or if given the *Default folder*. Additionally there also exists an option to transfer the settings from *FTP Input Options* by activating the *Use FTP input options* checkbox. Please note that folder names are case sensitive on many FTP servers and use slashes '/' instead of backslashes '\' to separate folders. The option *Place output next to input file* in combination with *Write output to FTP server* may only be used if the *Read input from FTP server* options is also selected (for further details refer to “*FTP Input Options*” (p. 15)).
- 5. The option *Output Format* lets you select the format for the output file(s) to create.

PDF Compressor can create PDF documents conforming with different variants of the PDF standard or different so-called conformance levels of PDF/A, and some raster image formats

The recommended default is PDF/A-2u.

a. *PDF 1.7* (ISO 32000-1)

The decomposition of images into different layers and masks is supported in a usable way since Adobe Reader 5.0 (PDF 1.4). PDF documents produced by PDF Compressor require a viewer to support at least PDF 1.4 to be displayed properly.

b. *PDF/A-1a* (ISO 19005-1)



Same as PDF/A-1b, but with automatically inserted rudimentary tagging information to conform to basic requirements of *PDF/A-1a*. This is the accessible variant of PDF/A-1. Therefore, it is strongly recommended to activate OCR in order to generate documents with textual content and structure information that can be processed by screen reader applications.

c. *PDF/A-1b* (ISO 19005-1)

This is the basic compliance level of the original PDF/A document standard for long-term preservation. It is based on PDF 1.5 with further restrictions regarding e.g. meta-data, encryption and transparency.

d. *PDF/A-2a* (ISO 19005-2)

Same as PDF/A-2b, but with automatically inserted rudimentary tagging information to conform to basic requirements of *PDF/A-2a*. This is the accessible variant of PDF/A-2. Therefore, it is strongly recommended to activate OCR in order to generate documents with textual content and structure information that can be processed by screen reader applications.

e. *PDF/A-2b* (ISO 19005-2)

PDF/A-2 is a revision of the original PDF/A-1 standard, allowing for some additional types of content, such as images compressed using JPEG 2000 and other PDF/A files embedded as attachments.

f. *PDF/A-2u* (ISO 19005-2)

This is the recommended default. The format is the same as *PDF/A-2b* with the additional advantage that all text within the PDF must have a consistent Unicode mapping.

g. *PDF/A-3a* (ISO 19005-3)

Same as PDF/A-3b, but with automatically inserted rudimentary tagging information to conform to basic requirements of *PDF/A-3a*. This is the accessible variant of PDF/A-3. Therefore, it is strongly recommended to activate OCR in order to generate docu-



ments with textual content and structure information that can be processed by screen reader applications.

h. *PDF/A-3b* (ISO 19005-3)

The PDF/A-3 standard is equivalent to PDF/A-2 with the only difference that it allows files of arbitrary type to be embedded as attachment to the PDF/A-3 documents.

i. *PDF/A-3u* (ISO 19005-3)

Corresponds to *PDF/A-3b*, with the additional requirement for text within the PDF to have a consistent Unicode mapping.

j. *TIFF*

Raster image file format that supports multi-page and various compression schemes. Black-and-white pages will be compressed with FAX G4 encoding, color and grayscale pages will use LZW compression.

k. *JPEG*

Raster image format, specifically optimized for compressing photographs. This is a single-page image format. Therefore *Output Splitting* must be activated with 1 page per chunk.

l. *PNG*

Raster image format with lossless compression. Like *JPEG* this is a single-page image format and *Output Splitting* is required to be activated with 1 page per chunk.

m. *No Output*

No main output file(s) will be generated. Other output, e.g. OCR results in different formats might be produced.

6. *Output Splitting* can be activated in two ways:

a. *Split output PDF files into chunks of <number> pages* let you limit the number of pages output to a single PDF document. Instead of writing huge PDF documents with many pages, a series of output files can be created, each with the specified number of pages. (Depending on the number of input pages, the last file of the output series will contain less pages.) This mode is typically used with *Output File Renaming* using a template string with %P, %L, or %C escape sequences (see item 6 below).


b. *Split output PDF files whenever this barcode is detected* is available if barcode detection is turned on at the *Barcodes* tab. (See also “*Configuring Post-Processing*” (p. 63).) If enabled, this function splits the output into multiple documents, whenever a new barcode of a particular profile is detected on a page. The next document begins with the first page that contains a new barcode. Via the dropdown box you can choose the barcode profile using its name, provided you have defined one already.

This mode is typically used with *Output File Renaming* using a template string with %V escape sequence representing the value of the detected barcode (see item 6 below).

Both *Output Splitting* modes imply the *Output File Renaming* to be turned on.

7. *Output File Renaming* can be used to customize the naming of PDF output files. The file name template is used to form the output file names. It uses escape sequences as specified in “*Template String Syntax Description*” (p. 73). If *Output Splitting* is enabled, you need to enter escape sequences to distinguish the output names of the different chunks.



A typical example for page number based splitting is %F_%P-%L, where %F is substituted by the input file name, %P by the first page number within the chunk, and %L by the last page number. See “*Template String Syntax Description*” (p. 73) for details and more examples or press the help button  next to the entry.

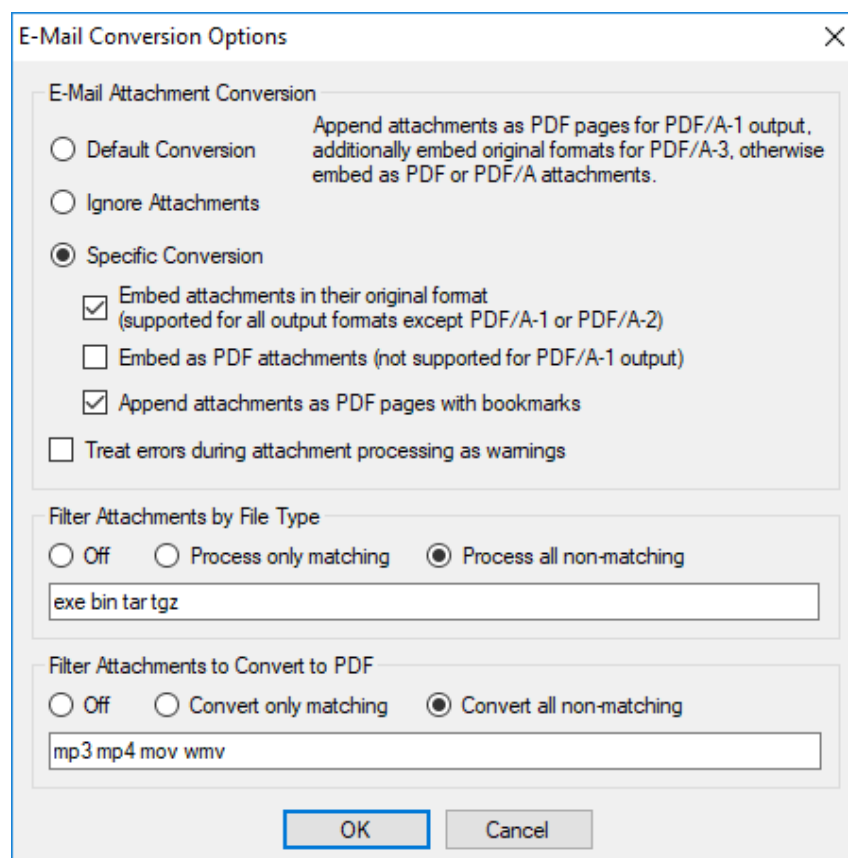
Note: If *Rename when file exists* is also enabled, that change is applied optionally after the renaming specified under *Output File Renaming*.

Configuring E-Mail Conversion Options

Clicking the *E-Mail Conversion Options* button brings up a dialog with further options regarding the conversion of attachments to e-mail messages.

Note: These options are available only if *E-Mail* as input format is chosen on the *Input* tab.

Note: Both e-mail message formats such as MSG and EML and the PDF format support the concept of attachments. These attachments are files of arbitrary formats embedded in the overall file, which acts as a container.



The dialog box titled "E-Mail Conversion Options" contains the following sections:

- E-Mail Attachment Conversion**
 - ☐ Default Conversion: Append attachments as PDF pages for PDF/A-1 output, additionally embed original formats for PDF/A-3, otherwise embed as PDF or PDF/A attachments.
 - ☐ Ignore Attachments
 - ☒ Specific Conversion
 - ☒ Embed attachments in their original format (supported for all output formats except PDF/A-1 or PDF/A-2)
 - ☐ Embed as PDF attachments (not supported for PDF/A-1 output)
 - ☒ Append attachments as PDF pages with bookmarks
 - ☐ Treat errors during attachment processing as warnings
- Filter Attachments by File Type**
 - ☐ Off
 - ☐ Process only matching
 - ☒ Process all non-matching
 - Text input field: exe bin tar tgz
- Filter Attachments to Convert to PDF**
 - ☐ Off
 - ☐ Convert only matching
 - ☒ Convert all non-matching
 - Text input field: mp3 mp4 mov wmv

Buttons: OK, Cancel



1. PDF Compressor offers the following options for converting the attachments of e-mail input files:
 - a. The *Default Conversion* option chooses the best conversion available for the selected PDF Output Format (see “*Configuring Data Output*” (p. 25))
 - When converting to PDF/A-1 the e-mail attachments of all supported input file types are converted to PDF/A-1 and appended as pages to the PDF output document, since PDF/A-1 does not allow for PDF attachments. In addition, a PDF bookmark with the name of the attachment will be created per e-mail attachment, which points to the first page of the corresponding content in the PDF output document.
 - PDF/A-2 permits attachments in PDF/A format. Therefore, per default each e-mail attachment is converted to a PDF/A-2 document and attached to the output document as a PDF attachment.
 - Since PDF/A-3 additionally allows PDF attachments of arbitrary file format, the default here is to convert e-mail attachments to PDF/A and append them as pages to the PDF output document. In addition the e-mail attachments are embedded as PDF attachments in their original format.
 - In all other cases, each e-mail attachment is converted to a PDF document and then embedded as an attachment to the PDF output document.
 - b. *Ignore Attachments* simply discards all e-mail attachments and merely converts the message bodies.
 - c. The *Specific Conversion* setting lets you specify in detail how attachments should be processed. Conforming to the selected *PDF Output Format* you may choose an arbitrary combination of appending e-mail attachments as pages or embedding them as PDF attachments in their original format or converted to PDF(/A). The options described below serve to configure this specific conversion.
 - d. The option *Treat errors during attachment processing as warnings* causes PDF Compressor to continue processing even if it failed to process one or more e-mail attachments. Otherwise this is considered an error and processing the corresponding job is halted.
2. The *Filter Attachments by File Type* discriminates among e-mail attachment types by specifying a list of file extensions. If settings chosen under *E-Mail Attachment Conversion Options* require the conversion of an e-mail attachment to PDF(/A) the PDF Compressor will try to convert attachments of all supported input file types. Embedding e-mail attachments as PDF attachments in their respective original format works for any file type.
 - If *Off* is chosen here, PDF Compressor will process attachments as described in the previous paragraph.
 - If *Process only matching* is selected only e-mail attachments with the specified file extensions will be processed; all others will be ignored.
 - If *Process all non-matching* is chosen e-mail attachments with the specified file extensions will be ignored and all others will be processed.

The file extensions must be separated by spaces, either with or without leading dot - i.e. both "mpeg mpg jpeg jpg" and ".mpeg .mpg .jpeg .jpg" are valid.



3. *Filter Attachments to Convert to PDF* provides options to further restrict the set of e-mail attachment types specified by the *Filter Attachments by File Type* options regarding the types of files to convert to PDF(/A).
 - If *Off* is chosen here, PDF Compressor will convert all attachments of supported input file types which have passed the previous filter.
 - If *Convert only matching* is selected only e-mail attachments with the specified file extensions will be converted to PDF(/A); all others will be excluded from the conversion.
 - If *Convert all non-matching* is chosen e-mail attachments with the specified file extensions will be excluded and all others will be converted to PDF(/A).

Configuring E-Mail Header Language

The headers of e-mails (such as "Subject") will be shown in the language that is selected for the user interface, for supported languages. If the language is set to *Automatic*, the system language is used. The currently supported languages are English and German. Refer to "*Language*" (p. 94) on how to select the language.

Configuring E-Mail Header Elements

If the preset template does not fit your need, you may adapt the template as follows: In your installation directory navigate `\var\Mail\Templates` and copy the "Foxit" folder, rename the copy to "Custom". Now edit the ".js" according to your needs.

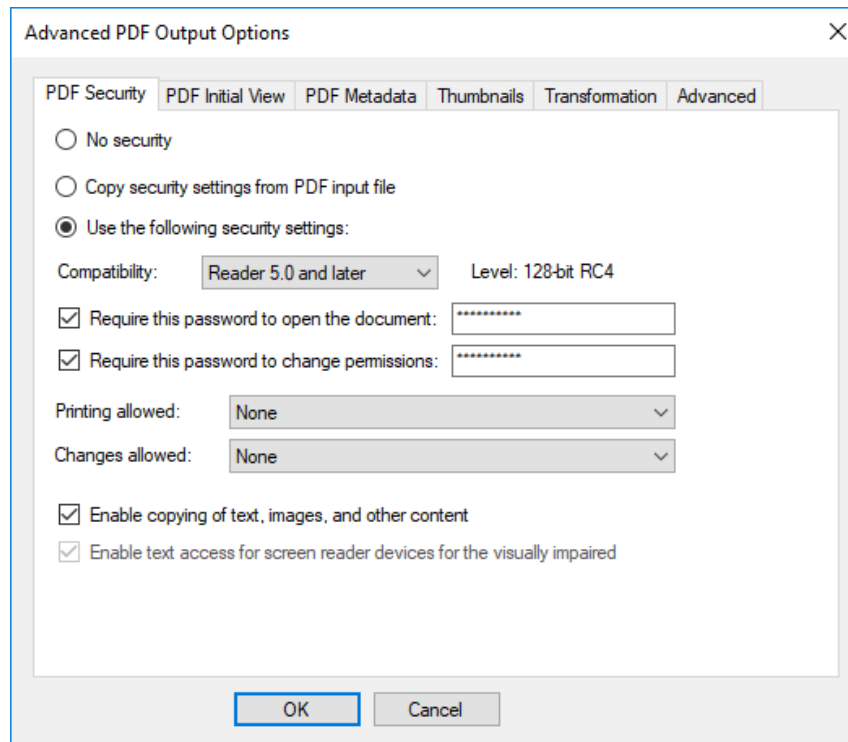
Configuring Advanced Output Options

Clicking the *Advanced PDF Output Options* button brings up a dialog with further settings regarding output generation.

PDF Security

The *PDF Security* tab lets you configure the security features of the output PDF documents. Currently, only security compatibility for Adobe Reader 3.0 and higher, and Adobe Reader 5.0 and higher is supported.

Note: PDF Security settings are disallowed for PDF/A compatibility.



1. You can either specify not to use any security on output PDF documents (*No security*), copy the security settings from the input files (*Copy security settings*), or use the settings given by this dialog (*Use the following security settings*). Copying security settings from input files is only possible in PDF optimization mode.
2. *Compatibility* lets you select Reader 3.0 or Reader 5.0 security compatibility. Depending on this setting, the following access permission options will slightly differ. See the *PDF Reference* document available from Adobe for details.
3. *Require this password to open the document*: encrypts the PDF content with the given password. This is true, even if the security settings from the PDF input file do not require a password. If a password is necessary to open the input file, it will be replaced by the one given here.
The given password is stored within PDF Compressor in an encrypted way. Note that this encryption does not meet the highest security demands.
4. *Require this password to open the document*: for preventing unauthorized users from changing the PDF file. As in the case above, this option overwrites respective security settings from the input file.
The change permissions password can be bypassed by third-party products that do not fully support or respect the access permissions settings.
The given password is stored within PDF Compressor in an encrypted way. Note that this encryption does not meet the highest security demands.

If a change permissions password is given, and *Use the following security settings* is chosen, the access permissions can be specified by the following options:

- a. *Printing allowed* can be set to None, Low resolution (with Reader 5.0 security compatibility), and Full resolution.

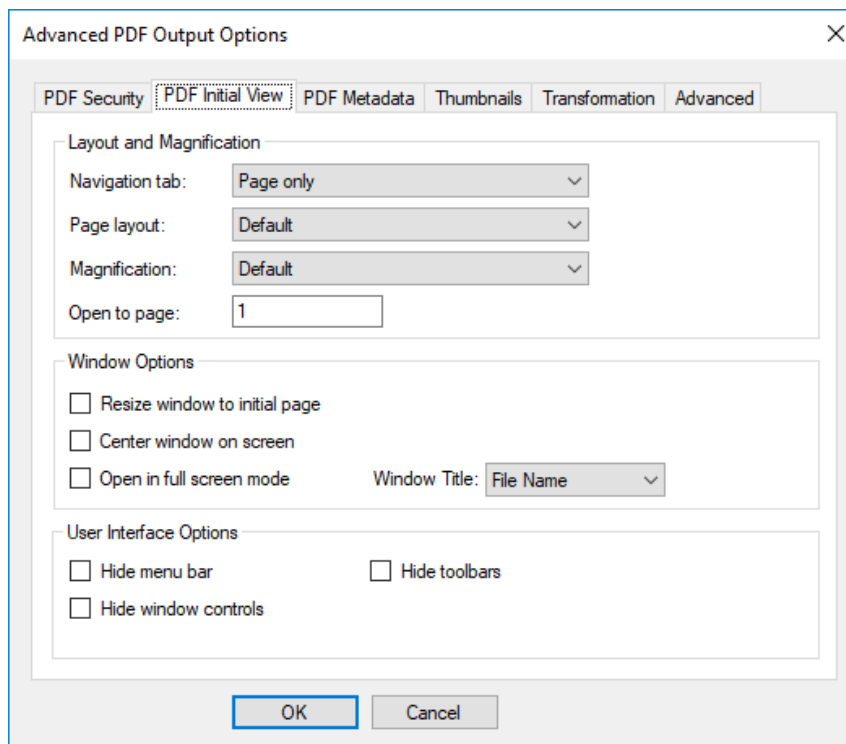


- b. *Changes allowed* can be set to various restrictions depending on the PDF security compatibility. Please refer to the PDF Reference document available from Adobe for details.
- c. *Enable copying of text, images, and other content* allows these operations on the protected PDF documents. If the security compatibility is Reader 3.0, this includes the access for the visually impaired.

With *Reader 5.0 and later* security compatibility, the text access for screen reader devices for the visually impaired can be selected independently from the access given under 7.

PDF Initial View

The *PDF Initial View* tab lets you configure the initial view of a PDF document when it is opened with a PDF reader. In addition, the appearance of the reader itself can be influenced.




1. *Layout and Magnification* sets the appearance of the opened PDF document:
 - a. *Navigation tab* lets you select the panel that is opened at the left-hand side of the PDF reader window. It can be bookmarks, the pages, or the layout panel, or no panel at all (Page Only).
 - b. *Page layout* sets the appearance (and number) of pages that will be shown in the PDF reader page window. There are settings for individual pages, a continuous page flow, and various options for two-page display.
 - c. *Magnification* sets the initial magnification of the document. It can be either a specific zoom level or an option to fit the page, its width or its height into the PDF reader window. *Actual Size* is a synonym for a zoom level of 100%.

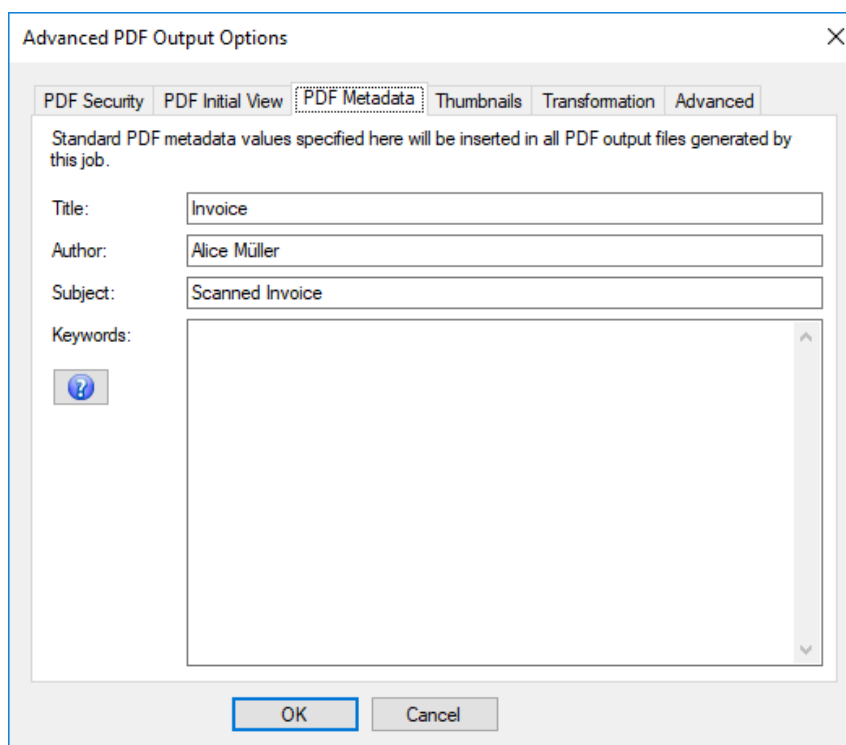


- d. The initially opened page of a document can be set by *Open to page*. If the given page is greater than the number of pages of the document, the initial opened page will be the last page of the document.
2. *Window Options* influence the appearance of the PDF reader window when a document is opened.
- a. The PDF reader window can be resized such that the initial page fits to the window.
 - b. The PDF reader window can be centered on the screen.
 - c. The PDF reader can go into *Full Screen mode* when a PDF document is opened.
 - d. You can select if the PDF file name or its title is shown in the PDF reader title bar.
3. *User Interface Options* can be used to hide various elements of the PDF reader window. You can
- a. *Hide menu bar*
 - b. *Hide toolbars*
 - c. or *Hide window controls* (such as scroll bars and navigation controls).

PDF Metadata

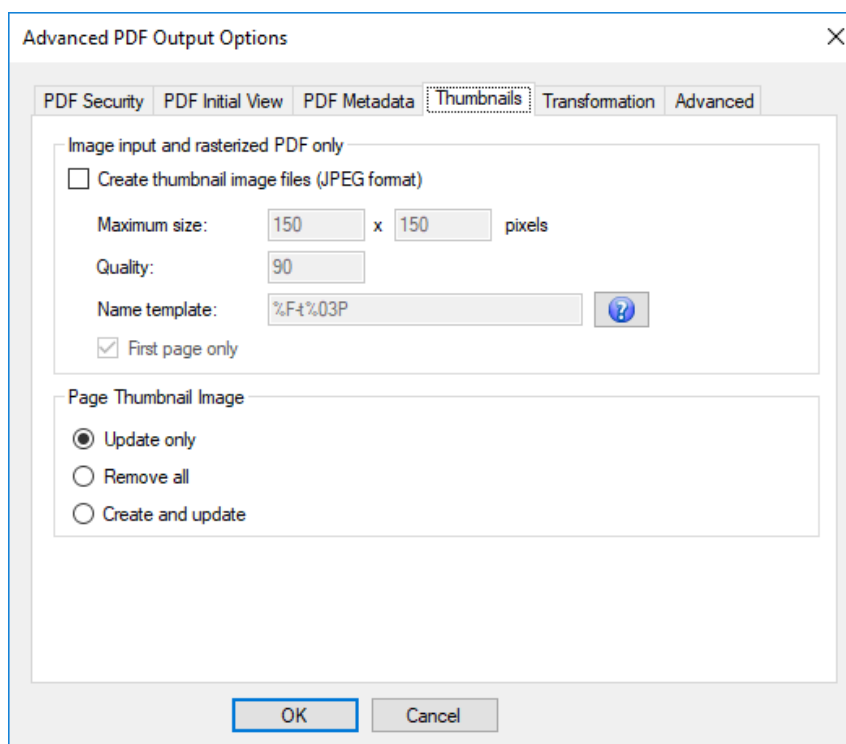
The data entered in the fields of PDF Metadata will be embedded in the output PDF document. You can find that information within Adobe Reader under the menu *File → Document Properties → Description*.

It is possible to put certain escape sequences into the PDF Metadata fields that will be substituted by information such as file name, number of pages, date, time, etc. The help button  displays a list of all possible substitutions. See also “*Template String Syntax Description*” (p. 73).



Thumbnails

The *Thumbnails* tab can be used to configure the page thumbnail images of the input documents.




1. The *Image input and rasterized PDF only* enables and configure the output of additional thumbnail images containing page views of the input documents. Thumbnails are output as JPEG files next to the PDF output file.



- a. Enable *Create thumbnail image files (JPEG format)* to turn the thumbnail output on.
- b. *Maximum size* contains the maximal dimensions of the output thumbnail images in pixel. The first number is the horizontal size, the second is the vertical size. The aspect ratio of a page will not be changed. A value of 0 means the original size as maximum for this dimension.

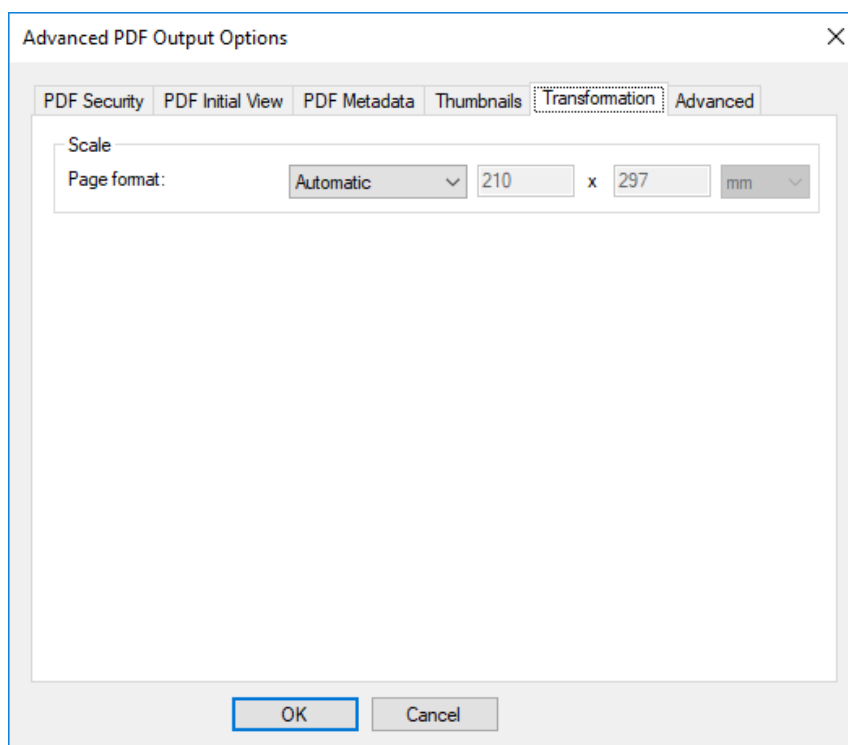
Example:

Assume *Maximum size* is set to 150 x 150.

- i. A 20 x 30 cm sized page will be scaled to 100 x 150 pixels.
 - ii. A 30 x 20 cm sized page will be scaled to 150 x 100 pixels.
- c. Set *Quality* to any value between 1 and 100. This is used to specify the image quality of the output JPEG files. Higher qualities result in larger file size.
 - d. The *Name template* is used to define the thumbnail file naming convention. It uses escape sequences as specified in Template String Syntax Description, "*Template String Syntax Description*" (p. 73). A typical sequence contains %F as a substitute for the input file name, and %P as a substitute for the current page number. See "*Template String Syntax Description*" (p. 73), for details and examples on output file naming, or press the help button  within the PDF Compressor application.
 - e. Enable *First page only* if you only want to output a thumbnail for the first page of a document. Otherwise thumbnails for all the document pages will be output.
2. The *Page Thumbnail Image* setting specifies how to handle the embedded thumbnails for the pages of the PDF output file. Embedding the thumbnails results in a larger size of the PDF file, but speeds up the display of page preview by older versions of the Adobe Reader.
- a. The *Update Only* updates the existing thumbnails, but does not create any new ones.
 - b. The *All Remove* setting removes all thumbnails of the output PDF document.
 - c. With the setting *Create and Update*, new thumbnails are created and embedded for all pages that do not yet have thumbnails. The existing thumbnails are updated.

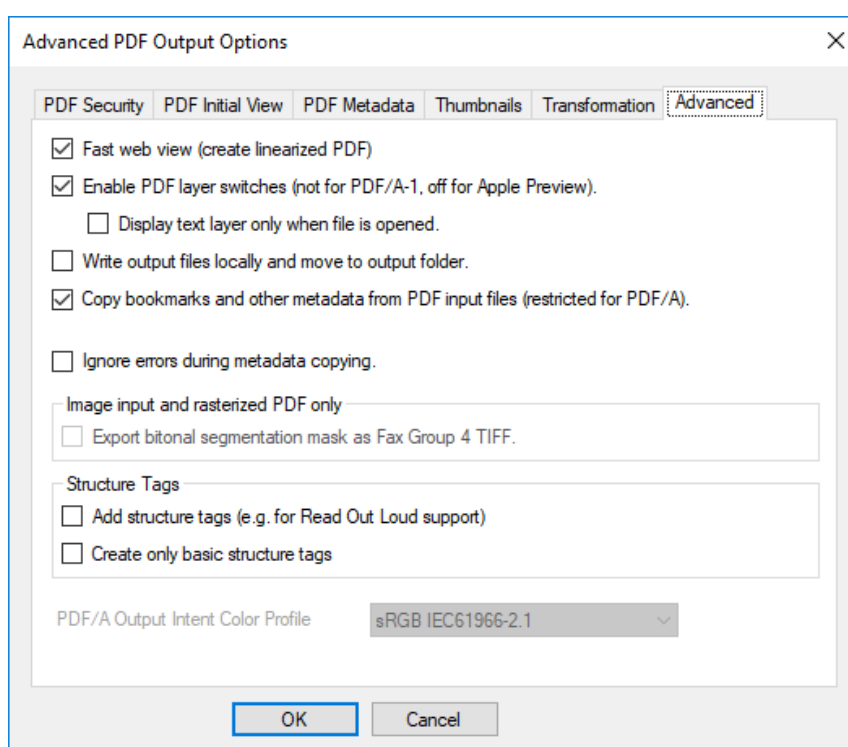
Transformation

This tab lets the user configure one of the given transformation options for the output PDF document.



1. *Scale* can be used for page size adjustment of the output PDF document while preserving the aspect ratio of each input page.
 - a. *Page format* setting allows to choose between one of the given standardized, a user defined or an automatic format selection. By default it is set to *Automatic* which means the dimensions from the input document are applied page for page.

Advanced





1. Enabling *Fast web view* will optimize the output PDF documents for viewing them within a web browser. Standard PDF documents need to be downloaded completely before the very first page is displayed within a web browser. This can take a significant amount of time, if the internet connection is slow or if the document contains a lot of pages. Web optimized PDF documents will show the first page immediately when this portion of the data is downloaded. Moreover, you can quickly jump to other document pages. The Adobe Reader plug-in will just download the page you selected. Depending on your Adobe Reader settings, the remaining pages will be downloaded in the background.
2. *Enable PDF layer switches...* lets you enable or disable the embedding of PDF layer switches within the output PDF document. When this feature is turned on, the Adobe Reader will show you switches within its layers tab that let you control the rendering of the internal layers:
 - a. PDF layer switches are not available for PDF/A-1 output compatibility. PDF/A-1 disallows the use of optional content.
 - b. By default, all layer switches are turned on when the PDF file is opened in Adobe Reader. If you want to display the text layer only when a file is opened in Acrobat, enable *Display text layer only when file is opened*.
3. *Write file locally and move to output folder* may be enabled to solve problems when the output file has to be written to a network share. Moreover, if the performance for transmitting small parts of data in the network is low, this option enhances performance.
4. You can optionally *Copy bookmarks and other metadata from PDF input files*. If configured accordingly under *Rasterize PDF Input*, PDF input files can be rasterized. In that case, all meta information contained in the PDF input files is lost, unless the *Copy bookmarks and other metadata* option has been selected.

With this option all bookmarks, XMP metadata, and data of the PDF information dictionary from the input file are copied to the corresponding PDF output file. PDF Metadata set in the *Advanced Output Options* dialog (i.e. non-empty strings) will overwrite any value copied from the input files. In case of PDF/A compliant output, only bookmarks that contain PDF/A-compliant actions are copied (e.g. go to page is allowed, whereas execute JavaScript is disallowed). Moreover, only XMP data from the schemas predefined in the XMP Specification are copied. PDF Compressor does not perform any schema validation. It is the responsibility of the user to ensure that the input files contain properly formed XMP when this option is used.

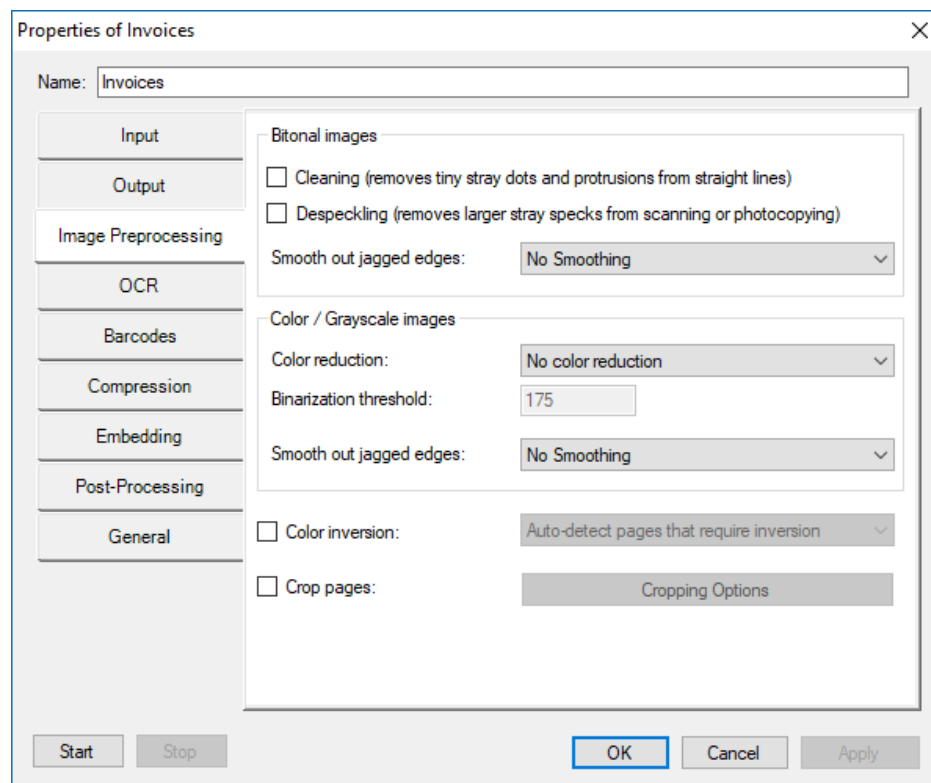
5. The option *Ignore errors during metadata copying* permits the suppression of corrupt metadata. Processing of the input document will not be aborted. Instead, PDF Compressor continues and handles the document as if it did not contain any metadata.
6. If the *Export bitonal segmentation mask as Fax Group 4 TIFF* is enabled the mask layer of the MRC Compression (sometimes referred to as text layer) is exported as a separate bitonal TIFF file. These TIFF files will be placed in the PDF output folder and will be named according to the PDF files. Multipage TIFF images are generated for multipage PDF output files.
7. To support tag-based interpretation of text by PDF software you can enable *Add structure tags...*. Tags are for example necessary to successfully read the PDF result using the "Read Out Loud" feature in Adobe Software. Creating PDF/A-1a output it is not necessary to enable this feature. Structure tags are always created in case of PDF/A conformance level 'a' output, even in case the option is disabled.



8. By selecting the option *Create only basic structure tags* you can enforce the PDF Compressor 7 behavior, to create relatively simple structure information by using the OCR result. If the option is not enabled (default), a advanced analysis of the document will be used to create advanced structure tags even for more complex documents.
9. The *PDF/A Output Intent Color Profile* allows the selection of several generic ICC profiles. The selected profile is used in PDF/A-Output to define the device independent standard color space. In case colors in the file are not already associated with a clearly specified color space they will be interpreted in the selected color space.

Configuring Image Preprocessing

On the tab *Image Preprocessing* several options can be chosen to optimize input images in before the actual conversion.



1. For *bitonal images* you can enable *Smoothing*. Three different modes - *Auto*, *Low* and *High* - are available which set the level of smoothing according to their names.
2. For *Color / Grayscale Images* further options are available:
 - a. *Color reduction* allows to change color images to grayscale or both color and grayscale images to bitonal images. In case of bitonal target color space several modes apply:
 - i. *Convert to grayscale*
 - ii. *Binarize with Dithering* emulates grayscales with different densities of black on white.
 - iii. *Binarize with Threshold* the smaller the value, the brighter the image
3. Again, you can enable *Smoothing*. Three different modes, *Auto*, *Less* and *More* are available, which set the level of smoothing according to their names.



4. *Color inversion* determines, whether images are to be inverted and if so, whether PDF Compressor decides where to apply inversion or if all images have to be inverted.
5. If you wish to crop pages, you'll find the respective settings in an extra dialog under *Cropping Options*:

Cropping Options

Mode: Automatic cropping zone detection (with constraints) ▾

Unit: ☒ cm ☐ in

Outer cropping zone

Point 1:
 Distance from left ▾ 0 cm
 Distance from top ▾ 0 cm

Point 2:
 Distance from left ▾ 0 cm
 Distance from top ▾ 0 cm

Inner cropping zone

Point 1:
 Distance from left ▾ 0 cm
 Distance from top ▾ 0 cm

Point 2:
 Distance from left ▾ 0 cm
 Distance from top ▾ 0 cm

Padding

Left: 0 cm Top: 0 cm ☒ Padding fill color:
 Right: 0 cm Bottom: 0 cm Black ▾
 White
 Black

Note: Automatic cropping is only supported for bitonal images. Color and grayscale images will not be cropped.

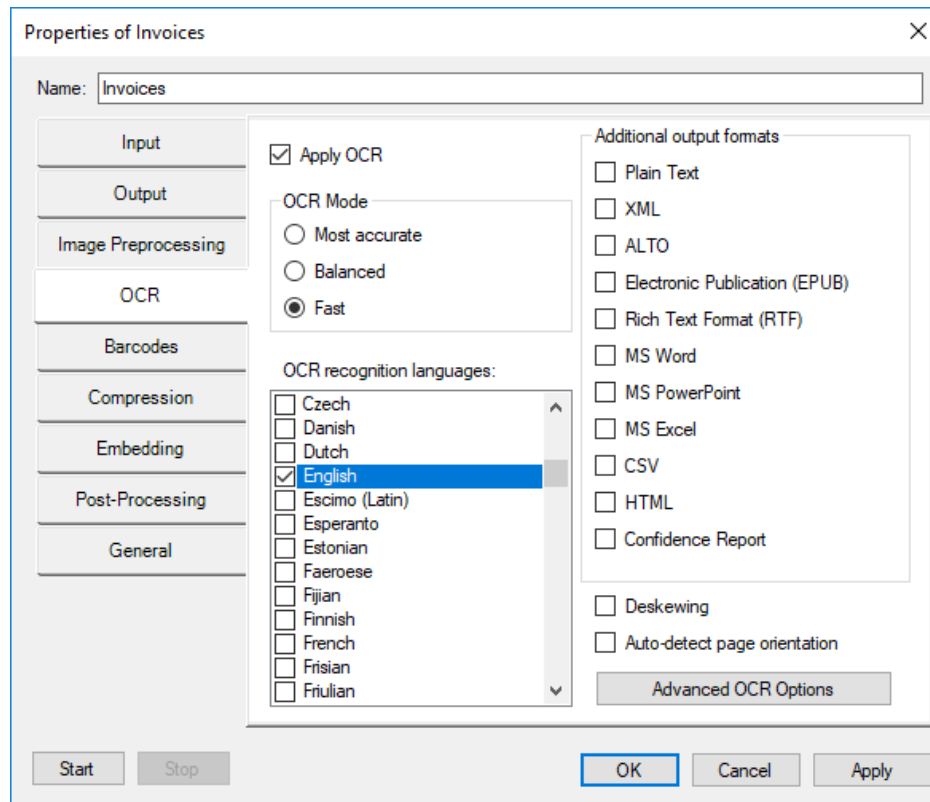
OK Cancel

1. *Mode* defines how cropping is applied:
 - a. *Automatic* lets PDF Compressor decide which images to crop and which not.
 - b. *Fixed* crops all images whose size exceeds the *Cropping Zone*.
 - c. *Constrained* mode determines automatically where to crop each page, but will only attempt to crop between the given *innerbox* and *outerbox*.
2. The *Unit* setting applies to all following dimensional properties.

In *Automatic* and *Constrained* mode a *Padding Zone* can be defined which is filled either with *white* or *black*.

Configuring OCR

The *OCR* tab contains options for configuring the optional text recognition from raster input documents.



1. The *Mode* section lets you control the tradeoff between OCR accuracy and speed.
 - a. *Most accurate* is the mode which gives the highest accuracy in character recognition. It should be used when the input image quality or resolution is low, and recognition time is not that critical.
 - b. *Balanced* is an intermediate mode between most accurate and fast mode.
 - c. *Fast* provides 2 - 2.5 times faster recognition speed at the cost of a moderately increased error rate (1.5 - 2 times more errors). On good print quality texts, the OCR engine makes an average of 1 - 2 errors per page, and such moderate increase in error rate can be tolerated in most cases.
2. Select the *OCR recognition languages* from the list. Text recognition will be more accurate if you select exactly those languages that appear in your documents.
 - Custom OCR Dictionary - In addition to the standard OCR languages PDF Compressor supports a user-supplied custom dictionary. This dictionary may contain words not directly pertaining to any language such as proper names or brand names. A custom dictionary has to be provided in form of a plain text file. The file must be UTF-8 encoded, contain one word per line and be named `pdfc_userdict.txt`. If you place the file in `ocr` folder of your installation directory, the custom dictionary will be used by the OCR.
3. You may choose *Additional Output Formats* to export the OCR results in various formats in addition to the PDF output.



Note: Whenever you select an additional OCR output, the overwrite protection (see Overwrite existing within the Output tab under “*Configuring Data Output*” (p. 25)) is extended to the additional file formats. Example: The output name is outfile.pdf, you select additional Plain Text OCR output, and Overwrite existing is disabled. If any of the files outfile.pdf or outfile.txt already exist, the PDF Compressor will not process the corresponding input file. You need to enable Overwrite existing if you want to overwrite existing files.

Important Note: Some of the additional OCR output formats may create auxiliary files (e.g. JPEG images to be used within an output HTML file). These auxiliary files are not subject to overwrite protection. Thus an existing file that has the same name as an auxiliary file will be overwritten regardless of the setting for Overwrite existing. See the list below for what formats create what auxiliary files.

a. *Plain Text (TXT)* exports OCR results as text. Extension is *.txt. No auxiliary files are created. One of four text formats can be chosen from the list next option:

- i. *Plain Text - UTF8* – unformatted Text (UTF8 code page).
- ii. *Plain Text - UTF8 formatted* – formatted Text (UTF8 code page).
- iii. *Plain Text - ANSI* – unformatted Text (ANSI code page).
- iv. *Plain Text - ANSI formatted* – formatted Text (ANSI code page).

The formatted text output is formatted using spaces and blank lines so that the text position is as close as possible to the text position in the input document.

b. *XML* exports OCR results in XML format. Extension is *.xml. One of four various XML formats can be chosen from the list next to the XML option:

- i. *Full* - all possible XML attributes will be written (large XML files)
- ii. *Simplified* - some optional attributes will be suppressed (smaller XML files)
- iii. *Word-level* - detailed character related information is suppressed. The recognized text is set as direct content of the formatting element. Resulting XML files are very



small (much smaller than simplified), they are human readable and easier to post process.

- iv. *Line-level* - similar to word-level, but without any formatting information.
- v. *Nuance XML* - exports OCR results in XML format according to Nuance specification.

Note: With the *Full*, *Simplified*, *Word-level* and *Line-level* formats the coordinates given within the XML files always relate to the resampled and not deskewed input image (compare “*Configuring Advanced Input Options*” (p. 20)).

- c. *ALTO* exports OCR results in the open XML standard ALTO (Analyzed Layout and Text Object).
- d. *Electronic Publication (EPUB)* exports OCR results in the EPUB format for electronic books. Extension is *.epub.
- e. *Rich Text Format* exports OCR results in Microsoft RTF format (to be opened with Microsoft Word). Extension is *.rtf. No auxiliary files. One of three various RTF formats can be chosen from the list next to the RTF option:
 - i. *MS Word 2000*
 - ii. *WordPad*
 - iii. *MS Power Point 97*
- f. *MS Word* exports OCR results in Microsoft Word format. Extension is *.docx.
- g. *MS PowerPoint* exports OCR results in Microsoft PowerPoint format. Extension is *.ppt. An auxiliary directory <output-name>_files is created. It contains various files used by the PowerPoint document.
- h. *MS Excel* exports OCR results in Microsoft Excel format. Extension is *.xls. No auxiliary files.
- i. *CSV* exports OCR results as comma separated values in plain text (ANSI code page). This is especially useful if the input document contains tables that should be imported into some other application. Extension is *.csv. No auxiliary files.
- j. *HTML* exports OCR results in HTML format. Extension is *.htm. Embedded images are stored as auxiliary files with naming convention <output-name>-<image-number>.jpg.
- k. *Confidence Report* - selecting this option will generate a separate file containing confidence information, essentially how “sure” the OCR engine is about the extracted data.
 - i. *Level* - confidence can be reported on a character, word, page or file basis.



- ii. *Metrics* - allows to select between mean, median or minimum measurement algorithm, which is used to calculate a confidence value for report levels like word, page and file.
4. Enable *Deskewing* if you want the OCR engine to deskew (align) your pages. This option only applies to raster input documents or PDFs which have been rasterized.

Please note that this option will improve the OCR result for PDFs which have been rasterized, but will not deskew them.

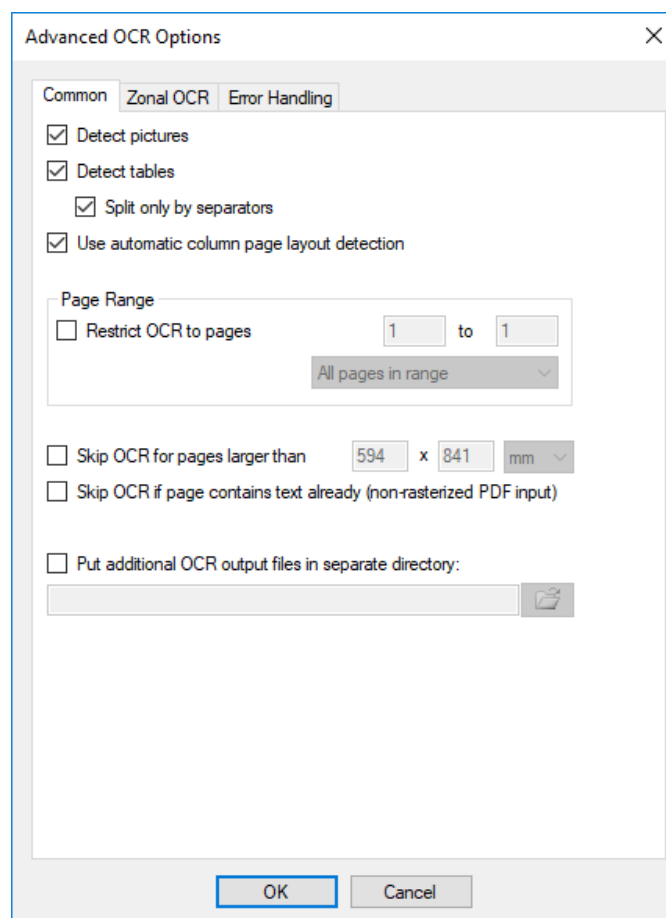
If the pages are additionally rotated in 90 degrees steps, for example 180 degrees, you should also enable *Auto-detect page orientation*.

5. *Auto-detect page orientation* rotates the page images in steps of 90 degrees to ensure that the text is correctly oriented. This is a useful feature when pages have been scanned with wrong orientation, but it only works for raster input documents or PDFs which have been rasterized. For this kind of documents, also the OCR result will be much improved.

If the page is not rotated exactly in 90 degrees steps, you should also enable the *Deskewing* option.

Advanced OCR Options

The *Advanced OCR Options* dialog is opened via the *Advanced OCR Options* button in the OCR tab and displays itself three tabs where you may configure detailed settings for OCR.





1. *Detect pictures* is enabled for faster recognition ignoring areas that are recognized as images during page analysis. Disable the option to recognize all text on a page even inside pictures.
2. *Detect tables* enables or disables the recognition of table structures. Checking *Split only by separators* advises the engine not to recognize table-like structures from appearance, but only from lines.
1. *Restrict OCR to pages...* determines, which pages will be considered for OCR.
2. The OCR can take a long time or even fail on very large pages, especially if these pages have little text content. Therefore, it makes sense to exclude such pages from OCR. Select *Skip OCR for pages larger than* option if you have mixed input that includes normal textual pages as well as very large sheets, e.g. drawings or maps. The maximum dimensions of pages that should be processed by OCR can be given in millimeter or inch.
3. When processing PDF input without rasterizing there is an option to *Skip OCR if page contains text already*. Please note that the OCR is suppressed for all pages that have any kind and amount of real text (set with fonts) on it. It is not ensured that the text covers the raster images that would be processed by OCR otherwise.
4. Select *Put additional OCR output files in separate directory* if you want the additional OCR output files not to be placed next to the output PDF files, but into its own directory. The browse button lets you select the directory. When processing a whole input directory tree (see “*Configuring Input Data*” (p. 11)), a corresponding subdirectory tree will be created under the specified OCR output directory.

OCR Zones

The image shows the 'Advanced OCR Options' dialog box with the 'Zonal OCR' tab selected. The 'Common' tab is also visible. The 'Zonal OCR' section has a checked 'Enable Zonal OCR' option and an unchecked 'Use for pages larger than' option with input fields for 0 x 0 cm. Below this is a list of 'OCR Zones' with one zone selected: 'Zone 1 (5 [left] / 5 [top] x 15 [left] / 10 [top] cm)'. Below the list are 'New' and 'Delete' buttons, and up/down arrow buttons. The 'Zone Properties' section has a 'Zone name' field, 'Point 1' with 'Distance from' dropdowns for 'left' (5 cm) and 'top' (5 cm), 'Point 2' with 'Distance from' dropdowns for 'left' (15 cm) and 'top' (10 cm), and a 'Unit' section with radio buttons for 'cm' (selected) and 'inch'. At the bottom are 'OK' and 'Cancel' buttons.



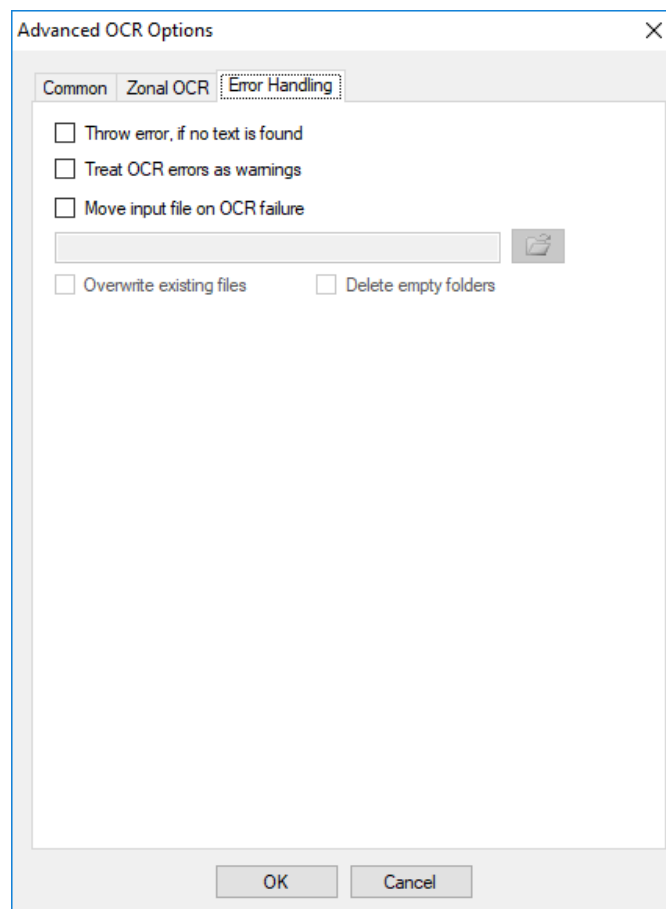
Under *Zonal OCR* you can define whether and how OCR is limited to regions:

1. *Enable Zonal OCR* restricts OCR to zones.
2. *Use for pages larger than...* delimits the previous option to pages that exceed one of the given dimensions.
3. *OCR Zones* are the regions that you have defined. Using the *New* button, you can create an arbitrary amount of zones. You may also name the zones like *adress* to make distinction easier. If you leave the name empty, the zone is named 'Zone XY' with continuous numbering. *Delete* removes the selected zone, whereas the up and down buttons shift it along the list.

Note: The order of appearance in the list determines the order of recognition results in the output stream. E.g., if you define *address* before *Invoice date* the address is written to the output file first, regardless of where the zones are placed on the page.

A zone is defined by two points, which can each refer to the left or right page edge and to the top or bottom, respectively. You can choose suitable units for these distances.

Error Handling



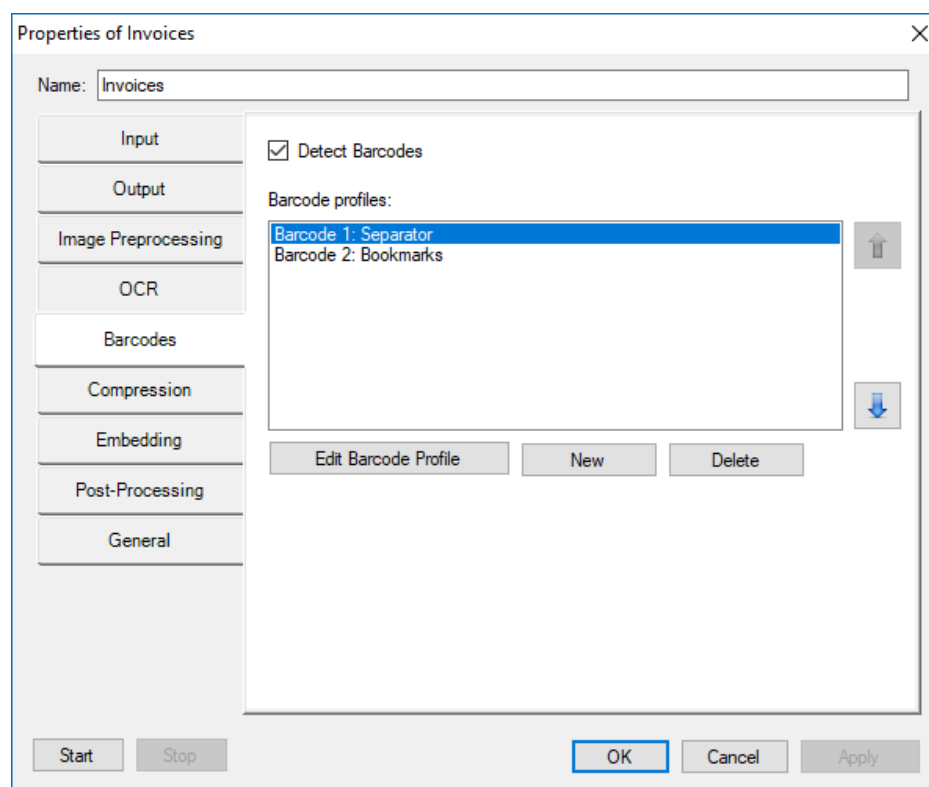
1. To detect that no text was found in the input document the option *Throw error, if no text is found* may be enabled.



2. *Treat OCR errors as warnings* means that errors during OCR are only reported as warning but will not stop the processing of the affected document.
3. Select *Move input files on OCR failure* if you want to move input files which caused an OCR failure to a special folder, e.g. OCR failure. The browse button lets you select the directory or you can enter the path manually or just use drag&drop to define the target directory.

Configuring Barcode Recognition

The *Barcodes* tab contains options for recognizing barcodes. Several types of barcodes can be recognized and the values of the barcodes can be used in several ways. In this tab you can specify which actions to perform upon detection of which type of barcode, or barcode value. These settings are called *Barcode Profiles*.



In the tab you see an overview of the Barcode profiles you have defined. You may create arbitrarily many profiles. For easier distinction you can give them meaningful names. With the *New* and *Delete* buttons you can create or delete profiles. Using the arrow buttons, you can change the order of the profiles.

Barcode Profile Options

The *Barcode Profile Options* dialog is opened via the *Edit Barcode profile* button of the *Barcodes* tab. Here you specify all the particulars of the profile.



Barcode Profile Settings [X]

Profile identifier:

Barcode type(s) to detect:

- ☒ Auto Detect
- ☐ Codabar
- ☐ Code 128
- ☐ Code 39
- ☐ Code 39 Extended
- ☐ Code 93
- ☐ DataMatrix
- ☐ EAN
- ☐ IATA 2 of 5
- ☐ Industrial 2 of 5
- ☐ Interleaved 2 of 5
- ☐ Matrix 2 of 5
- ☐ Patch Code
- ☐ PDF417

☒ Treat every barcode value as a new value

☐ Add bookmark for each barcode

☒ Exclude pages with detected barcode from output

Restrictions

☐ Search area on the page defined by two points: Unit:

Point 1: Distance from cm

Distance from cm

Point 2: Distance from cm

Distance from cm

Minimum barcode height: cm

Minimum value length:

Minimum number of bars:

☐ Match value with regular expression:

☒ Collect only first matching barcode

Recognized barcode values can be added to the PDF metadata (see “*PDF Metadata*” (p. 34)) or used to specify an output filename (see “*Configuring Data Output*” (p. 25)). Furthermore, a recognized barcode can trigger the start of a new output document (output splitting by barcode detection, “*Configuring Data Output*” (p. 25)).

1. At the top of the dialog, you can name the profile. The name will be used for instance in the profile overview, so you can recognize it.
2. From the list under *Barcode type(s) to detect* you choose barcode types you wish to detect. *Auto Detect* is the default. Note, some of the barcode types are incompatible to each other, i.e. they cannot be enabled simultaneously. In case you need a combination of incompatible types you have to define several profiles. The following restrictions apply:
 - a. *Auto Detect* tries to find barcodes of the following types on the page: 'EAN 8/13', 'UPC-A', 'UPC-E', 'ITF (Interleaved 2 of 5)', 'Code 39', 'Code 39 Extended', 'Code 128', 'Codabar', 'UCC/EAN Code 128', 'Code 93', 'PDF417', 'Data Matrix', 'QR Code (Quick Response)'. We recommend selecting a specific set of barcode types to improve the performance and detection quality.
 - b. Types *Patchcode*, *Postnet*, *PDF417*, *DataMatrix* and *QR-Code* may only be enabled alone.
 - c. You may select at most one of *Interleaved 2of5 (ITF)*, *Industrial 2of5* and *IATA 2of5*.



- d. You may select **either** *EAN* or *UPC-A*
 - e. Select no more types than necessary. Speculative selection increases error probability.
 - f. Select at least one type.
3. The option *Treat every barcode value as a new value* changes the behavior with respect to recurring barcode values. The default behavior is to ignore barcode occurrences if their value equals the value of the directly preceding barcode (of the same profile). When *Treat every barcode value as a new value* is selected, every barcode instance is treated as if its value were found for the first time.
 4. Enable *Add bookmark for each barcode* to add a bookmark pointing to the page of the barcode for each barcode that matches the restrictions given above. If page exclusion is activated, the bookmark will point to the next page. The bookmark value is used as the title.
 5. Enable *Exclude pages with detected barcode from output* to not output pages that contain a barcode that is recognized with respect to all the restrictions given. This is typically used when multiple input documents are separated by a plain page with barcode for reference.
 6. Select *Search area on the page defined by two points*: if you do not want to detect barcodes from the whole page area, but want to specify a certain region, where barcodes are going to be detected.

The rectangular region is defined by specifying two points (e.g. the upper left corner and the lower right corner of the rectangle). A point is given by its distance from the left or right edge of the page, and by its distance from the top or bottom edge of the page. Use the drop-down selectors to set left or right for the horizontal distance, and top or bottom for the vertical distance. The distances can be specified in cm or inch.

7. The option *Minimum barcode height*: defines the minimum height of the searched barcodes. Barcodes with lesser heights than specified will get ignored.
8. The option *Minimum value length*: defines the minimum length of the barcode value. Barcodes whose values have less characters than specified will get ignored.
9. The option *Minimum number of bars*: defines the minimum number of barcode bars. Shorter barcodes will get ignored.
10. The detection of barcodes can be restricted to barcode values that match a given regular expression. Enable *Match value with the regular expression*: and enter the corresponding expression into the given field. For a definition of regular expressions see the `<regex>` part of the syntax defined under “*Regular Expression Substitution*” (p. 77).

Example: If you only want to detect barcodes that begin with the digit 4, followed by any number of digits, use the regular expression `4[0-9]*`

11. *Collect only first matching barcode* - if this option is checked only the first barcode found on page or in the specified search area.

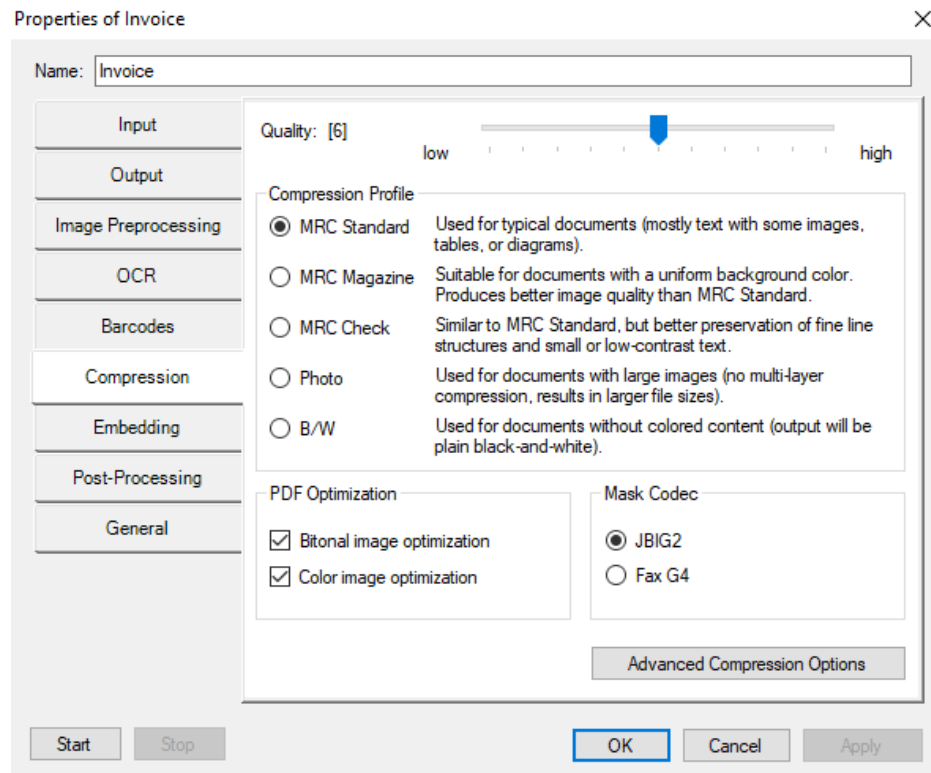


Note that the options specified in different profiles do not need to exclude each other. Therefore, a barcode may be matched by multiple profiles. This can have an effect on all the actions taken because of the barcode.

The %V template variable is set only by barcodes from the first profile.

Configuring Document Compression

The *Compression* tab lets you configure the options for the image compression.



These include the overall quality setting, adoption to special input document types and other options: . *Quality* affects the compression of grayscale and color images in the PDF. A higher quality value results in higher PDF image quality, but PDF file sizes will increase. A lower quality value corresponds to higher image compression ratios and thus to smaller file sizes.

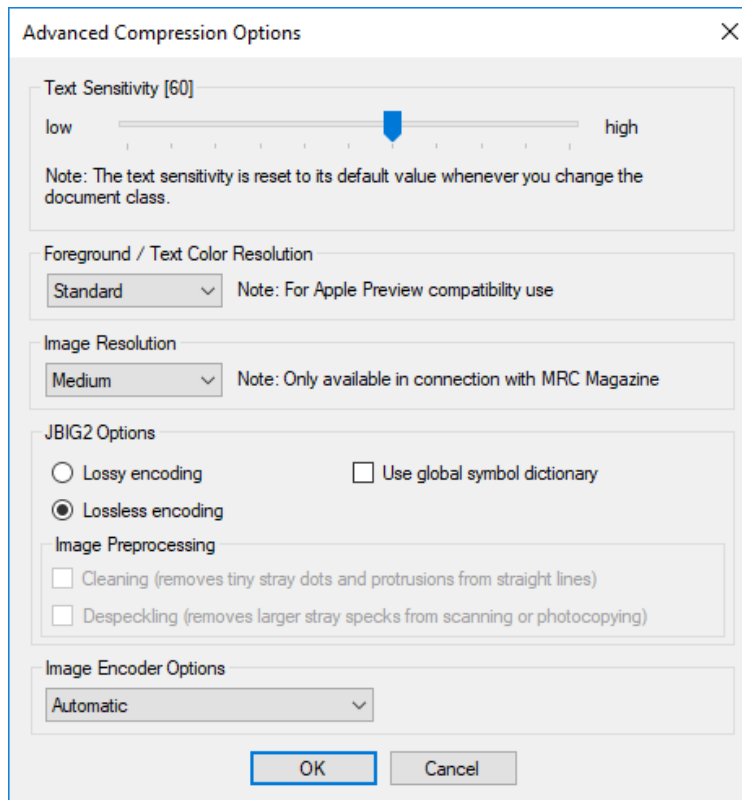
1. The *Compression Profile* can be specified to optimize the compression results for specific input document types:
 - a. *MRC Standard* should be used for all purpose input documents. This is the default setting.
 - b. *MRC Magazine* can be used for documents with mixed text and image content to obtain a better visual quality and compression rate for the result in most cases. The algorithm requires the pages to have a uniform background color. For pages not matching this requirement the *MRC Standard* profile is used as fallback automatically.
 - c. *MRC Check* should be used to obtain better quality if the input document contains a lot of fine structures like lines, thin graphical drawings, etc.



- d. *Photo* turns off segmentation into layers. Everything is put into the background image layer. This should be only used if your input consists of images with no or little text and/or graphics.
 - e. *B/W* (black & white) turns off image color. Your documents are stored in black and white color, as if they had been transmitted by a b/w fax machine.
2. *PDF Optimization* allows to recode bitonal and/or color images in the PDF document using more efficient compression algorithms. This optimization will always be performed unless you set the *Rasterize PDF Input* setting to *Always* (compare “*Configuring Input Data*” (p. 11)).
- a. Enable *Bitonal image optimization* to apply recoding of bitonal images using advanced JBIG2 compression. The JBIG2 options described under “*Configuring Advanced Compression Options*” (p. 51) will be used for this recoding.
 - b. Enable *Color image optimization* to apply recoding of color images found within the PDF document. The color images will be analyzed and, depending on their size and resolution, be replaced by an MRC compressed version (JPEG 2000, PDF/A-2 or PDF/A-3 compatibility has been chosen). The compression options described above apply here, too.
3. *Mask Codec* can be chosen between *JBIG2* and *Fax G4* compression for the text mask. Fax G4 uses lossless image compression, whereas JBIG2 can be used for lossless and lossy compression (further JBIG2 options can be set within the Advanced Compression Options dialog, see *Configuring Advanced Compression Options* below). In lossless mode JBIG2 has a better compression performance than to Fax G4, resulting in smaller file sizes with the same visual quality.

Configuring Advanced Compression Options

Use the *Advanced Compression Options* button on the *Compression* tab to access the advanced compression options dialog.



1. The *Text sensitivity* parameter controls how much of the text will be put into the text mask. A higher value for the text sensitivity may make the text appear bolder, whereas a lower value may make it appear thinner. Normally, values in the range 60 to 70 yield best results. However, if the input images have low contrast, an adjustment of the text sensitivity from its default value may give better results.

Note: Whenever you change the document class in the Encoder tab, the text sensitivity is reset to its default value which depends on the document class.

1. The *Foreground / Text Color Resolution* affects the visual quality of colored text or lines and the compatibility with some insufficient PDF Viewer applications. Some PDF Viewer applications do not render pages in full quality in case the resolution of the foreground image is lower than the foreground mask (standard). The following options are available:
 - a. *Standard* enables foreground sub sampling for minimal file size.
 - b. *Increased quality* will improve the visual quality of colored text and lines. The foreground sub sampling is decreased at the cost of increased file size.
 - c. *High quality* switches foreground sub sampling off. Additionally, a higher foreground image compression ratio is used to limit the increase in file size. Color reproduction for text and lines is further improved. This option may be used to work around problems with PDF Viewer applications that do not render the standard sub sampled foreground in full quality.
 - d. *Max quality* results in maximum text and line color quality at significantly increased output file size. (No foreground sub sampling and standard compression ratio)
2. There are four settings under *JBIG2 Options*.



- a. *Lossy encoding* produces the smallest file sizes.
 - b. *Lossless encoding* results in larger file sizes but does not change the text mask. Use this option for optimal quality and to prevent symbol/character mismatches in case of low resolution or low symbol/character quality.
 - c. The option *Use global symbol dictionary* can increase the compression rate notably and further reduce the output file size, especially for documents with many pages and homogenous type faces. This feature is only available for lossy JBIG2 encoding.
 - d. *Cleaning* removes small stray dots which are usually artifacts from a scan process.
 - e. *Despeckling* is a bit more ambitious than regular cleaning. It removes larger specks that are judged to be artifacts from scanning or photocopying. This computerized judgment call is not infallible, however, so the resultant document should be checked to ensure that no portion of the actual image was lost in this process.
3. *Image Encoder Options* determine which image codec is used for color images embedded into the layers of the created PDF documents (e.g. text color and background image layer when page segmentation is applied):
- a. *JPEG 2000* selects the newer JPEG 2000 compression.
 - b. *JPEG* selects the older DCT-based JPEG compression.
 - c. *Automatic* determines the selection of JPEG 2000 or JPEG compression based on the size (in pixels) of the encoded image. Using Adobe Reader, the display of document pages with large JPEG 2000 images can become very slow or even faulty, depending on the user interaction such as zooming or panning. To overcome this problem, the *Automatic* mode uses JPEG instead of JPEG 2000 compression whenever the width or height of an image is greater than 5120 pixels.

Note: This option is not available for PDF/A-1 compatibility. These compatibility modes must use JPEG compression by definition.

Note: This option is not available for PDF/A-2 or PDF/A-3 compatibility when the Document Class is set to Photo with Quality 11 (lossless). This special setting features a lossless compression for the page image. Since lossless compression is not possible with JPEG, this mode always uses lossless JPEG 2000 compression.

Note: The PDF layers option present in older versions of PDF Compressor has been moved to the Advanced Output Options dialog described under “*Configuring Advanced Output Options*” (p. 31).

Note: The PDF input rasterization option present in older versions of PDF Compressor has been moved to the Advanced Input Options dialog described under “*Configuring Advanced Input Options*” (p. 20).



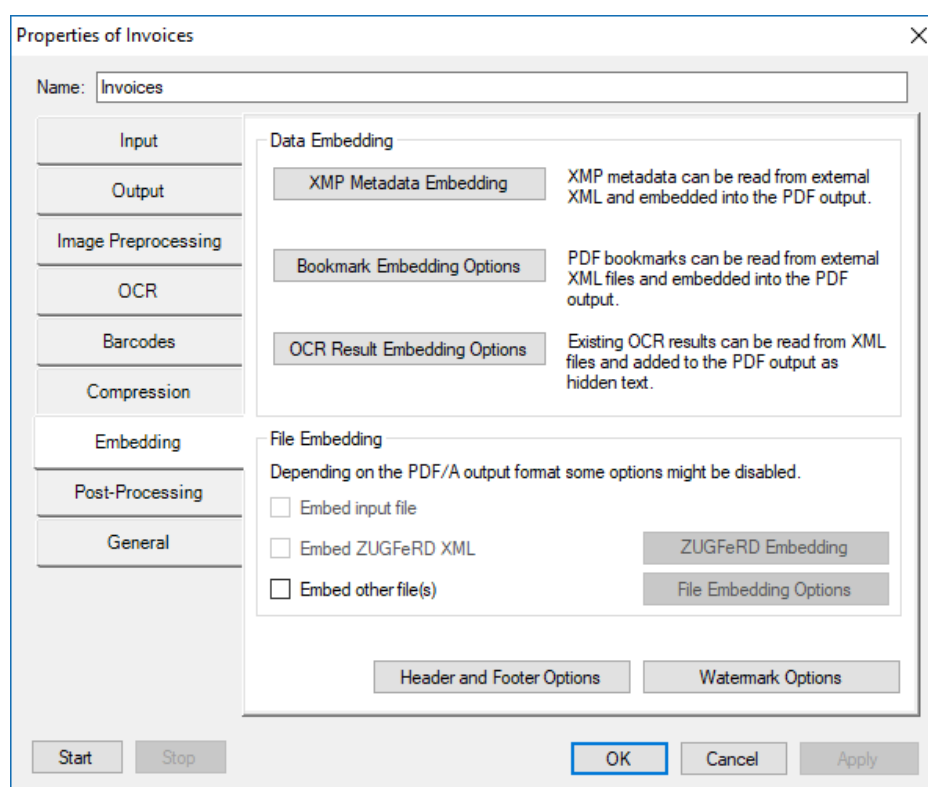
1. The *Image Resolution* affects visual quality of images in combination with *MRC Magazine* profile. The following options are available:
 - a. *Low* enables image sub sampling for minimal output file size.
 - b. *Medium* will improve the visual quality of images by decreasing the sub sampling at the cost of increased file size.
 - c. *High* results in maximum image quality at significantly increased output file size due to full resolution.

Configuring File and Data Embedding

The *Embedding* tab of the *Job Properties* dialog lets you configure options for several types of embedding files or supplementary data into your PDF output files.

Configuring Data Embedding

Three buttons under *Data Embedding* open special configuration dialogs which let you configure the embedding of three different kinds of supplementary data in your PDF output files.

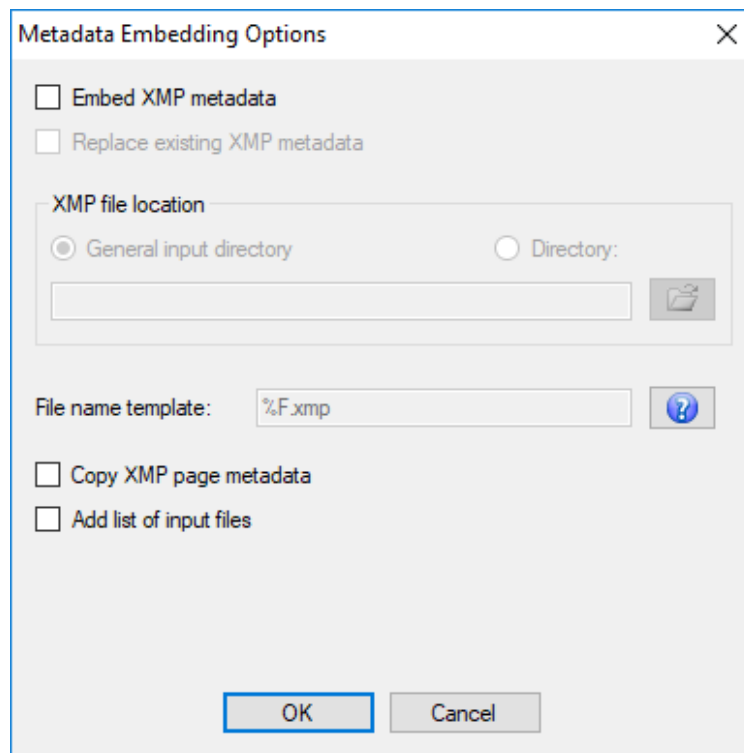


All three kinds of data - XMP metadata, PDF bookmarks and XML OCR results - are read from files containing XML of a corresponding dialect. The XML encoding must be UTF-8.

Note: Input processing as configured for the job (see “*Configuring Post-Processing*” (p. 63)) to move or delete the input files after processing is applied to these files only, if they are located next to the main input files (*General input directory*) and have non-static file names (escape sequences in their file name templates).



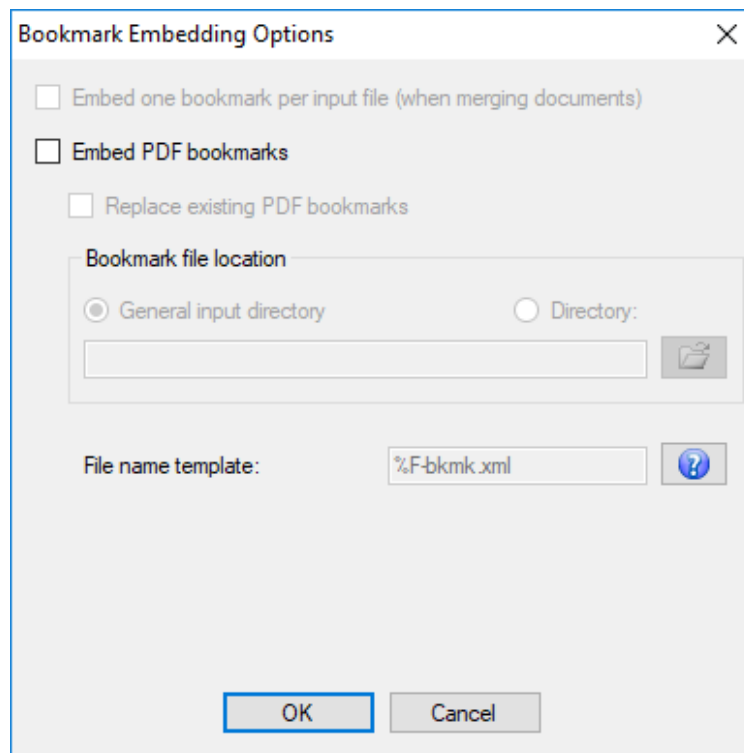
Options for Embedding XMP Metadata



1. Select *Embed XMP metadata* to embed existing metadata in XMP XML format into your output PDFs.
 - a. Select *Replace existing XMP metadata* to replace metadata already present in PDF input files. Otherwise, these data will be preserved and the additional metadata will be appended. This setting has no effect on input files in formats other than PDF.
 - b. Select *General input directory* to read XMP metadata files from the same directory as the main input files. Alternatively, choose *Directory* and specify a different source directory.
 - c. Under *File name template* you can specify a template to derive the name of each XMP metadata input file from the name of the corresponding input file. See “*Configuring Post-Processing*” (p. 63) for more information on template syntax.
 - d. Enable *Copy XMP page metadata* to copy XMP metadata embedded in TIFF files to the metadata associated to the resulting PDF page.
 - e. *Add list of input files* may be selected to add the full path of every input file to a list in the XMP document metadata.



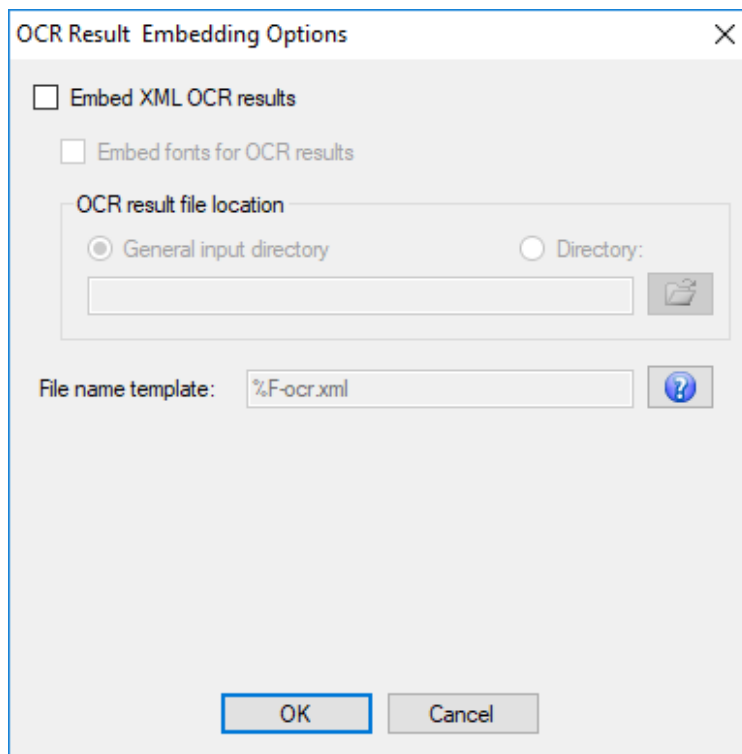
Options for Embedding Bookmarks



1. The option *Embed one bookmark per input file (...)* is only enabled when merging input documents. In that case it will insert a bookmark for each input file, pointing to the location of its original first page in the output file.
2. Select *Embed PDF bookmarks* to embed existing bookmarks in XML format into your output PDFs. The XML encoding must be UTF-8. A XML sample file and the schema definition can be found in the `\examples` subfolder in your PDF Compressor installation directory.
 - a. Select *Replace existing PDF bookmarks* to replace bookmarks already present in PDF input files. Otherwise, these entries will be preserved and the additional metadata will be appended. This setting has no effect on input files in formats other than PDF.
 - b. Select *General input directory* to read PDF bookmark files from the same directory as the main input files. Alternatively, choose *Directory* and specify a different source directory.
 - c. Under *File name template* you can specify a template to derive the name of each PDF bookmark input file from the name of the corresponding input file. See “*Configuring Post-Processing*” (p. 63), for more information on template syntax.



Options for Embedding XML OCR results



1. Select *Embed XML OCR results* to insert OCR results in Nuance XML format into your output PDF files. If such OCR results for your documents already exist, this is an alternative way to embed them as hidden text without submitting the files to OCR again. The XML encoding must be UTF-8.

Note: OCR results written to separate XML files by the PDF Compressor can only be embedded later if they were created either with the full or simplified option.

Note: If PDF input files already contain hidden text and are processed in PDF optimization mode, Embed XML OCR results will not remove or replace this text. Instead, PDF Compressor will add an additional layer for further hidden text it embeds.

- a. Some regulations for PDF documents require all fonts to be embedded, even those used only for hidden text. PDF and PDF/A do not require this and the default behavior of PDF Compressor is not to embed these fonts. To comply with the more restrictive rules, select *Embed fonts for OCR results*.
- b. Select *General input directory* to read XML OCR result files from the same directory as the main input files. Alternatively, choose *Directory* and specify a different source directory.
- c. Under *File name template* you can specify a template to derive the name of each read XML OCR result file from the name of the corresponding input file. See “*Configuring Post-Processing*” (p. 63), for more information on template syntax.



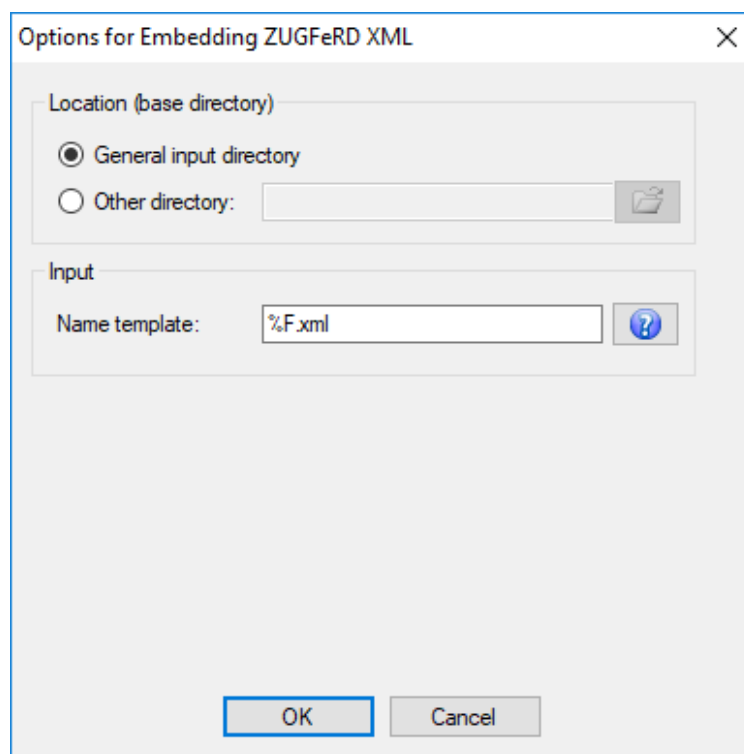
Options for File Embedding

The *File Embedding* group of the *File Embedding Options* dialog lets you configure options for adding files as attachments to your PDF output files.

Note: This feature is not available for PDF/A-1 output. PDF/A-2 documents are restricted to embed PDF/A-1 or PDF/A-2 only. If you want to create PDF/A documents and attach files of any type or create ZUGFeRD compliant invoices the PDF Output Format has to be set to PDF/A-3 (see “*Configuring Data Output*” (p. 25)).

Note: If input post-processing is configured for this job (see also “*Configuring Post-Processing*” (p. 63)) to delete or move input files, it also concerns files to embed - at least if the option *General input folder* is set and they are references per name schema like %F.

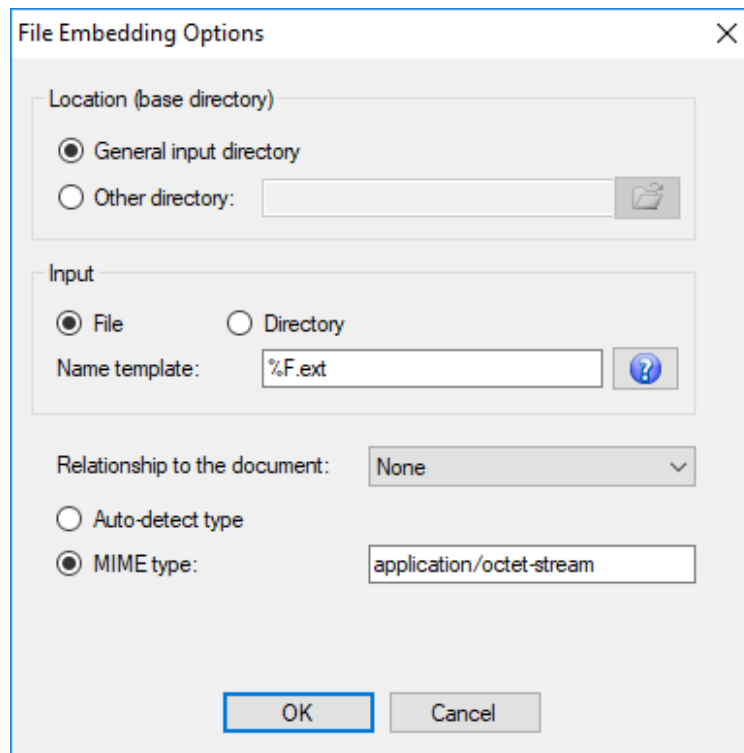
1. Select *Embed input file* to embed the original source file(s) into your output PDFs.
2. The *Embed ZUGFeRD XML* option refers to a German national standard for embedding XML invoice data as attachments into PDF/A-3 files. If this option is selected, PDF Compressor automatically checks and complies with the constraints required by the ZUGFeRD standard.
3. The button *ZUGFeRD Embedding* brings up the following dialog to configure the XML invoice data embedding.



- a. Select *General input directory* to read the XML invoice data files from the same directory as the input files. Alternatively, choose *Other directory* and specify a different source directory.



- b. Under *Input* you can adapt the *Name template* to derive the name of each XML data file attachment from the name of the corresponding input file. See “*Template String Syntax Description*” (p. 73) for more information on template syntax.
4. Select *Embed other file(s)* to embed any other file as attachment into your output PDFs.
5. Clicking on *File Embedding Options* brings up the following dialog with more configuration settings.



- a. Select *General input directory* to read the attachment files from the same directory as the main input files. Alternatively, choose *Other directory* and specify a different source directory.
 - b. Under *Input* select *File* to embed a single specific file per input file. Otherwise select *Directory* to let PDF Compressor embed all files contained in one specific directory per input file. The file or directory name must be configured via the *Name template*.
 - c. The *Name template* lets you specify a template to derive the name of each attachment input file or directory from the name of the corresponding input file. See “*Template String Syntax Description*” (p. 73) for more information on template syntax.
 - d. Select the *Relationship to the document* from the drop-down list. The following values are available:
 - i. *None*: Do not add any relationship information to the output PDFs.

Note: PDF/A-3 requires a relationship be specified for each embedded file. Thus PDF Compressor automatically sets the relationship to Unspecified when this option is selected and compatibility is set to PDF/A-3.



- i. *Unspecified*: To be used when the relationship is not known or can't be described using one of the other values below.
 - ii. *Source*: To be used if the attachment is the original source material for the document.
 - iii. *Data*: To be used if the attachment represents information used to derive a visual presentation - such as for a table or a graph.
 - iv. *Alternative*: To be used if the attachment is an alternative representation of content, for example audio.
 - v. *Supplement*: To be used if the attachment represents a supplemental representation of the original source or data that may be more easily consumable (e.g. A MathML version of an equation).
- a. PDF Compressor can auto-detect the type of the attachment file for many common formats like image files, office files etc. Alternatively, you can explicitly specify the *MIME type* for the attachment. Please use the correct syntax: `type/subtype`.


Configuring Header and Footer


The *Header and Footer Options* let you configure additional text that is added to each output PDF page. This can be used to add labels, e.g. "Draft", page numbers, date and time of the conversion, bates stamps, and other information. The given text can be customized concerning font, font face, color size and position.

For odd and even pages different text and horizontal positions can be specified. If the corresponding text field is empty, no such text will be added. Enable *Even pages same as odd pages* if you want the same text to be added to each page.

1. Enter your text at *Odd pages* and/or *Even pages*. The entered string may contain escape sequences that are replaced by corresponding substitute text once the header or footer is printed onto the page. See "*Configuring Data Output*" (p. 25), for a list of possible



escape sequences. The help button  will display such a list within the PDF Compressor application.

2. The *horizontal position* can be set to right, center or left.
3. The *Vertical position* can be set to top, middle or bottom.
4. The *Font size* can be chosen from the list. Alternatively, you can enter other positive integer numbers. Regular text is 10-12pt, oversized numbers yield an error or lead to absence of the header/footer.
5. The three fonts *Times*, *Helvetica* and *Courier* can be chosen with font faces *roman*, *bold*, *italic* and *bold italic* each.
6. Printed characters can be *filled* or *outlined* with the given color, or they can be *invisible*. Invisible text might be used to allow text searching within the output PDF without changing the original document appearance.
7. Click the color button  to select the font color.

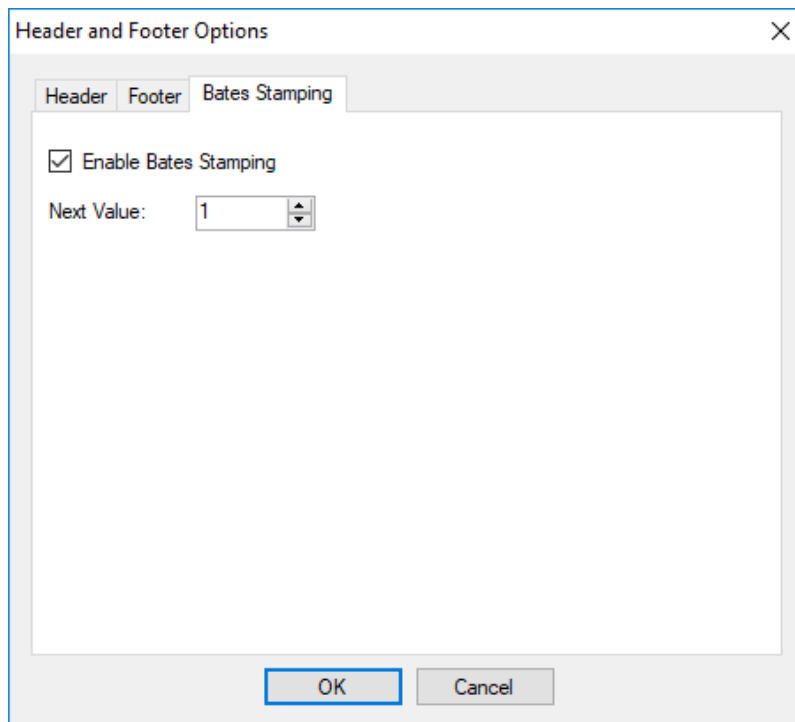
Bates Stamping

By using a bates stamp, you can number every converted page. So a number is printed on every page and automatically increment for every processed page.

Please keep in mind that by using this function, parallel processing of documents will be disabled for such a job. The documents will be processed one after another. This is required to have a consistent numbering in the output.

Configuration

You can use bates stamps within the header or footer, so configuration is also done in the same dialog.



You have to enable the bates stamp feature here to use it.



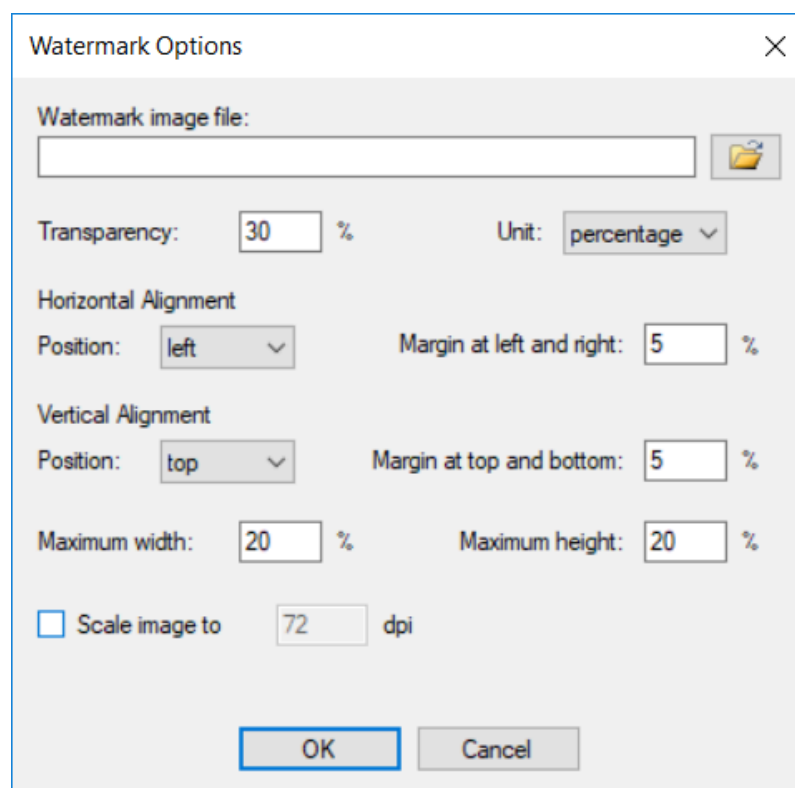
The next value that can be changed within the job settings is automatically increased after a document was successful processed. By that it is ensured, that the same job will not create documents containing the same stamp value. Failed conversions do not affect this value, so you can restart converting such a file without creating gaps in the numbering.


Usage

You can embed the bates stamp by using %s everywhere within a header or footer. By using a modifier, you can also specify the minimum count of digits for such a bates stamp. For example, %05s would be rendered as 00001 if the current number is 1.

Embedding Watermarks

The *Watermark Options* button displays a dialog which lets you add an additional image like a company logo to each of the output pages. The position, size, scaling and transparency can be specified for the image to add.



1. Select the *Watermark image file* to be added to each output page of the PDF file. Use the browse button  to select the file. See “*Choosing the Service Account*” (p. 102) for hints concerning the use of network shares. No watermark will be added if no filename is given. Supported image file formats are the same as listed on the *Input* tab (see “*Configuring Input Data*” (p. 11)).
2. Set the *Transparency* to a value in the range of 0% (opaque) to 100% (invisible).

Note: If you select *PDF/A-1* as output format, transparency is ignored because this format does not allow transparency.

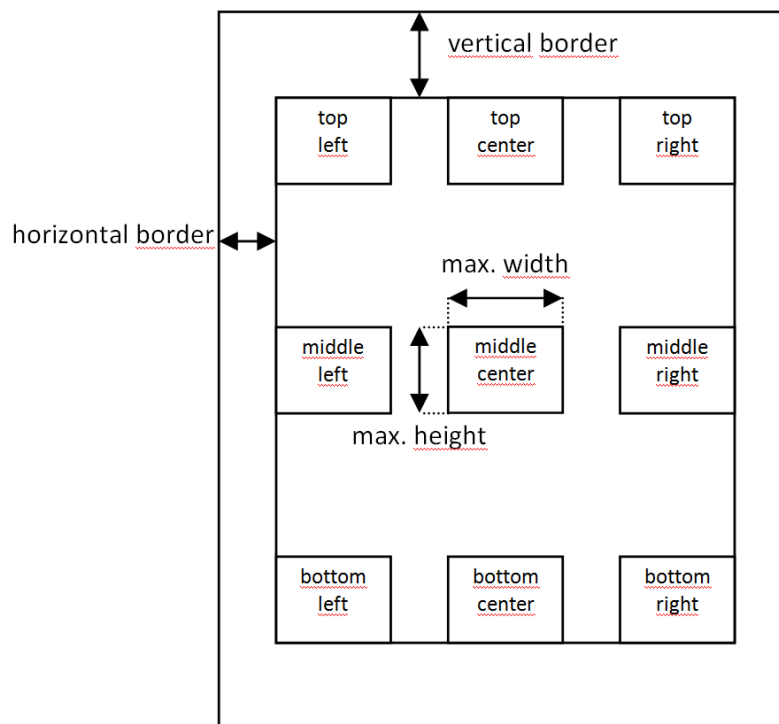
1. Select the *Horizontal* and *Vertical position* to place the image at any edge, any corner or at the center of the page.



2. Set the *Horizontal* and *Vertical Border* to specify the distance between the image and the page edges.
3. Set the *Maximum width* and *height* of the image. Keeping the aspect ratio of the original image, the image will be scaled to fit into a box of the given maximum width and height.
4. Image dimensions and border length can be given in absolute units (select mm), or as a percentage of the page's dimensions. (Horizontal border and maximum width as a percentage of the page width, vertical border and maximum height as a percentage of the page height.)
5. If the watermark image has high resolution, but is added with reduced size, then this may cause unnecessary overhead and larger output file sizes than necessary. If the watermark image is larger than necessary, then it is best to reduce its resolution by selecting *Scale image to* with an appropriate dpi value.

Example: The file *logo.tif* is a 4 by 2 inch image in 300 dpi (1200 x 600 pixels). It is added to the output pages at a size of 1 x 0.5 inch. The resulting effective resolution would be $4 \times 300 = 1200$ dpi. You should set *Scale image to* 75 dpi to get an effective resolution of 300 dpi. The embedded image will only be 300 x 150 pixels, thus causing smaller output files than without scaling.

The following diagram shows how the watermark image placement is calculated:



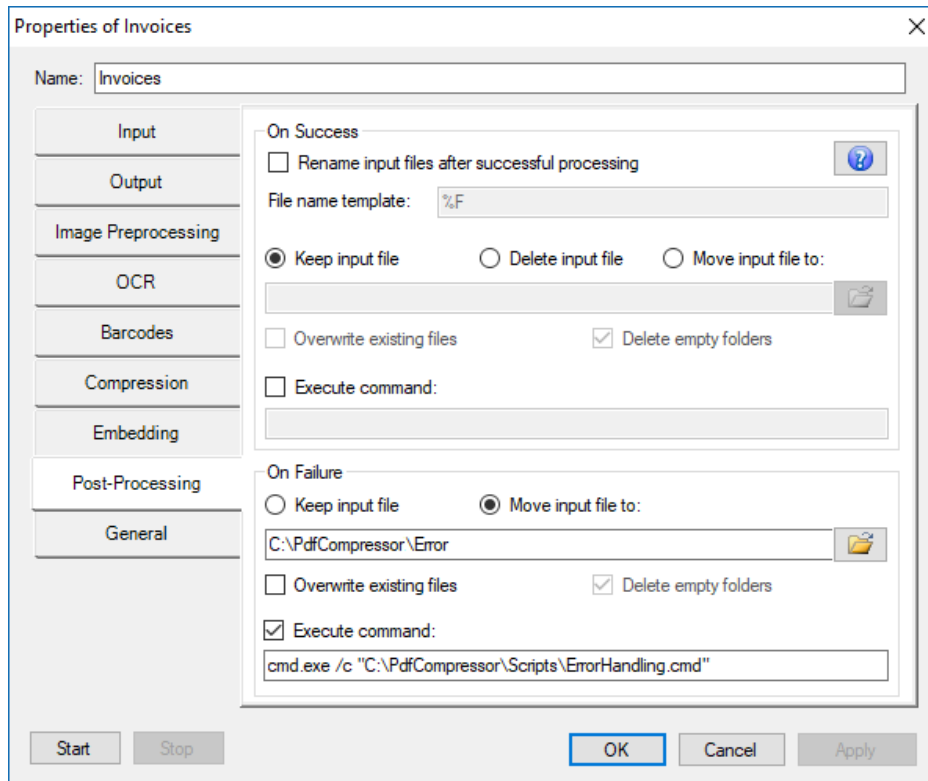
Output page

Configuring Post-Processing

The *Post-Processing* tab lets you configure additional actions to be performed on input files once they have been processed.




The two sections *On Success* and *On Failure* allow to define actions to be applied when either the processing of the input file was successful (a PDF document could be created), or if it failed (no PDF output). If the job was aborted by the user before the processing of an input file has been finished, the input file will not be moved or deleted because a user abort is neither considered as success nor as failure.



The creation of an output PDF file may fail for different reasons. The error message written to the log file (see “*Log Files*” (p. 91)) will give a detailed explanation. Common reasons are:

- The input file does not exist, or the PDF Compressor has not the necessary access rights,
- The output files already exist and overwriting is not allowed,
- PDF Compressor does not have the necessary rights to write or create files in the output folder,
- The input file is corrupted or its format is not supported.

Actions upon Successful Processing


1. *Rename input files* changes the filenames of the input files according to the given File name template. See “*Template String Syntax Description*” (p. 73) for a list of possible escape sequences. The help button  will display such a list within the PDF Compressor application.
2. Input file handling:
 - a. *Keep input file*: The file is left at its location, but may be renamed according to the *Rename input files* setting.



Note: Be careful when using this option. If you use *Check every <number> seconds* and overwrite existing output, you end up converting the same documents over and over. Each conversion will decrease the number of remaining pages in your license (see “*Managing Licenses*” (p. 94)).

- b. *Delete input file:* After the PDF output file is generated, the input file is deleted. Files used as input for embedding as file, bookmarks, metadata or hidden text are also deleted in case the option General input directory and name modifier (%F) are used.

Note: Be careful with this option! With this option, input data is irretrievably deleted. Upon processing failure, the input data will not be available for another processing run!


- c. *Move input file to a folder* sets a target folder to which successfully processed input files are moved. Type a path, use the browse button  or use drag&drop. In case the input directory is scanned recursively, a corresponding directory sub tree will be generated to hold the moved input files. Files used as input for embedding as file, bookmarks, metadata or hidden text are also moved in case the option General input directory and a name modifier (%F) are used. See “*Choosing the Service Account*” (p. 102), for hints concerning the use of network shares.

For this option you can select the behavior in case the destination directory already contains a file with same name as the input file: Either the input file is not moved, or the file within the destination directory will be overwritten.

- *Delete empty folders* is only available if you select *Move input file* and *Include sub-folders* for the input processing (compare “*Configuring Input Data*” (p. 11)). It will clean up the directory sub tree of moved input files once a job has been finished (or a hot folder becomes idle). PDF Compressor always creates required folders in advance, i.e. before any file is processed that needs to be moved into that folder on success. If errors occur such a folder may remain empty. Delete empty folders serves to remove unneeded folders resulting in a more compact file structure.
3. *Execute command:* allows specifying a custom command line to be executed. See section Post-processing command execution below for details.

Actions upon Processing Failure

1. Input file handling:

- a. *Keep input file:* The file is just left at its location (no action on input file).
- b. *Move input file to a folder:* Use the browse button  to select the destination directory where all failed input files should go to. If the input directory is scanned recursively, a corresponding directory sub tree will be generated to hold the moved input files. Files used as input for embedding as file, bookmarks, metadata or hidden text are also moved in case the option General input directory and a name modifier (%F) are used. See “*Choosing the Service Account*” (p. 102) for the use of network shares.



For this option you can select the behavior when the destination directory already contains a file with same name as the input file: Either the input file is not moved, or the file within the destination directory will be overwritten.

- *Delete empty folders*: This option is only available if you select *Move input file* and *Include subfolders* for the input processing (compare “*Choosing the Service Account*” (p. 102), and “*Configuring Post-Processing*” (p. 63)). It will clean up the directory sub tree of the moved input files once a job has been finished (or a hot folder becomes idle). PDF Compressor always creates required folders in advance, i.e. before any file is processed that needs to be moved into that folder on error. If no error occurs such a folder may remain empty. Delete empty folders serves to remove unneeded folders making it easier to locate files in the error folder.

2. *Execute command*: allows specifying a custom command line to be executed. See section Post-processing command execution below for details.

Post-processing *command execution*

PDF Compressor is able to run a custom command after successfully or unsuccessfully processing an input file (or multiple files in case of merging). This is a powerful feature to customize PDF Compressor Jobs to address requirements beyond the built-in post-processing features. Some common use cases are:

- Copy additional files from input folder to output folder
- Take custom action on processing errors, e.g. sending an e-mail
- Access the recognized barcode values
- Custom verification or validation tasks

The PDF Compressor creates a new process for the command. This process runs with the same credentials and permissions as the PDF Compressor service and in the same context, namely in the *Session 0* of the Windows Operating System. Thus, the same restrictions apply as for the PDF Compressor itself.

The process for the command inherits the environment of the PDF Compressor process. Additionally the environment is extended with a set of special variables that provide access to the input and output files and other information (see section *Environment variables* below).

The PDF Compressor redirects the standard output and standard error streams of the process. Any console output written by the command will be appended to the PDF Compressor log file.

Usage

The command must specify the fully qualified path to a valid executable file, optionally with arguments. Ensure to enclose the path in quotes if it contains any blanks.

It is possible to use special environment variables (see *Environment variables*) directly in the command line. Ensure proper quoting: Values of expanded variables may contain blanks. Before execution the PDF Compressor expands any symbols that match the format %Name%. If the variable name is not known, it is replaced with an empty string. To use the percent sign '%' in the command line it must be escaped by another percent sign: %%.



The PDF Compressor evaluates the exit code of the process. Any value different from 0 is considered as error and will be notified in the log file. **Please note:** Since the command is executed after the complete processing in all its aspects an error in the post-processing command will NOT change the processing status to "Failed". In other words, a successful processing will still be considered successful even if the post-processing command fails.

Examples

The following examples should help getting started with a custom post-processing command.

```
cmd.exe /c "C:\MyScripts\postproc.bat"
```

This line executes a batch script as a custom command.

```
powershell.exe "C:\MyScripts\postproc.ps1"
```

This line executes a Windows PowerShell script as a custom command.

```
C:\ToolDir\MyTool.exe "%FX_OutputFilePath0%"
```

This line runs the MyTool.exe as custom command and passes the output file as the first argument.

```
cmd.exe /c set
```

This line prints the whole set of environment variables to the log file. This can be useful to get an idea how the environment looks for the command.

```
cmd.exe /c echo %FX_BarcodeValue0% > "%FX_OutputFilePath0:~0,-4%.txt"
```

This might be the simplest way to store a recognized barcode in a text file next to the regular output file. Note the use of double percent characters to ensure that the `FX_OutputFilePath` variable is not already expanded by PDF Compressor. Instead it is expanded later by the command interpreter (cmd.exe) allowing to use the substring syntax (`:~0,-4`). In a batch script enable delayed expansion and use exclamation marks instead of percent characters.

For production it is recommended to add some error checking, e.g. if a barcode was detected at all, and to put everything in a reusable batch script.

Environment variables

The following table lists the environment variables that are provided by the PDF Compressor to the command.

| Variable name | Description |
|-----------------------|---|
| %FX_ProcessingStatus% | <i>Success</i> or <i>Failure</i> . The value is set to <i>Success</i> if processing of the input file(s) succeeded or to <i>Failure</i> if an error occurred. |
| %FX_ErrorMessage% | This variable is only available if the processing failed. It contains a description for the error that happened. |
| %FX_JobName% | The name of the Job that is processed. |
| %FX_InputFileCount% | The number of input files that were processed. Typically, the value is 1 unless the job has the "Merge all pages of a folder to a single PDF document" flag set or additional input files specified on the "Embedding" tab. |



| Variable name | Description |
|--|---|
| %FX_InputFilePath0% %FX_InputFilePath1% ... | <p>The fully qualified paths to the processed input files.</p> <p>There will be %FX_InputFileCount% many entries, counted from 0 to (%FX_InputFileCount% - 1). The order in the list is the processing order of the files. Additional input files (e.g. for embedding) are listed directly after the main input file to which they belong.</p> <p>Note: The command is executed after the complete processing which includes the built-in post-processing. Thus the provided paths can be different from the original input paths when move or renaming actions were applied. If the job is configured to delete the input files after successful processing the custom command will be provided with the original input file paths but at the time the command is executed these files have already been deleted.</p> <p>Note: For standard jobs there is exactly one input file and the variable %FX_InputFilePath0% contains the fully qualified path (see also description of %FX_InputFileCount%).</p> |
| %FX_OutputFileCount% | <p>The number of output files that were generated. Typically the value is 1 unless the job produces additional output files like OCR results as XML or TXT, thumbnail images, or the output splitting option is enabled.</p> <p>If the processing failed the number of output files is always 0 because the PDF Compressor removes incomplete output.</p> |
| %FX_OutputFilePath0% %FX_OutputFilePath1% ... | <p>The fully qualified paths to the generated output files.</p> <p>There will be %FX_OutputFileCount% many entries, counted from 0 to (%FX_OutputFileCount% - 1). The order in the list is the order of generation. Additional output files (e.g. OCR results) are listed directly after the main output file to which they belong.</p> <p>If the processing failed there are no output files at all because the PDF Compressor removes incomplete output.</p> <p>Note: For standard jobs there is exactly one output file and the variable %FX_OutputFilePath0% contains the fully qualified path (see also description of %FX_OutputFileCount%).</p> |
| %FX_OutputPageCount% | <p>The number of pages of the output files that were generated.</p> <p>If the processing failed the number of pages of the output files is always 0 because the PDF Compressor removes incomplete output.</p> |



| Variable name | Description |
|--|--|
| %FX_BarcodeCount% | The number of recognized barcode values. This variable is only set if barcode recognition was activated. %FX_BarcodeCount% is also provided. |
| %FX_BarcodeValue0% %FX_BarcodeValue1% ... | The list of recognized barcode values. There will be %FX_BarcodeCount% many entries, counted from 0 to (%FX_BarcodeCount% - 1). The order in the list is the order of detection. |
| %FX_BarcodeType0% %FX_BarcodeType1% ... | The types of the recognized barcodes. The types are similar to the names in the dialog but simplified to uppercase letters and digits. |
| %FX_BarcodePage0% %FX_BarcodePage1% ... | The page numbers in the input file where each barcode was recognized. Page numbers start at 1. In case of multiple input documents, the page number keeps counting, as if they were a single concatenated document. |
| %FX_BarcodeProfile0% %FX_BarcodeProfile1% ... | The name of the profile of the barcode. Note: A barcode can be matched by multiple profiles and thus reported several times. |

Configuring General Job Properties

The *General* tab lets you configure general properties of the selected job. You can define the priority of the job here and you can configure a delayed start and periodical stop and restart of the job.

Properties of Invoices

Name: Invoices

Input

Output

Image Preprocessing

OCR

Barcodes

Compression

Embedding

Post-Processing

General

Job Priority: 5

Time Scheduled Processing

☒ Start job at: Fr, 30. Nov 2018 19:49

☒ Duration: 5 Minutes

☐ Restart interval: 0 Days

☐ Processing timeout: 0 Seconds

Advanced Job Options

Start Stop OK Cancel Apply



Job List Processing Order


The order in which jobs are processed may be set in different manners: One option is to define *Job Priorities* and to enable *Priority Processing Order*, another is to run through the job list in order of appearance. In either case jobs are sliced into *job units*.

A *job unit* contains all input files that contribute to one output PDF document¹. If *Merge all pages of a folder to a single PDF document* is enabled (see “*Configuring Input Data*” (p. 11)), this includes all files that are merged into a single document, otherwise it is just one single input file.

Hot folder jobs (see “*Configuring Input Data*” (p. 11)) are considered done when the respective input folder is empty. Once new input files are detected, a hot folder job becomes active again and is processed according to its priority setting.

Priority Processing Order

If *Priority Processing Order* is enabled jobs with higher priorities are processed before lower priority jobs. All jobs with the same priority are executed concurrently by processing its job units in interleaved order: Once a unit is done, the next unit of the next job with same priority is processed. When more than one CPU core is licensed and enabled, more than one job unit will be processed at the same time. The list of progress bars located in the lower part of the main window shows information for each CPU core, including job units that are currently processed.

If *Priority Processing Order* is enabled the *Job Priority* may be set. Small numbers reflect high priority whereas larger number reflect low priority. Hence, 1 is highest and 10 lowest priority. In order to switch *Priority Processing Order* on or off, all jobs must be paused. Once done, navigate to the menu *File* → *Priority Processing Order* or use the toolbar button .

Priority Processing Order allows setting up complex job dependencies. High priority hot folder jobs can be used to dynamically suspend other lower priority jobs. The priority of a job can be changed at any time, even when it is running. Thus it allows you to start an interim job to be processed immediately without stopping a currently running job.

When processing in parallel using several cores, a better utilization is possible by changing the global setting *Complete all jobs of highest priority before starting further jobs* under *File* → *Options...* → *Configuration* → *Priority Processing*. If this option is activated, free cores are only assigned to a job if all jobs with higher priority have already been processed. Please find more information on this setting in the Administration chapter under “*Priority Processing*” (p. 94).

List Processing Order

Priority Processing Order is the preferred way to control the order in which jobs are processed. For backward compatibility *List Processing Order* is still available. To enable *List Processing Order*, disable *Priority Processing Order* from the *File* menu.

In *List Processing Order* mode, all jobs are processed in the order they appear in the job list. Priorities are ignored, and jobs are executed sequentially. There is no way to dynamically suspend a running job.

¹Please note that a PDF input file might be split into multiple output files. For further information, see “*Configuring Data Output*” (p. 25).



Time Scheduled Processing and Processing Timeout

Settings for *Time Scheduled Processing* are enabled by selecting their check boxes. The following parameters can then be set for a time-scheduled job:

1. The desired start time is entered in the *Start job at* field. You can either type in a date and time or pick it from the calendar widget, which appears you click on the arrow right next to the entry. Processing this job will start at this point in time.
2. The *Duration* specifies how long the job should run. You can select days, hours or minutes to specify the duration. Once this time has elapsed the job is stopped.
3. The *Restart interval* denotes a second time span, after which the job is started again and runs for the given duration. This value can only be given in days. Once the job has been started, it is restarted periodically, unless the configuration is changed.

Note: The *Duration* and *Restart* interval settings can be enabled independently. This means on the one hand that a job can run just once for a given duration without restart. And on the other hand, a job can be periodically restarted, without a fixed duration, assuming that it will finish once it has processed all its current input.

With the *Processing Timeout* option, you can set a maximum duration for the processing of a job unit. (The definition of job unit is given above under “*Priority Processing*” (p. 94).) If processing of a job unit takes longer than the specified duration, the processing of this unit is aborted. Such an abort is considered an error, which is recorded in the log file. Concerning optionally configured post-processing steps, an abort is also treated as an error.

Configuring Advanced Job Options

The *Advanced Job Options* dialog lets you configure additional error handling and Foxit Europe maintenance settings.



The image shows the 'Advanced Job Options' dialog box with three sections: 'Error Handling', 'Audit Log', and 'Foxit Maintenance'. The 'Error Handling' section has a checkbox for 'Continue job on critical error'. The 'Audit Log' section has an 'Activate' checkbox, a 'Client ID' text field, and an 'Audit Log File Path' text field with a folder icon button. The 'Foxit Maintenance' section has a warning message and a 'Key' text field. At the bottom are 'OK' and 'Cancel' buttons.

1. If the option *Continue job on critical error* is turned on, PDF Compressor will not stop a hot folder when a critical error occurs (such as 'input folder does not exist' or 'output folder cannot be created') but to go idle instead and try again later.
2. For information regarding please refer to the following chapter “*Audit Log*” (p. 72).
3. The *Foxit Maintenance* entry should not be used unless advised by the Foxit Europe Support Team. The *Key* must be left blank for normal processing.

Audit Log

The PDF Compressor can create audit logs per job. The audit logs provide information about which input was processed and when, which output was generated, and if the processing was successful. An audit log has the following format:

```
This audit log file was created by PDF Compressor 8.x on Mon
Jan 07 15:18:13 2019
```


```
Finished,Started,Duration_ms,Input,Output,Pages,Client_Code,Bates,
Status
2019/01/07 15:18:13,2019/01/07 15:17:37,35345,\\fileserver\input\
pages.tif,\\fileserver\output\pages.pdf,10,My Client,,OK
```

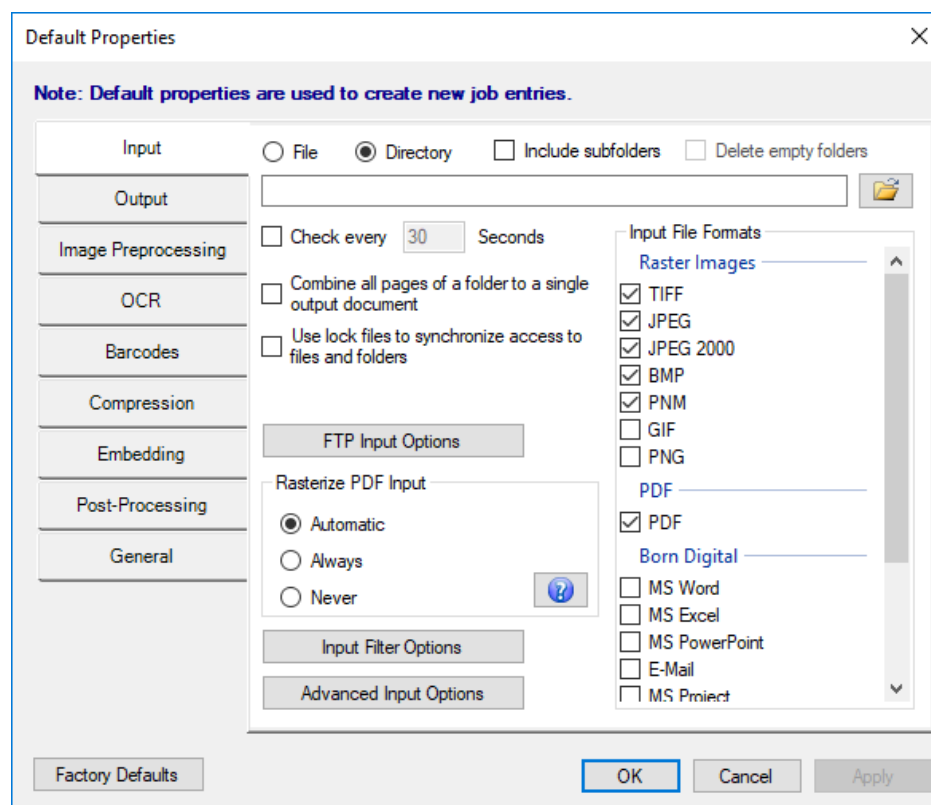
The audit log is configured per job at the *General* tab in the *Advanced Job Options* in the *Audit Log* group. The *Activate* checkbox enables or disables the audit log. The *Client ID* is an arbitrary text string which will be written to the audit log into the "Client_Code" column. The *Audit Log File Path* must be an absolute file path to a location where the PDF Compressor Service has permissions to create and modify files. Multiple jobs, probably running in parallel may share the same audit log file.



Setting up the Default Properties

Each job list entry has its own properties settings. Additionally, there is a set of properties which is not related to any specific job list entry: the *Default Properties*. These are used as a template whenever you create a new entry in the list. You can open the *Default Properties* dialog by any of the following actions:

- Use the default properties button  in the toolbar,
- Use the menu *Entry* → *Edit Default Properties*,
- When a properties dialog of an entry is already opened: Unselect the currently selected entry by clicking within the list window outside any entry.



As depicted above, the *Default Properties* dialog is similar to a job entry *Properties* dialog. The only differences are, that there is no name to be edited and of course, the *Start* and *Stop* buttons are missing as there is no related job to run. In place of the *Start* button you find a reset button which lets you restore *Factory Default* values for all parameters.

Template String Syntax Description

Template strings are used for various purposes. They are used for PDF metadata (see “*PDF Metadata*” (p. 34)), headers and footers (see “*Configuring Header and Footer*” (p. 60)), input file renaming (see also “*Configuring Post-Processing*” (p. 63)), output file naming (see “*Configuring Data Output*” (p. 25)), thumbnail file naming (see “*Thumbnails*” (p. 35)), data embedding and file embedding (“*Configuring File and Data Embedding*” (p. 54)).

Restrictions



| Used in | Restriction |
|---|---|
| PDF Metadata | Must not contain %Q, %s or its versions with modifiers. |
| Header and Footer | Must not contain %L, %N, %C, %V or its versions with modifiers. |
| Input file renaming | Must contain %F or its version with modifiers, or regular expression substitution for F. Must not contain %P, %Q, %L, %N, %C, %V, %s or its versions with modifiers. Must not contain any characters disallowed for file names. |
| XMP metadata, PDF bookmarks, XML OCR results, file attachments: | Must not contain %Q or its versions with modifiers. Must not contain any characters disallowed for file names. May be constant strings without any % escape sequences. |
| Output file naming | Must at least contain one of %P, %L, %C or its versions with modifiers when page-based output splitting is enabled. Must at least contain one of %P, %L, %C, %V or its versions with modifiers when barcode triggered output splitting is enabled. Must not contain %Q, %s or its versions with modifiers. Must not contain any characters disallowed for file names. |
| Thumbnail file naming | Must not contain %L, %N, %s or its versions with modifiers. Must not contain any characters disallowed for file names. |

Escape Sequences

| Escape Sequence | Substitution |
|-----------------|--|
| %% | Percent sign "%" |
| %F | Input file or directory name. File name in case of single file input processing. When merging files from a folder it may also represent the folder name. In header and footer as well as in file name templates for additional input %F always prints the file name and not the name of the folder. The file name is printed without extender in file name templates and PDF metadata. |
| %E | Input filename extension or empty string for directory. |
| %P | First page number that is output to the PDF file (metadata or output file name) or Current page number in output document (header, footer, or thumbnail output) |
| %Q | Current page number in output chunk |
| %L | Last page number that is output to the PDF file |
| %N | Number of pages output to the PDF file |
| %s | Bates stamp |
| %C | Number of PDF file when using output splitting |
| %V | Value of the last detected barcode (from the first profile only) |
| %a | Abbreviated weekday name |
| %A | Full weekday name |



| Escape Sequence | Substitution |
|-----------------|--|
| %b | Abbreviated month name |
| %B | Full month name |
| %c | Date and time representation appropriate for locale |
| %d | Day of month as decimal number (01 - 31) |
| %H | Hour in 24-hour format (00 - 23) |
| %I | Hour in 12-hour format (01 - 12) |
| %j | Day of year as decimal number (001 - 366) |
| %m | Month as decimal number (01 - 12) |
| %M | Minute as decimal number (00 - 59) |
| %p | Current locale's A.M./P.M. indicator for 12-hour clock |
| %S | Second as decimal number (00 - 59) |
| %U | Week of year as decimal number, with Sunday as first day of week (00 - 53) |
| %w | Weekday as decimal number (0 - 6; Sunday is 0) |
| %W | Week of year as decimal number, with Monday as first day of week (00 - 53) |
| %x | Date representation for current locale |
| %X | Time representation for current locale |
| %y | Year without century, as decimal number (00 - 99) |
| %Y | Year with century, as decimal number |
| %Z, %z | Time-zone name or abbreviation; no characters if time zone is unknown |
| %#c | Long date and time representation, appropriate for current locale. For example: "Tuesday, March 14, 1995, 12:41:29". |
| %#x | Long date representation, appropriate to current locale. For example: "Tuesday, March 14, 1995". |
| %[, %], %?, %: | Used for conditional and regular expression substitutions. See "Conditional Substitution" (p. 76) and "Regular Expression Substitution" (p. 77). |

Moreover, adding the character '#' to shortcuts that output numbers, will remove leading zeros in most cases.

Modifiers

| Escape Sequence | Modifier | Function |
|-----------------|-------------|---|
| %P | %0<number>P | Minimal number of output characters. This modifier defines the number of preceding zeros used to format a page number. The modifier starts with the digit '0' and is followed by one or more digits, represent- |
| %Q | %0<number>Q | |
| %L | %0<number>L | |
| %N | %0<number>N | |
| %C | %0<number>C | |



| Escape Sequence | Modifier | Function |
|-----------------|---------------------|---|
| %s | %0<number>s | ing the minimal number of output characters as a decimal number. |
| %F | %<number>-<number>F | Defines a substring of the given file name This modifier is composed by two integer numbers separated by a minus sign '-'. If the modifier is present, only a substring of the filename will be substituted. The numbers represent the first and the last character index (counting from 0) of the substring. Any number greater or equal to the string length will be adjusted to represent the last character of the string. |

Modifier Examples

| Escape Sequence | Page Number | Substitution |
|-----------------|---------------------|-------------------------------|
| %02P | 1 | 01 |
| %02P | 9 | 09 |
| %02P | 10 | 10 |
| %02P | 100 | 100 |
| Escape | File name | Substitution |
| %3-10F | 01_myfile01_xyz.tif | myfile01.pdf |
| %3-99F | 01_myfile01_xyz.tif | myfile01_xyz.pdf |
| %50-60 | 01_myfile01_xyz.tif | z.pdf |
| %10-3F | 01_myfile01_xyz.tif | <i>considered as an error</i> |

Examples

When a 10 pages document "MyFile.tif" is output into chunks of 4 pages, the output file template

```
"%F, Part %C, Pages %02P - %02L"
```

will output the files

```
"MyFile, Part 1, Pages 01 - 04.pdf"
```

```
"MyFile, Part 2, Pages 05 - 08.pdf"
```

```
"MyFile, Part 3, Pages 09 - 10.pdf"
```

Conditional Substitution

The following syntax can be used to make a template string substitution dependent from a condition:

| Condition Syntax | Substitution |
|---------------------------------------|-----------------------------------|
| %[<condition>%?<string1>%:<string2>%] | <string1> if <condition> is true, |



| Condition Syntax | Substitution |
|------------------|---------------------|
| | <string2> otherwise |

Currently the following conditions are defined:

| <condition> | Description |
|-------------|--|
| C | True if the number of output chunks is greater than 1 or unknown. False otherwise. |
| V | True if the value of the last detected barcode (%V) is not empty (a barcode was detected). |

Examples

Output Chunks

The template

```
"%F, %[C%?Part %C, %:~]Pages %02P - %02L"
```

outputs the 10 pages document "MyFile.tif" with a chunk size of 5 pages to

```
"MyFile, Part 1, Pages 01 - 05.pdf"
"MyFile, Part 2, Pages 06 - 10.pdf"
```

and the 5 pages document "MyShortFile.tif" with the same chunk size to

```
"MyShortFile, Pages 01 - 05.pdf"
```

Barcode Renaming

The template

```
"%[V%?%V%:%F%]"
```

names the output file according to the value of the barcode found in the document (%V). If barcode detection fails for a document however, the output is named like the input file (%F) as a fallback.

Barcode Splitting

The template

```
"%F - %[V%?%V%:PREFIX%] (%02P - %02L) "
```

outputs for a document "MyFile.tif" with barcodes "Chapter 1" on page 3 and "Chapter 2" on page 7 the following chunks:

```
"MyFile - PREFIX (01 - 02).pdf"
"MyFile - Chapter 1 (03 - 06).pdf"
"MyFile - Chapter 2 (07 - 10).pdf"
```

Regular Expression Substitution

The regular expression substitution syntax is defined by:

```
%[<input>%?<regex>%:<replacement>%]
```



<input> identifies the input string the regular expression substitution is applied to:

| <input> | Description |
|----------------------|--|
| F | Input file (or directory) name. File name for single file input processing, directory name when merging files from a directory into a single PDF output file. The file name is printed without file extension for PDF metadata, input file renaming, and output file renaming. |

<regexp> is the regular expression. It implements the commonly used syntax for regular expressions defined as follows:

A **regular expression** is zero or more branches, separated by '|'. It matches anything that matches one of the branches.

A **branch** is zero or more pieces, concatenated. It matches a match for the first, followed by a match for the second, etc.

A **piece** is an atom possibly followed by '*', '+' or '?'. An atom followed by '*' matches a sequence of 0 or more matches of the atom. An atom followed by '+' matches a sequence of 1 or more matches of the atom. An atom followed by '?' matches a match of the atom, or the null string.

An **atom** is a regular expression in parentheses (matching a match for the regular expression), a range (see below), '.' (matching any single character), '^' (matching the null string at the beginning of the input string), '\$' (matching the null string at the end of the input string), a '\' followed by a single character (matching that character), or a single character with no other significance (matching that character).

A **range** is a sequence of characters enclosed in '[']. It normally matches any single character from the sequence. If the sequence begins with '^', it matches any single character not from the rest of the sequence. If two characters in the sequence are separated by '-', this is shorthand for the full list of ASCII characters between them (e.g. '[0-9]' matches any decimal digit). To include a literal ']' in the sequence, make it the first character (following a possible '^'). To include a literal '-', make it the first or last character.

<replacement> is the replacement string to be used to build up the output string if the regular expression found a successful match. (If the regular expression has no match, the output string will be a copy of the input string.) In case of a successful match, each character in the replacement string will be copied to the output string except for the following special characters:

| Sequence | Output |
|-----------------|-----------------------------|
| & | The complete matched string |
| \1 | Matched sub-string 1 |
| | and so on until ... |
| \9 | Matched sub-string 9 |

Regular Expression Substitution Examples

| Description | regular Expression |
|--|---------------------------|
| A single number in a range between 1 through 5 | [1-5] |



| Description | regular Expression |
|--|--------------------|
| Exclude number 5 in a range between 0 through 9 | [0-46-9] |
| Combination of a capital, a number in a range between 2 through 7 and a lower-case letter | [A-Z][2-7][a-z] |
| Combination of a range between 1 through 9 and an optional letter | [1-9][a-zA-Z]? |
| Combination of multiple number ranges | [0-2][0-9][0-9] |
| Zero or multiple number ranges | ([0-9])* |
| One or multiple number ranges | ([0-9])+ |
| Matching first and one of the following numbers in a range | ([0-1])[0-1]*\1 |
| Combination of a capital letter at the beginning, three subsequent arbitrary characters and a number in a range between 0 through 7 at the end | ^[A-Z]...[0-7]\$ |

Note: Using curly braces "{}" to signify range repetition is not allowed. Each range has to be denoted explicitly.

The template

```
"%[F%?(.*)_From_[0-9]+_To_[0-9]+$%:\1%]_%P_To_%L"
```

outputs the 10 pages document "MyFile_From_1_To_10.tif" with a chunk size of 5 pages to

```
"MyFile_ 1_To_5.pdf"
```

```
"MyFile_ 6_To_10.pdf"
```



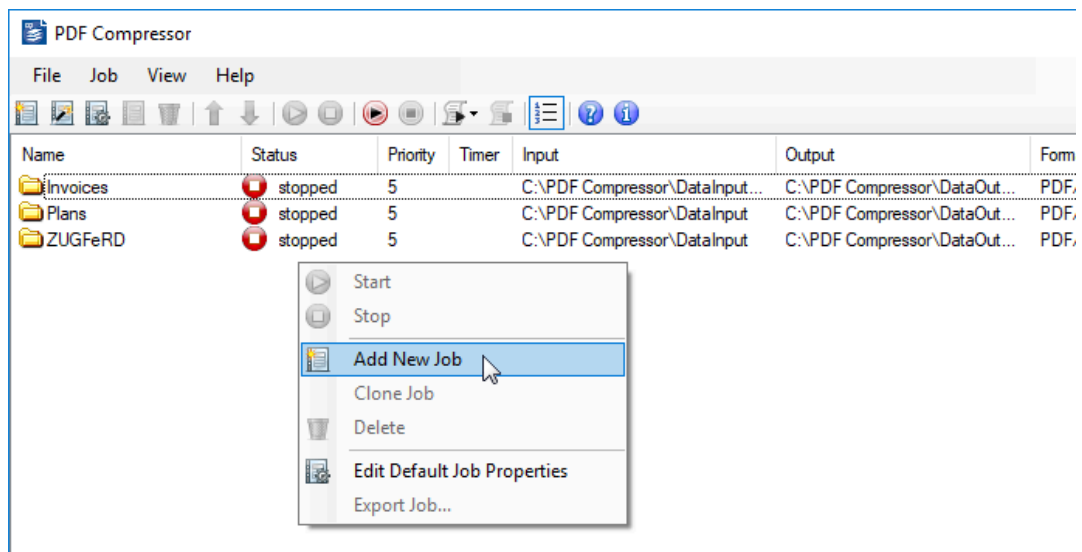
Chapter 6. Managing the Job List


This section describes in detail all operations that can be performed on the job list.

Adding, Deleting and Copying Entries

New entries can be added by any of the following actions:


- Using the context menu within an area of the list that is outside any existing entry.



- Use the *new* button  in the toolbar.
- Use the menu *Entry* → *Add New Entry*.
- Drag & drop files and/or folders from the Windows Explorer to the list window.

When a new entry is created, all settings from the Default Properties are copied to the new entry. The name is set to "Entry <num>", where <num> is the first available number. When the entry is created by drag & drop the input path is set accordingly. The output path defaults to the input path in this case.

Entries can be deleted by any of the following actions:

- Use the context menu on top of the entry to be deleted. Choose *Delete*.
- Select an entry and use the *delete* button  in the toolbar.
- Select an entry and use the menu *Entry* → *Delete*.

All entries can be deleted by use of the menu *File* → *Delete All Job Entries*. Sometimes it is useful to create a copy of an existing entry (e.g. you might only want to change one conversion option and the output folder to compare the results).

An entry can be copied by:



- Using the context menu on top of the entry to be copied. Choose *Clone Job*.



- Selecting an entry and using the menu *Entry → Clone Job*.

Changing the Sequence of Job Entries

To change the order of the list, you can move entries up and down by:



- Selecting an entry and use the *up*  and *down*  buttons in the toolbar.
- Selecting an entry and use the menu *Entry → Move Up* or *Entry → Move Down*.

See “*Priority Processing*” (p. 94), for details on the processing order of job list entries.

Starting and Stopping Job Entries



Individual Entries

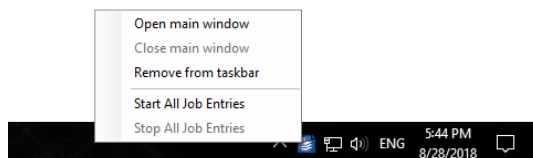
A selected entry can be started and stopped (aborted) by any of the following actions:

- Use the context menu of the entry to be started or stopped. Choose *Start* or *Stop*.
- Select an entry and use the *start*  and *stop*  buttons in the toolbar.
- Select an entry and use the menu *Entry → Start* or *Entry → Stop*.
- Use the Start and Stop buttons of the entries Properties dialog.

All Entries



All entries can be started or stopped simultaneously by any of the following actions:

- Use the toolbar buttons  or .
- Use the menu *File → Start All Job Entries* or *File → Stop All Job Entries*.
- Use the context menu of the PDF Compressor taskbar icon (right mouse click) and choose *Start All Job Entries* or *Stop All Job Entries*.









Monitoring Jobs

The status of all jobs is listed within the job list. It can be one of the following:

| Status | Description |
|--|--|
|  no service | The underlying PDF Compressor service is stopped or has not been installed. See “ <i>Managing the PDF Compressor Service</i> ” (p. 101) for details. |
|  stopped | The job is stopped. No file conversion will take place. |



| Status | Description |
|--|---|
|  starting | A signal has been sent from the GUI to the PDF Compressor service to start the job. The status will soon change to started or working. |
|  working | The job has been started, and a file conversion is currently performed. |
|  working | The job has been started and is within the group of same priority jobs that are about to be processed in an interleaved order. |
|  monitoring | The job is a hot folder that is idle and waiting for new files to appear. |
|  running | The job has been started and is waiting for other higher priority jobs to finish before it becomes working itself. In case <i>priority processing</i> is switched off, the job is waiting for preceeding jobs to finish working. The job will start after all preceeding jobs in the list are finished. See also “ <i>Priority Processing Order</i> ” (p. 70). |
|  stopping | A signal has been sent from the GUI to the PDF Compressor service to stop the job. The status will soon change to stopped. |

Importing and Exporting Job settings

The settings of entries within a job list can be exported to a file or restored from a file by use of the menus:



- *Entry* → *Export job...* (single entry)
- *File* → *Export job list...* (all entries)
- *File* → *Import job list...* (overwrites existing)
- *File* → *Append job list...* (imports a job list by appending its entries to the current list)

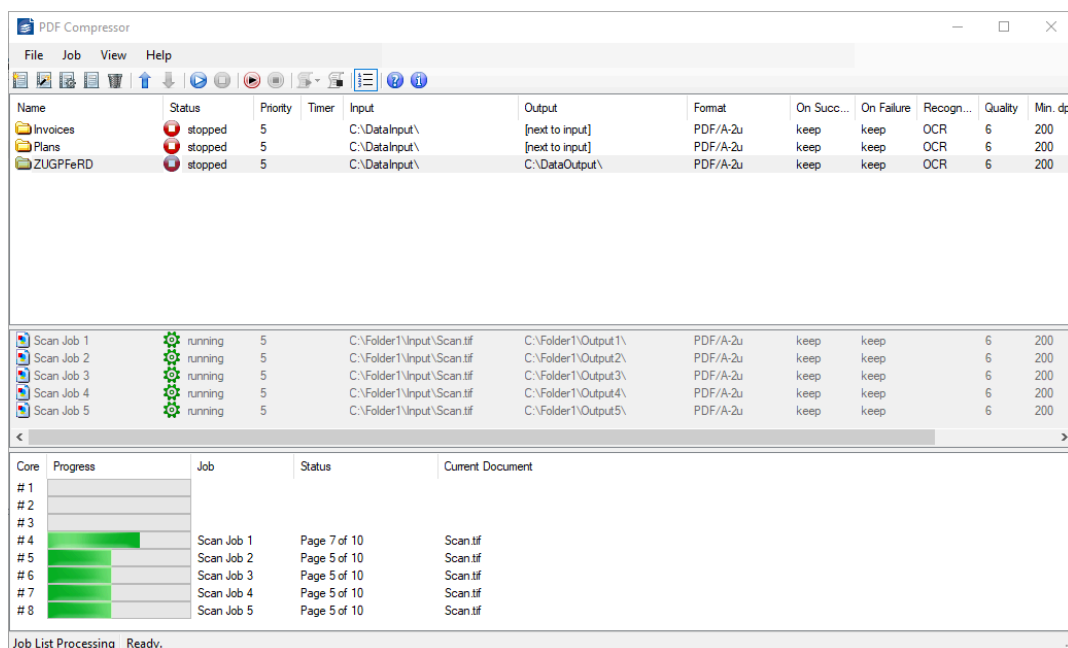
This can be used for backup purposes. *Append job list...* can be used to merge multiple job lists into a single list. Settings of a single job can also be exported by right-clicking on the job entry and selecting *Export Job...* from the context menu.



Chapter 7. Automatic Job List Processing

PDF Compressor can automatically load and process predefined job lists. Job lists are read from plain text files that can be created by external applications. Please refer to “*Job List File and Control File Syntax*” (p. 84), for a description of the syntax. Together with corresponding control files this enables PDF Compressor to integrate with many different workflows.

Automatic job list processing is either started by the menu *File* → *Start Automatic Processing* → *Start Job List Processing* or using the toolbar button . Use the menu *File* → *Stop Job List Processing* or the button  to leave the job list processing mode.



Job list processing is indicated in the status bar. In this processing mode the job list is being split into two parts. Automatically created jobs will show up in the lower part of the list. They can only be monitored. It is not possible to control (start/stop) or modify jobs created by the PDF Compressor service from within the GUI.

Finding and Processing Job Lists

The PDF Compressor service looks for job list files and its associated control files within the folder that is specified by the *Options* dialog (see “*Job List Processing*” (p. 93)). Whenever it finds one or more control files *.go it starts to process the oldest of these control files (timestamp). This implements a first in, first out strategy. It is also possible to set priorities for each job list, as described in “*Job List File Order and Priorities*” (p. 84).

When a file

```
list0001.go [signal a job list to be processed]
```

is going to be processed, PDF Compressor tries to load a job list from the file

```
list0001.dat [the job list to be processed]
```





You will see the corresponding jobs in the PDF Compressor list window. The format of the *.dat files are the same as used for job list import and export.

During processing, the control file *.go is renamed to *.wrk. When all the entries of a list have been processed without any error, the control file *.wrk is renamed to *.rdy:

```
list0001.rdy [processing is ready, all files succeeded]
```

If an error occurred with any of the processed files, the control file is renamed to *.err:

```
list0001.err [processing is ready, an error or abort occurred]
```

If the stop button  is pressed while there are still running entries in the list, all entries are aborted immediately, and the control file signals an error (*.err). This ensures that you will only find list0001.rdy if all files have been successfully processed.

Job List File Order and Priorities

After a list has been loaded, all its entries are automatically started. As soon as an entry has been finished (and is shown as stopped) it is removed from the list.¹ The next job list is loaded when at least one processor core becomes idle and there are no jobs left waiting to get processed. Thus, jobs of different lists may be processed in parallel to ensure all available processor cores run at highest possible capacity.

To enforce pure sequential processing of job lists and the contained jobs, the processing order can be set to *List Processing Order* mode (compare “*Job List Processing*” (p. 93)). If this option is chosen, the next job list file is not loaded until processing all jobs from the previous file has been completed.

Instead of assigning priorities to individual jobs, in *Automatic Job List Processing* mode the whole job list may be assigned a priority. Individual job priorities are ignored. To assign a priority to a job list, append the respective value to the file extension: `samplelist.go` becomes `samplelist.go2` to set priority to 2. As for individual jobs, lower values indicate higher priority, wherefore `.go1` means highest and `.go10` lowest priority. Any file with a plain `.go` extension is given a priority between `.go5` and `.go6`. Among files of the same priority the oldest ones are processed first.

The option *Complete all jobs of highest priority before starting further jobs* does not affect job list processing.

Job List File and Control File Syntax

The Control File has the same file name as the corresponding Job List file. The extension is `.go` (see details above at “*Finding and Processing Job Lists*” (p. 83) and “*Job List File Order and Priorities*” (p. 84)). The control file is a plain text file with no content. You can create them with any ASCII editor manually.

Job list files *.dat are plain text files, too. You can view and edit them in a any text editor. It is recommended to work with automatically generated job list files as follows:

Set up your job list within PDF Compressor first. Export the job list to a file. Modify only the basic entries explained in the list at the end of this section.

A job list file has the following structure:

¹Hotfolder jobs will run as normal jobs and only check for input files once at start. Jobs need to stop automatically for automatic processing.



A line with the keyword

```
FOXIT_PDF_COMPRESSOR_ENTRY_START_UTF8 <num>
```

starts a new entry in the list. <num> is the number of the entry counting from zero. This number is written when exporting a list but is ignored for import. The "_UTF8" flag serves to clarify the text encoding. It is highly recommended to use Unicode UTF8 encoding. UTF-8 will properly represent all sorts of special characters such as accented vowels, German umlauts and Slavic or Cyrillic characters in paths, file names and other text strings.

As a fallback, the keyword

```
FOXIT_PDF_COMPRESSOR_ENTRY_START_ANSI <num>
```

can be used in cases where use of a Unicode compliant encoding is impossible and an ANSI encoding must be used instead.

Properties of a job list entry are given by a number of lines, each of the form

```
<property_name> <value>
```

A <property_name> is an ASCII string defined by PDF Compressor. It does not contain any white space. PDF Compressor reads known names, and it ignores lines with unknown names.

A <value> is either a decimal integer like 300 or a string in quotation marks like "Entry 01". Note that strings do not escape any contained quotation character ". A string begins after the first occurrence of the character " in a line and ends just before the last occurrence of the character ". Examples:

```
name "Entry 01" # set the name of the entry to Entry 01
name "Entry "01"" # set the name of the entry to Entry "01"
```

Line feeds (hex 0A), carriage returns (hex 0D) and the character # are escaped by the following sequences: #0A, #0D, and ##. No other substitutions occur. Example:

```
keywords "Key ##1#0D#0AKey ##2"
```

sets the keyword metadata field to two lines Key #1 and Key #2 using a carriage return / line feed for the line separator.

A job list entry within a job list file doesn't have to contain all the defined <property_name> keywords. When loading an entry, the value of a property is calculated as follows:

1. Use the default value. See Section "Setting up the Default Properties" (p. 73), on how to setup the general default values. For a single job list file with several jobs/entries the general default can be overwritten by a file specific default. Every job list entry containing the property *listdefault* with a value of 1 overwrites the default with its own property values.
2. If a corresponding <property_name> is found, use the specified value instead of the default.

The following list contains the most important keywords used to create custom job list files:

| <property_name> | Type | <value> |
|-----------------|--------|-------------------|
| name | string | Name of the entry |



| <property_name> | Type | <value> |
|-----------------|---------|--|
| pathIn | string | Path of the input directory or file (see also isDir) |
| isDir | integer | 0 = pathIn is a file 1 = pathIn is a directory |
| recurse | integer | 0 = do not scan input directories recursively 1 = scan input directories recursively |
| pathOut | string | Path of the output directory (see also outNext) |
| outNext | integer | 0 = place output files in the directory specified by pathOut 1 = place output files next to input files |
| pathMove | string | Path where to move input files after successful processing (see also modelIn) |
| modelIn | integer | 0 = do not move input files 1 = delete input files 2 = move input files to the directory specified in pathMove |
| pathErr | string | Path where to move input files after an error occurred during processing (see also moveOnErr) |
| moveOnErr | integer | 0 = do not move input files on error 1 = move input files on error to the directory specified in pathErr |
| title | string | PDF metadata "title" |
| author | string | PDF metadata "author" |
| subject | string | PDF metadata "subject" |
| keywords | string | PDF metadata "keywords" |

New property names and/or more possible values of existing names will be added in future releases of PDF Compressor. However, Foxit Europe will keep existing names, values and semantics compatible as far as possible.



Chapter 8. Submitting Jobs via the PDF Compressor API

The PDF Compressor provides a way to be programmatically driven. The 'Application Programming Interface', called API, was introduced in version 7.2. It enables 3rd party Windows applications and services to programmatically submit jobs to a PDF Compressor instance running on the same computer. For now, the API does not provide for remote access.

To make use of the API an external program must link against a DLL distributed with the regular PDF Compressor installation. No restrictions on usage or extra licenses apply. The API DLL comes in three different flavors:

- a native 32 bit DLL with a C programming interface,
- a native 64 bit DLL with a C programming interface,
- a .NET 2.0 Assembly for 64 bit environments, which will internally load the appropriate native version.

As all two interfaces offer equivalent functionality they are collectively referred to as *API DLL*. An application that has incorporated the PDF Compressor API DLL becomes an *API client*, whereas the PDF Compressor service is considered as *Server*. The client can create job configurations and submit them to the PDF Compressor via synchronous or asynchronous function calls. The progress of these jobs can then be monitored via *callback* functions and the caller can directly receive success or error messages.

In order to accept jobs submitted via the API the PDF Compressor Service must be put in *API Processing Mode*. Similar to the *Job List Processing Mode* covered in the previous chapter Chapter 7, *Automatic Job List Processing* (p. 83), it is a separate mode in which the PDF Compressor accepts jobs issued through the API. The *API Processing Mode* is mutually exclusive with Job List Processing Mode.

| Name | Status | Priority | Timer | Input | Output | Format | On Succ... | On Failure | Recogn... | Quality | Min. dpi |
|-----------|---------|----------|-------|---------------------------|---------------------|----------|------------|------------|-----------|---------|----------|
| Invoices | stopped | 5 | | C:\Data\input\ | [next to input] | PDF/A-2u | keep | keep | OCR | 6 | 200 |
| Plans | stopped | 5 | | C:\Data\input\ | [next to input] | PDF/A-2u | keep | keep | OCR | 6 | 200 |
| ZUGFeRD | stopped | 5 | | C:\Data\input\ | C:\Data\output\ | PDF/A-2u | keep | keep | OCR | 6 | 200 |
| API Job 1 | working | 5 | | C:\Folder1\input\Scan.tif | C:\Folder1\Output1\ | PDF/A-2u | keep | keep | | 6 | 200 |
| API Job 2 | working | 5 | | C:\Folder1\input\Scan.tif | C:\Folder1\Output2\ | PDF/A-2u | keep | keep | | 6 | 200 |
| API Job 4 | running | 5 | | C:\Folder1\input\Scan.tif | C:\Folder1\Output4\ | PDF/A-2u | keep | keep | | 6 | 200 |
| API Job 5 | running | 5 | | C:\Folder1\input\Scan.tif | C:\Folder1\Output5\ | PDF/A-2u | keep | keep | | 6 | 200 |
| API Job 3 | running | 5 | | C:\Folder1\input\Scan.tif | C:\Folder1\Output3\ | PDF/A-2u | keep | keep | | 6 | 200 |

| Core | Progress | Job | Status | Current Document |
|------|-------------|-----------|--------------|------------------|
| # 1 | <div></div> | API Job 1 | Page 6 of 10 | Scan.tif |
| # 2 | <div></div> | API Job 2 | Page 6 of 10 | Scan.tif |
| # 3 | <div></div> | API Job 4 | Page 4 of 10 | Scan.tif |
| # 4 | <div></div> | API Job 5 | Page 4 of 10 | Scan.tif |
| # 5 | <div></div> | API Job 3 | Page 6 of 10 | Scan.tif |
| # 6 | <div></div> | | | |
| # 7 | <div></div> | | | |
| # 8 | <div></div> | | | |

API Mode Processing Ready.



API mode processing is indicated in the status bar and the job list is being split into two parts. Automatically created jobs will show up in the lower part of the list. They can only be monitored. It is not possible to control (start/stop) or modify jobs created by API clients from within the GUI.

API Job Lifecycle

In contrast to the rather persistent job entries in *Standard Mode* of the PDF Compressor, in *API Mode* jobs have a more limited lifespan.

An API client creates an instance of a job configuration, either from default settings or by loading a configuration from a *.dat file (compare “*Job List File and Control File Syntax*” (p. 84)). Then he can access and modify all individual settings in this job configuration via *set* and *get* functions. The entire scope of options of standard and Job List Processing mode is available in API Mode, too.

Note, that input and output files are referenced through file system paths. For an API job to succeed, these paths must be accessible by the PDF Compressor service. The API does neither provide for passing the entire content of an input file programmatically to the service, nor for receiving the result in a similar fashion. File input and output are performed via the file system.

Once properly configured, the job can be submitted to the PDF Compressor for processing. Submission can be done synchronously or asynchronously. In the first case the call will return after the job has completed, in the latter case it returns immediately.

The API client can provide a callback function to receive progress and status information regarding the submitted job. This information can also generally be retrieved using the reference handle of a client to a submitted job.

Jobs can be submitted in a multithreaded and asynchronous fashion. If all processor cores on the server side are busy the PDF Compressor service will queue the submitted jobs and process them according to priority (“*Priority Processing Order*” (p. 70)) and submission time.

A client can prompt the server to abort a running job, if necessary.

The C/C++ API

The two native API DLLs reside in the *api* subfolder of the PDF Compressor installation folder as *PDF_CompApi_32.dll* and *PDF_CompApi_64.dll*. The C header and library files required to incorporate them in an API client can be found in the *api\include* folder. The *api* folder also contains a PDF manual for the API and sample source code in the *api\samples\ApiDemo* sub folder.

The .NET API

The *api* subfolder of the PDF Compressor installation also contains the .NET version of the API DLL as *PDF_CompApi.NET.dll*. It is also accompanied by documentation and sample source code in the *api\samples\ApiDemo.NET*. The API Assembly itself requires .NET 2.0, but can be used with higher versions. The assembly is strongly named and can therefore be installed in the Global Assembly Cache (GAC).

Note: The .NET API Assembly comprises a wrapper around the native API DLLs. Therefore, all three DLLs are required by a .NET API client at runtime.

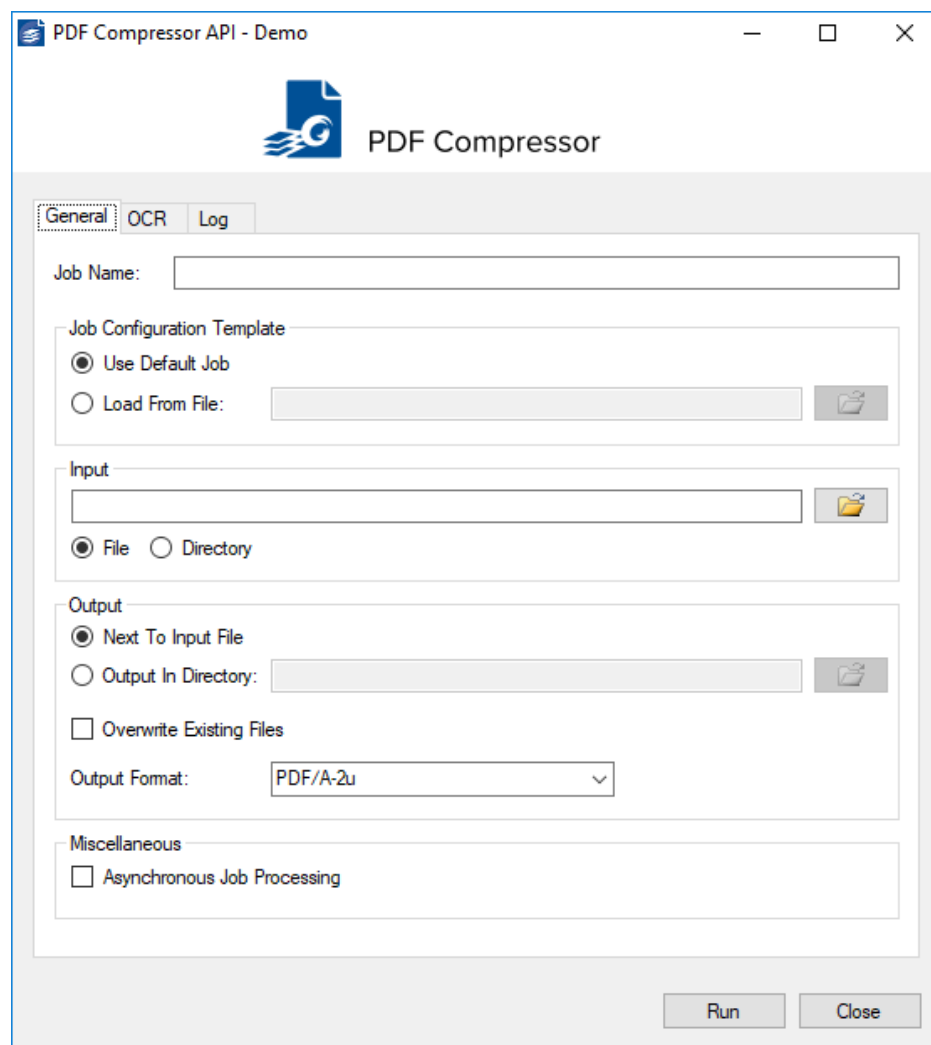


API Demo Applications

Two sample demo applications are deployed with PDF Compressor:

- **PDF_CompApi_Demo.exe** is a simple command-line client using the native C DLL that matches the version of your operating system.
- **PDF_CompApi_DemoGUI.exe** comprises a simple WinForms user interface, permitting a user to interactively configure a handful of common job settings, launch the corresponding job and receive its progress messages.

These two applications correspond to the compiled versions of the sample source code projects found under `api\samples` in the PDF Compressor installation folder.

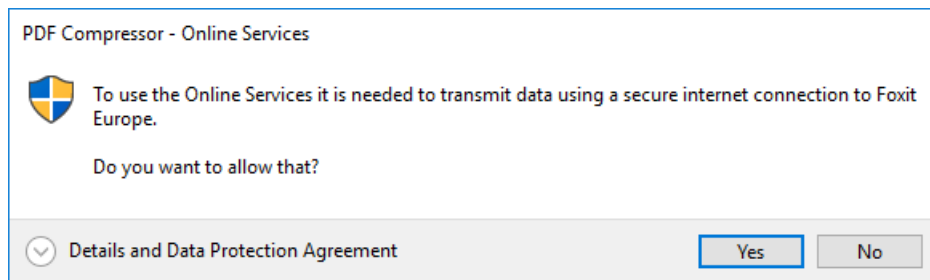




Chapter 9. Administration and Licensing

Administration and license management can be performed via the *File* → *Options* menu. The Options dialog has several tabs which are described in the following sub sections.

Online Services and Data Protection



PDF Compressor offers multiple service functions that need a online connection and transfer data to Foxit Europe.

These functions are:

- “*Online Check for Updates*” (p. 93)
- License Deactivation
- Chapter 13, *Support and Contact* (p. 119)

Typical these functions will transfer the license status in form of a license request to Foxit Europe. This request contains for example information about available pages, the used serial numbers, data to identify the machine and the user account name that used the function.

Some features like the Chapter 13, *Support and Contact* (p. 119) and *License Deactivation* can used in a limited form without a online connection and direct data transfer.

Other features like the “*Online Check for Updates*” (p. 93) can’t be used without.

We will use your data only to provide these functions and will not disclose it to third parties.

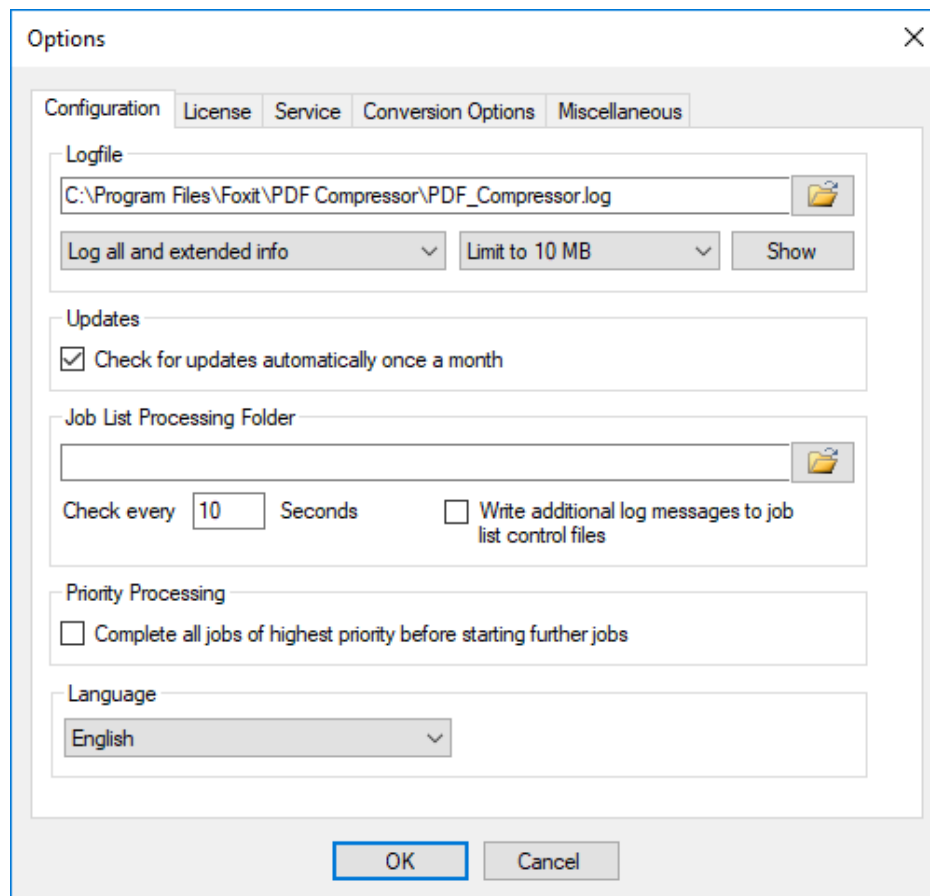
The processing of the data transmitted takes place exclusively based on your consent (Art. 6 (1) lit. GDPR).

The GUI will ask the user to confirm the described data transfer for the installation at first use of such a function and allow them only if this is confirmed. You can revoke these confirmation for future requests at any time in the GUI under *Options* → *Miscellaneous* → *Allow Online Services*. See “*Miscellaneous*” (p. 106).

A revocation of already given consent is possible at any time via an informal e-mail message.




General Configuration Settings



Log Files

The section *Logfile* of the Options dialog lets you configure PDF Compressor logging system.

1. You can specify path and file name of the log file. Type the path, use the browse button  or drag & drop the desired file from the Windows Explorer.
2. The type of logged information can be selected by the corresponding drop-down box:
 - a. *Log all and extended info* includes extended information about the input and output (resolution, size etc.). Apart from that it works like *Log errors, warnings and info*.
 - b. *Log errors, warnings and info* is the recommended setting. It logs errors, warnings and other basic information. With this setting the log file will give you detailed information on what files have been converted and how many pages from the license have been used.
 - c. *Log errors and warnings* will not log the additional conversion information.
 - d. *Log error messages only* will log only errors.
3. The size of the log file can be limited or set to unlimited by the corresponding drop-down box:
 - a. *Unlimited logfile size* will enable the log file to grow until all your disk space is used. This gives you the full history of all your conversions. When your server is performing



high volume conversions, and when it is not frequently monitored, it is recommended to limit the log file size by any of the other settings.

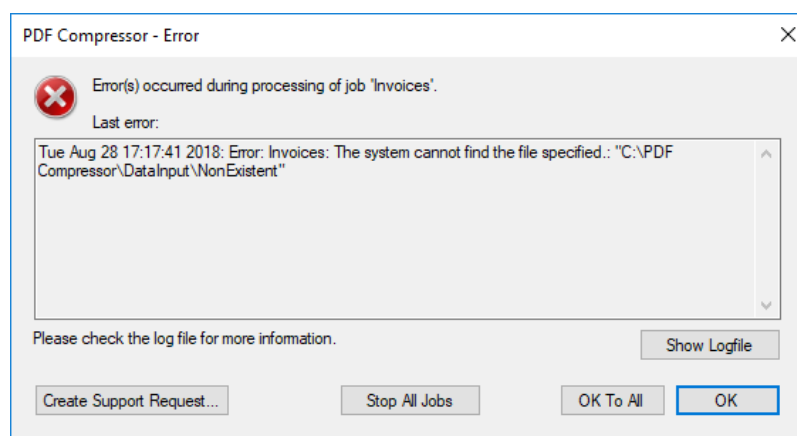
- b. *Limit to <num> MB* limits the log file size to the given amount. Once a log message is written that lets the log file size reach its limit, the file <log_file>.log is renamed to <log_file>.log.bak within the same directory, overwriting any present file of that name. A new, empty log file <log_file>.log is created. This ensures that at least the given number of bytes is always present as the logging history.
4. The *Show* button starts your default text editor application to show the content of the current log file.

Log File Analysis

Each line within the log file represents one message, consisting of the following fields:

1. Day, date and time of the message followed by a colon;
2. An optional CPU core number followed by a colon¹;
3. Type of message (Error, Warning, Info, InfoEx [extended information]);
4. If the message is related to a job entry: Name of the entry as shown in the job list, followed by a colon;
5. Clear text message like *error description*, *warning reason*, *information* - often related files are given with their full path in quotation marks.

Some errors are directly shown by a popup box when the GUI is opened:



Statistics Data in Log Files

PDF Compressor stores statistical data regarding processing times and OCR in the logfile, if the log level *info* is enabled under *File* → *Options* → *Configuration* → *Logfile*. These entries have the following form.

```
Thu Jan 03 20:30:36 2019: 1: Info: Entry 3: Output file written
"C:\Temp\Test.pdf"
```

¹If such a number is present, the message was issued by the PDF Compressor responsible to handle processing assigned to the given core number. Otherwise the message was issued by the PDF Compressor Windows service.



```
Thu Jan 03 20:30:36 2019: 1: Info: Entry 3: stats: 0:00:22
p=5 c=2955 w=397 C=2.39 O=2.04 F=0.01 E=0.01 T=4.44
```

The numbers in the second line denote the following values.

| Identifier | Description |
|------------|--|
| p | Number of pages in the result file |
| c | Average number of characters per page recognized by the OCR in this file |
| w | Average number of words per page recognized by the OCR in this file |
| C | Average time for the compression of one page |
| O | Average time for the OCR of one page |
| F | Average time for formatting of one page |
| E | Average time for exporting one page |
| T | Average time for the overall processing of one page |

In this context average time denotes the average over all pages of the corresponding output file. The values are given in seconds and their formatting is independent of the language settings of the PDF Compressor.

Online Check for Updates

If your computer is connected to the Internet, you can check whether updates for the PDF Compressor are available. Use *Help → Check For Updates...* to display the installed and the most recent version of PDF Compressor.

If the license check was successful, this dialog will also show the download link. You can open the link just by click on the *Installation Package* link or copy the link into the clipboard by right click and *Copy Link to Clipboard*. If the link is unavailable, please contact pdf_compressor@foxitsoftware.com .

Click *Browse Release History* to view a list of release notes.

In the *File → Options...* dialog you can enable *Check for updates automatically once a month* for a monthly update check. If a newer version is available, an *Update Information* dialog box shows up.

Please also check “*Online Services and Data Protection*” (p. 90) for information about the online services and General Data Protection Regulation.

Job List Processing

In the *Options* dialog you can configure the *Job List Processing Folder*, where PDF Compressor picks up the job lists to process. *Check every <number> seconds* defines how often the PDF Compressor checks the input folder for new job lists. Zero seconds is valid value and tells PDF Compressor to use the minimal possible waiting time, which is approx. 500 milliseconds.

PDF Compressor can collect all log messages that apply to jobs of the same job list and write these messages to the associated job list control file (*Write additional log messages in job list control files*).



See Chapter 7, *Automatic Job List Processing* (p. 83), for further details regarding job list processing.

Priority Processing

The option *Complete all jobs of highest priority before starting further jobs* only takes effect if *Priority Processing Order* is chosen as the general processing order. For more information on priority processing please compare “*Job List Processing Order*” (p. 70).

By default, the option *Complete all jobs of highest priority before starting further jobs* is not selected. This means that PDF Compressor will still start jobs in order of their priorities, but will allow for jobs of different (lower) priorities to run at the same time, thereby using the full capacity of all available cores.

If this option is selected, this may be desirable in some situations, but may lead to unnecessary idle times with more than one licensed processor cores.

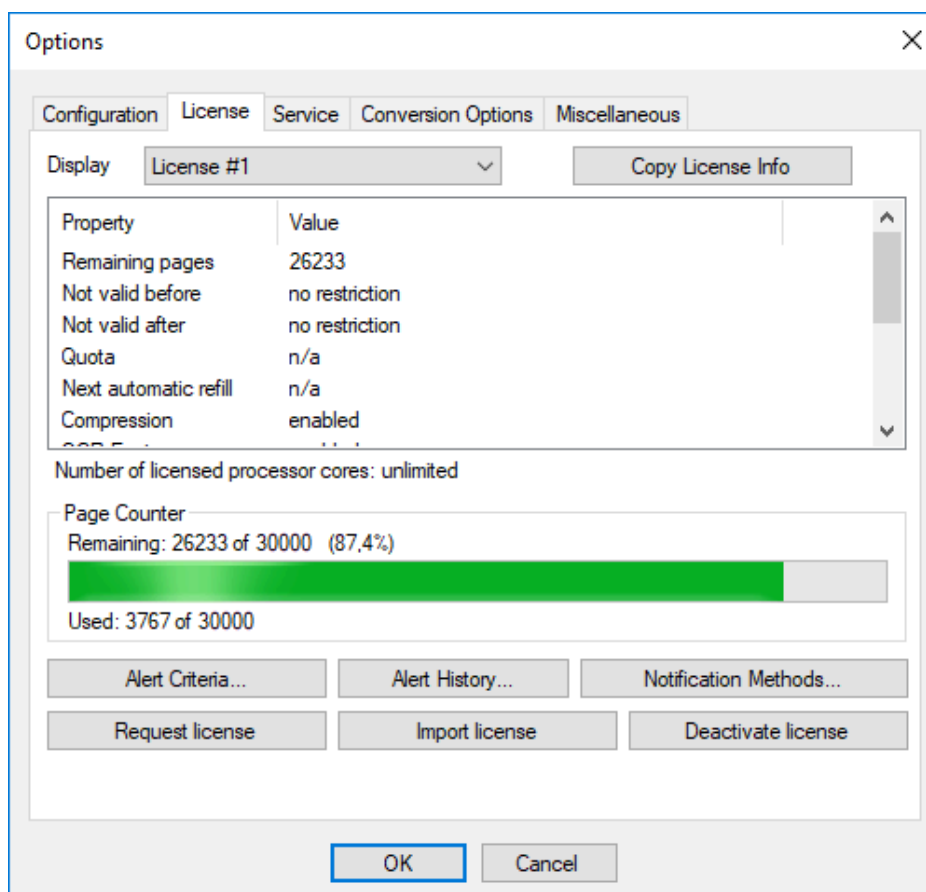
In either case, newly started jobs of higher priority will take precedence over the currently running ones (as soon as any active processor core has finished creating its current output file). If the *Complete all jobs (...)* setting is changed, it will immediately take effect, but will not cause cancellation of any current processing.

Language

The user interface language default is *Automatic*. PDF Compressor uses the system language if files with translations are available or English in case suitable language files are missing. The language option allows changing this and setting a language explicitly. A change in language will only be effective after restarting the user interface. To do so, the PDF Compressor window must be closed and the tray icon must be removed from task bar. This setting also determines the translation of email headers when emails are converted (see “*Configuring E-Mail Header Language*” (p. 31)).

Managing Licenses

PDF Compressor uses page-based licenses which may restrict the compression, or other features to a certain number of pages. It can maintain multiple licenses, each providing a contingent of remaining pages for certain set of features. A license can also be limited in time, thus be valid only during a certain period (*Not valid before, Not valid after*). The page contingent can be limited or unlimited. And there is also a license type that has a refilling page quota, e.g. the remaining pages count fills up to 5000 pages each month.



Trial Mode

When the PDF Compressor is installed for the first time, it is running in trial mode, which causes all output document pages to contain a visible watermark. OCR is possible in trial mode, but you are not allowed to output the OCR results in additional formats such as XML, HTML, etc. (see also “*Configuring Post-Processing*” (p. 63)). Since this mode is meant for evaluation purposes, it is sometimes also called evaluation mode.

This mode automatically ends after 30 days.

If you need a longer evaluation period, please contact our sales team (sales-eu@foxitsoftware.com).

License

For PDF Compressor different license models are available:

1. Volume Pack

One time contingent of pages, starting at 15.000 pages upto several millions of pages. Uses all available CPU cores of the system.

2. Annual Volume

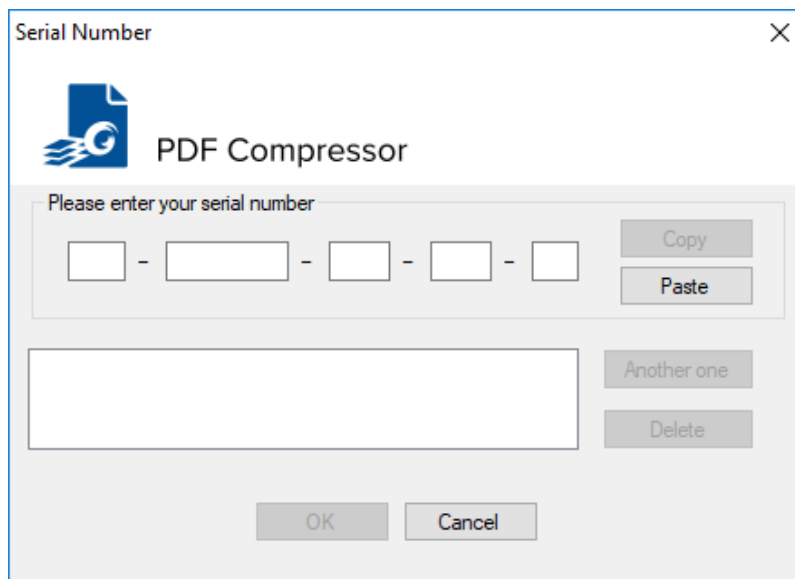
Recurring contingent of pages per Year. Uses all available CPU cores of the system.

3. Core



Unlimited time and volume-based license. Use the total number of licensed CPU cores.

You have to order the license by e-mail. This is done as follows:



1. Use the *Request license* button in *Options*→*License* tab to open the dialog shown above. To generate the request, you must enter the serial number as received with your invoice or along with the download link for PDF Compressor. If you obtained two or more serial numbers for the computer in question, please use the *Another one* button to enter all of them. This may apply when purchasing a regular license and an additional page contingent license.
If you cannot determine your serial number, please contact Foxit Europe support.
When you click *Ok* you are prompted to save the save a license request file. You can use an arbitrary file name such as `request.txt`.
2. Send the request file (in the example `request.txt`) via e-mail to `pdf_compressor@foxitsoftware.com` . Please also state the invoice number or the company which purchased the license.
3. We will reply by sending a license file `license.txt`.
4. Use the *Import license* button to open the file `license.txt`.

Afterwards you will see the number of remaining pages (which can be unlimited), and the license capacity within the *Options* dialog. For a limited license the number of remaining pages will decrease when you generate PDF output. When it reaches zero you have to order a new license by redoing steps 1 to 4 as shown above.

PDF Compressor will not fall back into trial mode once your license is empty. It will just stop creating further PDF output. A corresponding message will be displayed and written into the log file. This prevents you from generating PDF output with visible watermark accidentally.

The Number of licensed processor cores is shown in the *License* tab of the *Options* dialog. You can configure the actually used number of processor cores (up to the licensed maximum) in the *PDF Compressor Service* tab of the *Options* dialog.



Caveats

Your license is bound to a single installation.

To transfer the license to a different system or to restore it after re-installing the system, it is necessary to deactivate the existing license first. To accomplish that, use the wizard that can be started using the *Deactivate license* button within the Options Dialog. Do not uninstall the software before you do this successfully. You can find further information about this process at “*Moving Licenses*” (p. 97).

In case this is not possible, for instance due to hardware malfunction, consult the support department pdf_compressor@foxitsoftware.com ; we will inform you how to proceed. Under rare circumstances (e.g. after a system crash) it might be possible that your license becomes invalid. If this happens, please send a license request file to pdf_compressor@foxitsoftware.com . We will reply with a special repair license.

Updating Licenses

If you replace the PDF Compressor application by a newer version, an existing page-based license will remain valid. However, depending on the updated version of the PDF Compressor, you might need an update license for the new application. In such a case the Options Dialog will show an additional button *Set update license* which allows you to set the update license.

Moving Licenses

The existing license needs to be deactivated before it can be issued again for a different machine. To do this, use the wizard that can be started using the *Deactivate license* button within the Options Dialog. After that you can create a license request from the new installation. To receive a new license, send a license request from the new installation pdf_compressor@foxitsoftware.com . Please mention that you want to move the license.

Useful hints:

- Remember the account used to start the PDF Compressor service
- Export the settings using *File → Export Job List...* and import them into the new system

License Monitoring

To minimize the risk of unexpected license expiry you can set license alerts to notify you in time. This helps to ensure continuous PDF Compressor operation when dealing with limited runtime or page contingents. To activate one or more license alerts, select *Alert Criteria...* from the *License* tab of the *Options* dialog.



Alert Criteria [X]

License Alert Criteria

The program will display an alert when the license has expired. To minimize the risk of such interruptions, it is recommended that you also enable at least one of the following optional alerts:

Page counter-based warnings

☐ Warn when the number of pages falls below this number: 1000

☒ Display a warning when the % of remaining pages falls below each of the following values:

- ☒ 99% remaining
- ☐ 98% remaining
- ☐ 97% remaining
- ☐ 96% remaining
- ☐ 95% remaining
- ☐ 94% remaining

Selected: 99% remaining

Time period-based notifications

☐ Time since last activity: 1 Minutes

☒ Number of days remaining until license expiration falls below 7

Write license status to Windows Event Log

☐ Daily at 00 : 00 (24-hour format)

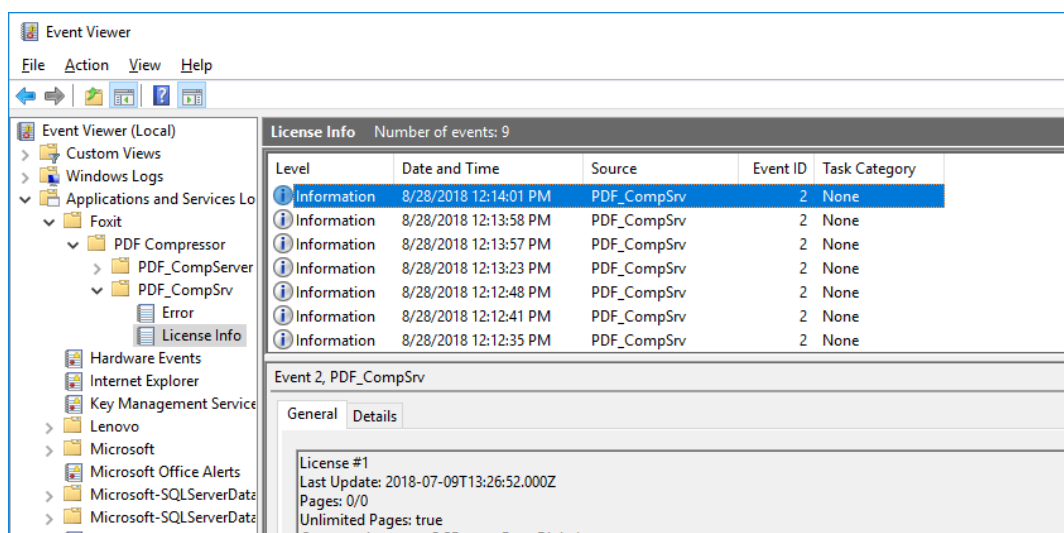
☒ After every 50 pages processed (one value for all counters)

Number of seconds between activity checks: 4

OK Cancel

1. For *Page counter-based warnings* you have the options
 - Warn when the number of pages falls below this number...
 - Display a warning when the % of remaining pages falls below each of the following values...
2. For *Time period-based warnings* you can specify
 - Time since last activity
 - Number of days remaining until license expiration falls below...
3. Under *Write license status to Windows Event Log* the PDF Compressor can be also configured to write license status information to the Windows Event Log on a regular basis.
 - You can configure the trigger when to write the updated license status to the event log - either each day at certain time or based on page consumption.

The license status events are written to the Microsoft Event Log: Applications and Services Logs → Foxit → PDF Compressor → Compressor-PDF_CompSrv → License Info. If your license contains sub-licenses you will see a single event for each sub-license.



You can use the Windows Event Viewer to check the events but more likely you will use system management or monitoring tools to evaluate the data.

Each license status event contains the following details.

| Name | Type | Description |
|-------------------|-------------|---|
| LicenseIndex | Int | The index of the license. This corresponds to the index shown in the Options dialog. |
| LastUpdateTime | DateTime | The time when the page contingent was last changed by importing a license, e.g. installing a new cartridge. |
| LastUpdatePages | UnsignedInt | The value of the page counter at the LastUpdate-Time. |
| CurrentPages | UnsignedInt | The current value of the page counter, i.e. the remaining pages of this license. |
| HasUnlimitedPages | Boolean | True if the license has unlimited pages, False otherwise. |
| HasCompression | Boolean | True if the license contains the compression feature. |
| HasOcr | Boolean | True if the license contains the OCR feature. |
| HasBornDigital | Boolean | True if the license contains the Born Digital feature. |
| Version | String | Product name and version number. |

Finally, you can choose the *Number of seconds between activity checks*

The button *Alert History* on the *License* tab of the *Options* dialog shows an overview of previous license alerts.



Clicking the *Notification Methods* button on the *License* tab of the *Options* dialog how you are notified when an alarm is triggered.

You can select any combination of the following three notification methods:

1. *Log alerts to Windows Event Log*
2. *Display alerts on screen for ... seconds*
3. *Enable e-mail alerts*

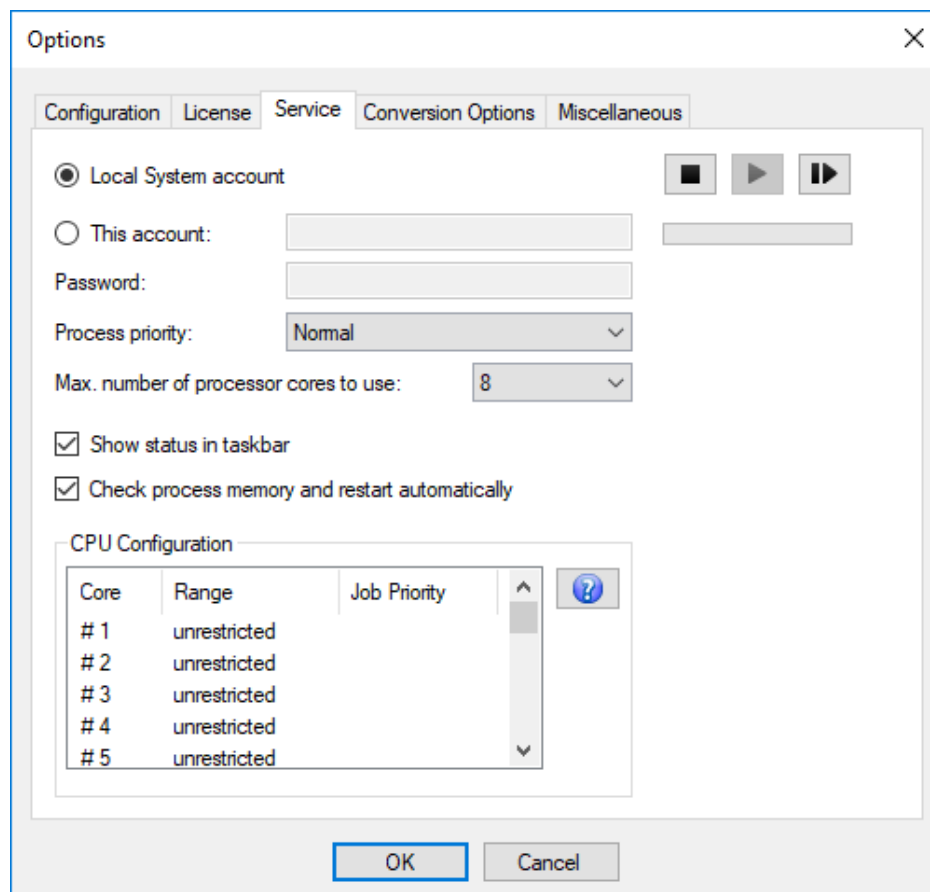
For the latter you must enter data concerning the mail server and user account under *E-Mail-Setup*. In addition you can specify *Sender name*, *Message subject* and a *Custom mes-*



sage for the alert mails to send under *Message Properties*. Information about the system (hostname), the detailed license warning and the license details will be added automatically to the e-mail message.

Managing the PDF Compressor Service

The PDF Compressor Service can be directly administrated from the Options dialog of the PDF Compressor GUI. Alternatively the Windows service administration tools can be used. Here only the administration using the PDF Compressor GUI is described.

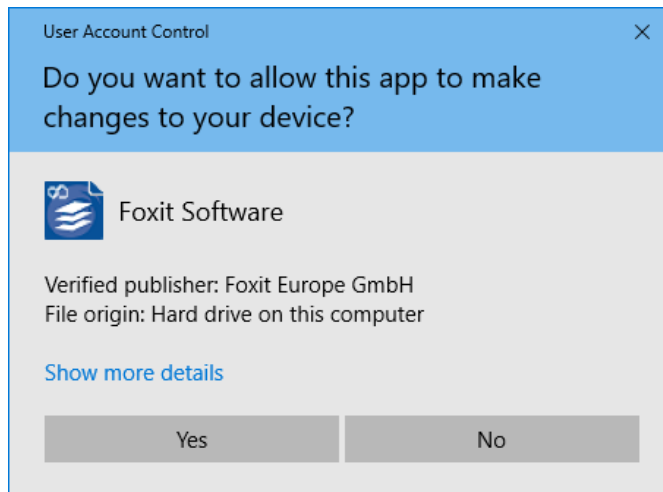




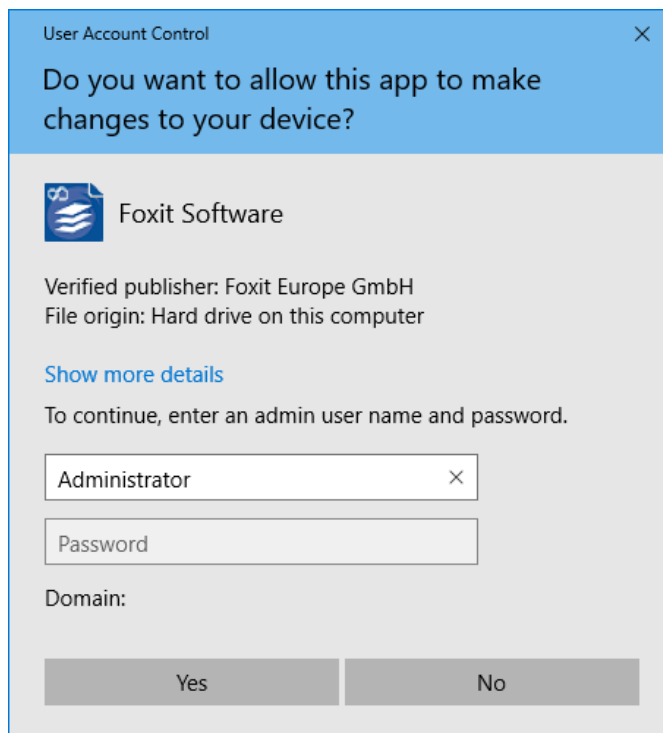
Required Privileges


The administration of a Windows service needs administrative privileges. If you haven't started the GUI with such privileges, they are requested if necessary.

You can grant them yourself:



Or, if you don't have the privileges yourself, by an administrator:



If the confirmation of the Options dialog will require such privileges, this will be indicated to you by the shield symbol: 

Choosing the Service Account

The PDF Compressor Service tab of the Options dialog shows the current status of the service (normally running) and the current service account (*Local System* after installation).



One of the strengths of the PDF Compressor lies in its ability to do automatic PDF conversions even if no user is logged in. This also enables PDF Compressor to do conversions on remote computers by use of network shares. The default service account *Local System* typically does not have the right to access files on remote machines of a Windows domain. You can change the service account by use of the Options dialog. When you select a new account, make sure it has access to:

- the PDF Compressor installation directory
- the input folders of the jobs you want to run
- the output folders of the jobs you want to run
- the log files


If you have changed the account, the OK button triggers a restart of the service. If the given password is not correct, you will get an error message. Starting, stopping and restarting the service can be also done by the corresponding buttons ■ ► ▢ within the Options dialog. Please mind that the PDF Compressor user interface must run with administrative privileges to change and restart the service. Alternatively you can use the Windows service manager instead.

Important Note: When configuring a network share for input or output files, make sure that you use the syntax `\\host\share\directory` to specify the location within the Input, Output, and Options tabs. Mapped network drives like `Y:\directory` cannot be used, since they are not available to the PDF Compressor Service even when you have changed the service account.

Setting up the Service Priority

Depending on your workflow you might want to change the process priority of the PDF Compressor service. In case your workflow is designed to run PDF Compressor on its own PC that is only used for document processing, it is best to set Process priority to *Normal*. When your PC has only a single core processor (CPU), and you are working with the PC while PDF Compressor processes its jobs in the background, it may help to set the Process priority to *Low*. That way PDF Compressor interferes less with your work. Keep in mind that the priority *Low* is typically used for screen savers. Therefore, screen savers are better turned off when PDF Compressor priority is set to *Low*.

Taskbar Status Icon

When you log in, a PDF Compressor icon  will be shown in the Windows taskbar. Its color reflects the status of the PDF Compressor, and it can be used to open the GUI (double-click or context menu). If you want the icon to disappear, please disable the check box *Show status in taskbar* of the *Options* dialog.

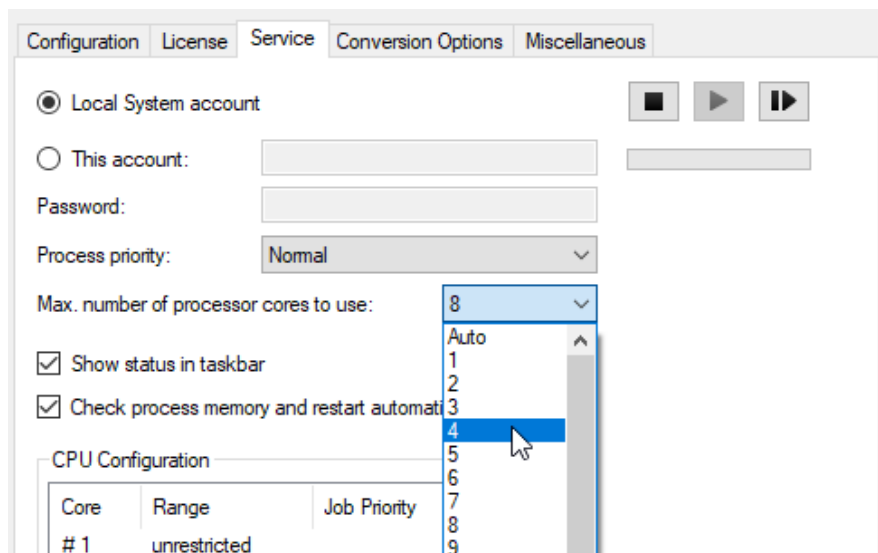
CPU Configuration

The PDF Compressor can process multiple jobs in parallel when *Priority Processing Order* is enabled (see “*Priority Processing Order*” (p. 70)). Jobs with the same priority will be scheduled in a circular order (round-robin scheduling). Jobs with a higher priority take precedence



over jobs with a lower priority. The job priority is a value in the range of 1 (highest) to 10 (lowest).

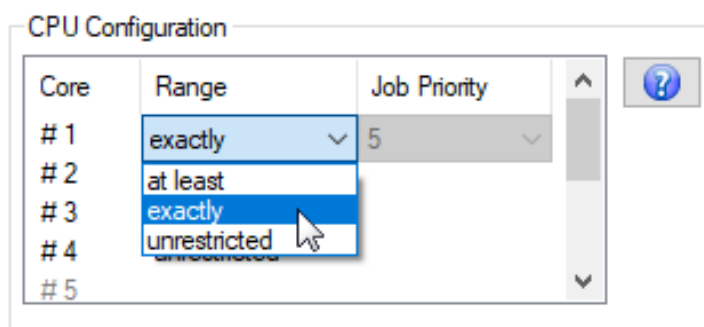
The amount of processor cores used for conversion can be adjusted manually through the dropdown menu. By choosing "Auto" mode PDF Compressor will either use the maximum licensed cores or the amount available in hardware, depending on the smaller number. Furthermore, this value is set as default after the first installation and each successful license import.



The PDF Compressor is built to operate at the full capacity of the available resources, namely the available and licensed processor cores, to reach the maximum throughput. However, there are scenarios where you want to spare some resources for high priority tasks. Or you might want different priority jobs to run on different cores. In either case you can adjust the *CPU Configuration* accordingly.

In the *CPU Configuration* table, you find one row per licensed processor core. By clicking a row, the entry switches to edit mode and lets you modify the settings for the selected core.

Note: The CPU Configuration is read-only as long as there are jobs running. You need to stop all running jobs in order to edit the CPU Configuration.



The *CPU Configuration* allows you to specify restrictions for single cores on the basis of the job priority. The column Range offers three different settings.



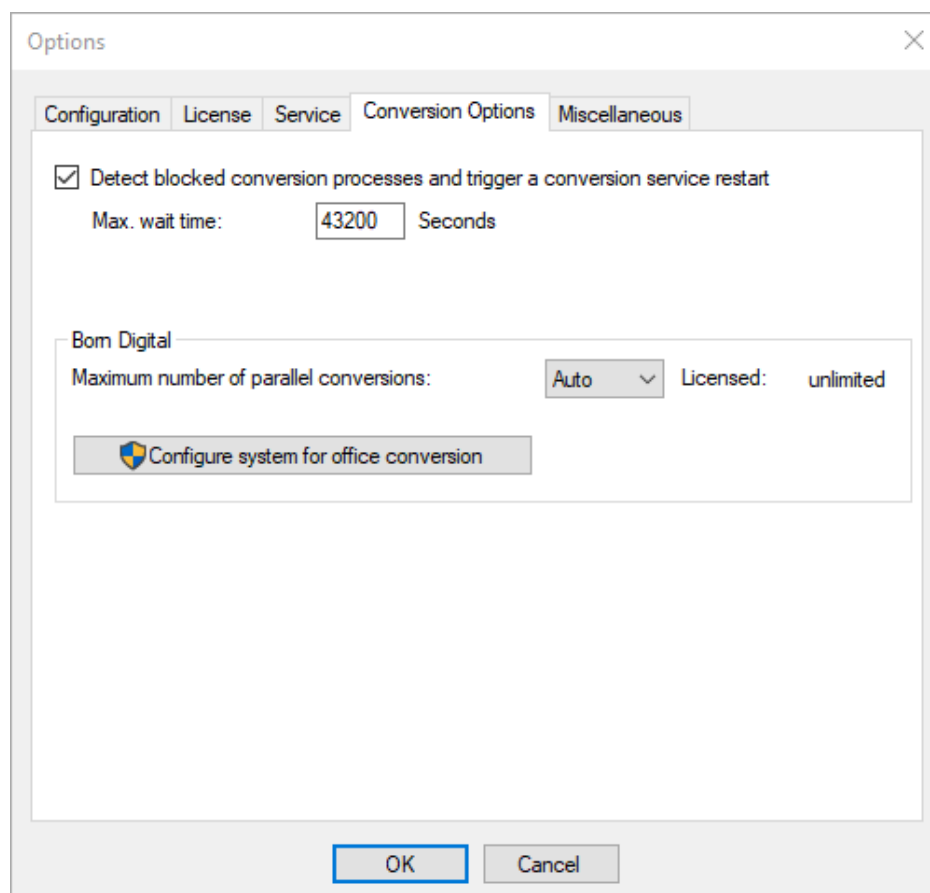
| Setting | Description |
|---------------------|--|
| unrestricted | This is the default setting. The core will be used to process jobs of any priority. |
| at least <priority> | The core will be used to process jobs with a priority of <priority> or higher. A setting like "at least 3" will restrict the core to only run jobs with priority 1, 2, or 3. |
| exactly <priority> | The core will be used to only process jobs with a priority of <priority>. |

Examples

1. To dedicate one core for important jobs: + Change the configuration for core #1 to "at least 4", to have PDF Compressor to run only jobs with a higher than standard priority (1 to 4) on that core. Leave the configuration of all the other cores "unrestricted".
2. To share four cores between two jobs:
Give the two jobs different priority values, e.g. 1 and 2. In the CPU Configuration setup cores #1 and #2 for "exactly 1" and set up cores #3 and #4 for "exactly 2". This way both jobs will run on two cores each exclusively.

Born Digital Options

The *Options* dialog contains a tab, called *Conversion Options*, for administering options specific to the PDF(/A) conversion.





The Born Digital component utilizes third-party applications (e.g. Microsoft Office) during the conversion process. It may happen that these applications become unavailable or get stuck so that the conversion process cannot finish. To avoid blocking whole conversion in that case, there are some options for automatic clean restarting.

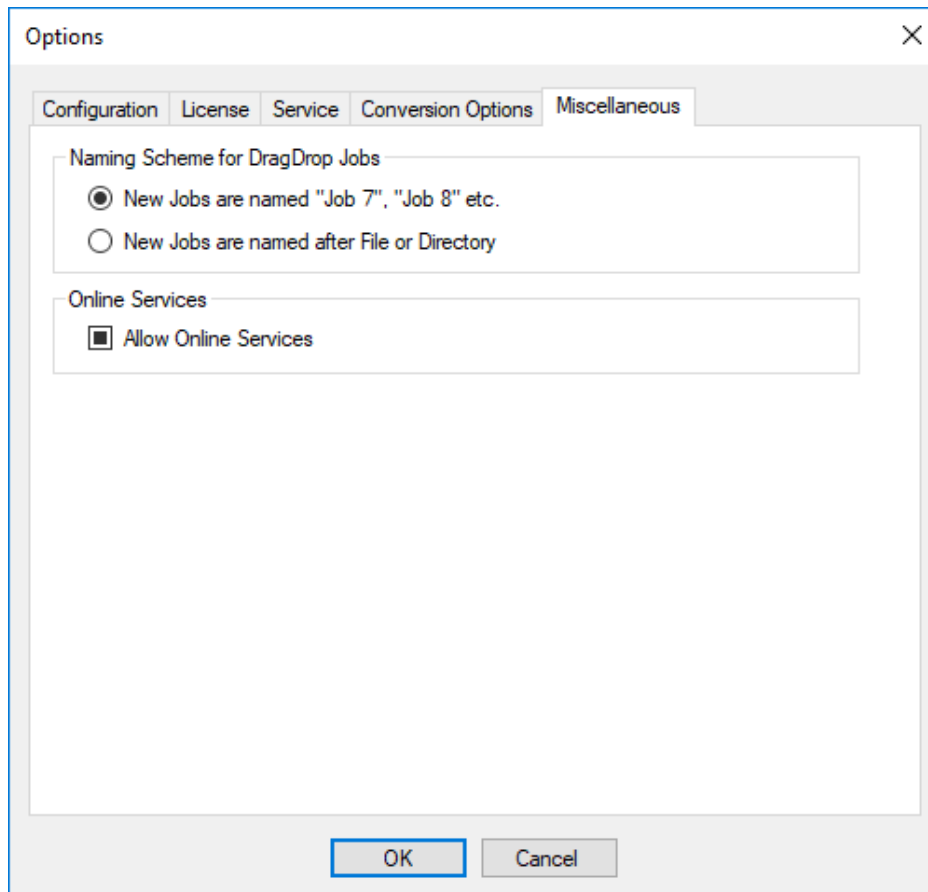
Note: In order to enable PDF Compressor to successfully restart the conversion service the PDF Compressor service user account must be set to LocalSystem or to an account which is member of the local Administrators group (see “*Choosing the Service Account*” (p. 102)).

1. *Detect blocked conversion processes:* Define how long the service waits for an ongoing conversion process to send any feedback at maximum. The default value of 43200 seconds means 12 hours. If the system gets stuck and the situation is not noticed at least after half a day the conversion service will be recovered. Usually a single conversion takes much less time, so that a value of 600 (10 minutes) might be more convenient.
The *Max. wait time* does not define the maximum overall processing time. Instead it defines how long the service waits for an external task to get finished, e.g. Microsoft Office, that does not feedback any progress. This means that the overall processing time can be much longer than the *Max. wait time* e.g. when processing an e-mail with multiple attachments.
2. *Maximum number of parallel conversions* specifies how many Born Digital conversion processes PDF Compressor should attempt to run in parallel. The maximum available number is the number of processor cores licensed for Born Digital conversion. The additional "Auto" mode allows an automatic selection of maximum available number of processor cores. It is enabled by default after the first installation and each successful license import.
3. Use the button *Configure system for office conversion* to run a special configuration tool. The tool can be used to change the PDF Compressor service account. Additionally it sets a number of permissions and properties (DCOM) to enable the conversion of Microsoft Office™ documents in the background.

Note: The tool requires administrative privileges (see “*Required Privileges*” (p. 102)).

Miscellaneous

On the Tab *Miscellaneous* you find further options.



As stated earlier you can create new jobs via dragging and dropping files or folders into the main window of the GUI. Here you may set, whether new jobs are continuously numbered (*New Jobs are named "Job 7", "Job 8" etc*) or whether the job will be derived from the dropped file or directly (*New jobs are named after file or directory*).

Additionally you can enable or disable the online service here. For more information please check *"Online Services and Data Protection"* (p. 90).



Chapter 10. Tips and Tricks

This section contains some hints you may find useful.

- The easiest way to create a new entry in the job list is to drag & drop files or folders from the Windows Explorer to the list window.
- **All** text fields that represent file or folder names accept the dragging & dropping of files or folders.
- There are more valid shortcuts for header and footer text than shown in the GUI. Please refer to *“Template String Syntax Description”* (p. 73), for more details.
- *Header and Footer Options* are merely the captions of the corresponding buttons on the Properties → Embedding tab. You can configure these imprints e.g. to set left and right bounded text in the header, or to a large outlined text at the center of a page.
- To create different variants of output (e.g. high resolution, low resolution and bitonal versions) from the same input file, you can configure several job list entries in a cascade:
 1. Use Entry A to make Version A at Destination A, after processing move the input to the input folder of Entry B.
 2. Use Entry B to make Version B at Destination B, after processing move the input to the input folder of Entry C.
 3. Use Entry C to make Version C at Destination C, after processing move the input to the final destination folder (which could be the input of Entry A).
- You can duplicate job list entries (see *“Adding, Deleting and Copying Entries”* (p. 80)). This lets you quickly set up two entries that only differ in a few parameters (e.g. encoder quality), and can be useful for comparison purposes.



Chapter 11. Troubleshooting

| Problem | Possible Reason / Solution |
|--|---|
| The output PDF has a visible watermark. | PDF Compressor is running in trial (evaluation) mode. Please order and install a license or an update license (see “ <i>Managing Licenses</i> ” (p. 94)). |
| The output PDF looks strange within Mozilla Firefox or other applications with PDF rendering capabilities. | Unfortunately, some viewers do not properly render and display PDF layers and transparencies as generated by PDF Compressor. In order to verify the integrity of your documents try viewing them with an Adobe Reader 7.0 or higher or the Foxit Reader. |
| The input document is not converted to PDF. | Check if your license has expired, i.e. the number of remaining pages is zero and you are not running in trial (evaluation) mode. If necessary, order and install a license (see “ <i>Managing Licenses</i> ” (p. 94)). |
| | Verify you selected the corresponding input file format. The Input tab (see “ <i>Configuring Input Data</i> ” (p. 11)) lets you choose among several input file formats. |
| | Check your input file is of a supported file format. PDF Compressor only accepts TIFF, JPEG, JPEG 2000, PDF, BMP or PNM as input format. See “ <i>Configuring Input Data</i> ” (p. 11), for further formats and options. Moreover, TIFF has a large variety of sub formats. Currently bilevel, grayscale or RGB color TIFF images are supported. CMYK or CIELab color, as well as palletized images are not supported. |
| | Is the underlying PDF Compressor service running? Check the status text within the PDF Compressor job list and/or the Options dialog. See Managing the PDF “ <i>Managing the PDF Compressor Service</i> ” (p. 101). |
| | Is your input folder a network share where the underlying PDF Compressor service has no access rights? See “ <i>Choosing the Service Account</i> ” (p. 102) for details. |
| | Did you use a mapped network drive letter to specify the input location? You cannot use mapped network drives. Use the syntax \host\share\ instead. See “ <i>Configuring Input Data</i> ” (p. 11) for details. |
| Processed input files are not moved to the given destination folder. | Check whether a file of the same name already exists at the destination and whether the overwrite option is set? Check the log file (see “ <i>Log File Analysis</i> ” (p. 92)). |
| | The overwrite option is set, but a file with the same file name at the destination cannot be deleted, because it is write protected or opened and locked by another application. Try to delete the file in the destination folder using the Windows Explorer. Please check the log file (see “ <i>Log File Analysis</i> ” (p. 92)). |
| | Is your destination directory a network share where the underlying PDF Compressor service has no access rights? See “ <i>Choosing the Service Account</i> ” (p. 102) for details. |



| Problem | Possible Reason / Solution |
|--|---|
| Processed input files are not deleted, even if you use the delete option. | Maybe the input file is write protected. Try to delete the file by use of the Windows Explorer. Please also check the log file (see <i>"Log File Analysis"</i> (p. 92)). |
| The output PDF file seems to be an old version from a former PDF conversion, but you are using the overwrite option. | Is the output file write protected or opened by another application? Try to delete the file by use of the Windows Explorer. Please check the log file (see <i>"Log File Analysis"</i> (p. 92)). |
| No log file is written, or the existing log file is not appended. | Check if the configured log file destination really exists. If it exists, check write rights. Maybe it is opened and locked by another application. Try to move the log file out of the way by use of the Windows Explorer. |
| | Check if the log file is located on a network share where the underlying PDF Compressor service has access rights. See <i>Choosing the "Choosing the Service Account"</i> (p. 102), for details. |
| PDF Compressor displays a message 'Invalid license'. | Does the underlying PDF Compressor service have full access rights to its installation folder? Within the Options dialog, try to stop and restart the service (compare <i>"Managing the PDF Compressor Service"</i> (p. 101)). Does a message box tell you that the service has not the rights mentioned above? Change the service's account (as described under <i>Choosing the Service Account</i> , p. 85), or the access rights of the installation directory to meet the above requirements. |
| | Please send a license request file to pdf_compressor@foxitsoftware.com (see <i>"Managing Licenses"</i> (p. 94)). |

Born Digital Conversion

Please check if you met all *"Requirements of the Born Digital component"* (p. 2) in the first place.

| Problem | Possible Reason / Solution |
|--|--|
| Required macros within my documents are not executed. | Make sure that the processing of Office macros is not disabled in the security center of Office. Keep in mind that this option has to be set correctly for the user who is selected by <i>"Born Digital Options"</i> (p. 105). |
| | Make sure no policy, for example for your Windows domain, forbids the execution of Office macros. Such a policy would overwrite the settings you made in Office. |
| Office macros are not processed correctly or the processing hangs. | Make sure that the macros do not require any user interaction (input, buttons, ...). This includes also message boxes. |
| | Not all constructs are available for macros processed by PDF Compressor: The properties <i>ActiveDocument</i> and <i>Selection</i> for example are related to the graphical user interface of Office, which is not available during the background processing. Due to that, also the |



| Problem | Possible Reason / Solution |
|---|--|
| | usage of these properties is not possible. Use for example <i>This-Document</i> to access the current document from within a macro. |
| A job configured for converting Born Digital input (e.g. Microsoft Office documents) does not finish and there is no processing noticeable for a long time. | It can happen that a Born Digital conversion process gets stuck (see " <i>Born Digital Options</i> " (p. 105)). If that is the case, you can change the appropriate setting right away. Within the <i>Options</i> dialog, change the <i>Max. wait time</i> to e.g. 600 seconds (10 minutes). The new value will be evaluated instantly and cleaning up and restarting process is initiated. Note that this process may take a considerable amount of time. |



Chapter 12. License Agreement

Foxit Europe Software License

PLEASE READ THIS SOFTWARE LICENSE AGREEMENT ("LICENSE") CAREFULLY.

THE "FOXIT EUROPE SOFTWARE PRODUCT" INCLUDES COMPUTER SOFTWARE AS WELL AS ACCOMPANYING PRINTED MATERIAL AND "ONLINE" OR ELECTRONIC DOCUMENTATION. BY INSTALLING, COPYING OR USING THE FOXIT EUROPE SOFTWARE PRODUCT IN ANY WAY, YOU ARE AGREEING TO BE BOUND BY THE TERMS OF THE FOLLOWING LICENSE. YOU NEED NOT NOTIFY FOXIT EUROPE OR YOUR DEALER IF YOU AGREE TO THIS LICENSE. IF YOU DO NOT AGREE TO THE TERMS OF THIS LICENSE, DESTROY THE SOFTWARE OR RETURN THE SOFTWARE TO THE PLACE WHERE YOU OBTAINED IT.

1. Copyright

The license issuer is Foxit Europe GmbH, or its local subsidiary. The Foxit Europe Software Product is protected by national and international copyright law. You are obliged to treat the Foxit Europe Software Product in the same way as any other material protected by copyright law.

2. License

The Foxit Europe Software Products are licensed to users, not sold. The license issuer retains the right of use of the enclosed software and any accompanying fonts (collectively referred to as Foxit Europe Software Product) whether on disk, in read only memory, or on any other media. You own the media on which the Foxit Europe Software Product is recorded but Foxit Europe and/or LuraTech's licensor(s) retain title to the Foxit Europe Software Product. The Foxit Europe Software Product in this package and any copies which this license authorises you to make are subject to this license.

Using the Foxit Europe Software Product for creating a commercial or non commercial software solution that enables either the creation or display files in the LuraWave or LuraDocument File Formats requires that a Runtime-License-Contract has be concluded. This applies among other, but not limited to, the integration into other software application or to the distribution of images or documents on electronical data carrier devices (e.g. CD) or in the internet.

To receive an offer or to enter into a run-time license agreement please contact the head office: Neue Kantstr. 14
14057 Berlin
GERMANY

Tel.: +49 2191 / 58960 - 0
Email: info@foxitsoftware.com

or one of its local subsidiaries.

3. Usage

This license allows you to install and use the Foxit Europe Software Product on a single computer. This license does not allow the Foxit Europe software to exist on more than one computer at a time. You may make one copy of the Foxit Europe Software Product in machine-readable form for backup purposes only. The backup copy must include all copyright



information contained on the original. All other rights and title to the Foxit Europe Software Product are reserved.

4. Transfer of Rights

You may only transfer your rights under this license with written consent from Foxit Europe and only to a party who agrees to accept the terms of this license. After transferring your rights you must destroy any other copies of the Foxit-Europe-Software-Product you have in your possession. You may not rent, lease, loan or distribute the Foxit Europe Software Product.

5. Restrictions

You may not (i) decompile, reverse engineer or disassemble the Foxit Europe Software Product in any way, (ii) modify, adapt or create derivatives from the Foxit Europe Software, or (iii) sell, rent, lease, loan or distribute the Foxit Europe Software Product. Your rights under this license, and your rights to use the Foxit Europe Software Product, will terminate automatically, without notice from Foxit Europe, if you fail to comply with any of term(s) of this license.

6. Updates

If this Foxit Europe Software Product is marked as an update then you must have a licensed version of the appropriate full software product. The Foxit Europe Software Product supplements the licensed version of the full product. All updates are liable for cost.

7. Demo Version

Foxit Europe demo software, shareware and the associated documentation are copyright protected. Foxit Europe allows the license holder of the demo software/shareware version to make as many copies of the software as required. The license holder is forbidden from using the demo software/shareware for commercial purposes or to distribute for a fee.

8. Student Version

If this Foxit Europe Software Product is marked as a student version then the license holder must have the status of a registered student to have the right to use the product.

9. Special Conditions

The product "LuraWave OCX-SDK" contains a license file "lurawave.lic". This file must be kept secure and not be accessible or given out to third parties

The license holders access codes must be kept secure and not be accessible or given out to third parties.

Foxit Europe Software Products may include the TIFF-library. Copyright reference for the TIFF-library:

Copyright © 1988 - 1996 Sam Leffler

Copyright © 1991 - 1996 Silicon Graphics, Inc.

Foxit Europe Software Products may include the JPEG library of the Independent JPEG group. Copyright reference for the JPEG-library (libJPEG 6b, 27-Mar-1998):



Copyright © 1998 - Thomas G. Lane

Permission to use, copy, modify, distribute, and sell this software and its documentation for any purpose is hereby granted without fee, provided that (i) the above copyright notices and this permission notice appear in all copies of the software and related documentation, and (ii) the names of Sam Leffler and Silicon Graphics may not be used in any advertising or publicity relating to the software without the specific, prior written permission of Sam Leffler and Silicon Graphics.

Foxit Europe Software Products may include Nuance OmniPage. Copyright reference for Nuance:

The software and materials provided with this agreement are licensed, not sold, and are available for use only under the terms of this license agreement. Please read this agreement carefully. By downloading, installing, copying, or otherwise using the software, you agree to be bound by the terms and conditions of this agreement and become a party to this agreement. If you do not agree with all of the terms and conditions of this agreement, do not download, install, copy or otherwise use the software.

This Nuance Communications, Inc. ("Nuance") End User License Agreement ("Agreement") accompanies a Nuance software product in machine-executable binary code and related explanatory written materials ("Software"). The term "Software" shall also include any modified versions, updates or upgrades of the Software that may be licensed to you ("Licensee") by Nuance, but does not include source code for the Nuance software product. As used in this Agreement, "Software Package" means the Software and the accompanying documentation. You may install and use such a modified version, update, or upgrade of the Software only if you have a validly licensed full version of the Software being modified, updated, or upgraded. If you download, install, copy, or otherwise use such a modified version, update, or upgrade of the Software, then this Agreement terminates as to the previous version of the Software, and you have a license only to such modified version, update, or upgrade of the Software under the terms of this Agreement. The Software is licensed to you, Licensee, as the end user, subject to all of the terms and conditions of this Agreement.

LICENSE GRANT. Subject to the terms and conditions of this Agreement, Nuance grants Licensee a non-exclusive and non-transferable license only to: (a) install and use for personal or internal business purposes one copy of the Software on a single computer; (b) install and use for personal or internal business purposes one copy of the Software on either a single portable computer or a single home computer, provided that such copy is not used concurrently with the copy in section (a) above; (c) make a single copy of the Software solely for archival purposes; and (d) store or install a copy of the Software on a storage device such as a network server, used only to install or run the Software on your other computers over an internal network, provided that you acquire and dedicate a separate license for each separate computer on which the Software is installed, run or otherwise accessed from the storage device. A single license for the Software does not allow you to share the Software or use it concurrently on different computers or for others other than the Licensee to access, install, download, copy or otherwise use the functionality of the Software.

MULTIPLE LICENSE PACK. If Licensee has purchased a Multiple License Pack, Licensee may make additional copies of the Software not to exceed the number of licenses purchased. Licensee may use each copy solely in the manner specified in this Agreement.

ACTIVATION. According to its discretion, Nuance may include features in the Software to prevent unlicensed use of the Software. You agree that Nuance may do so. In particular, use of the Software may require that Licensee activate the Software through the Internet (this



process may be more fully described during the installation set up of the Software). During such activation, Nuance may collect certain non-personal technical information from your computer concerning your computer or network. You agree that Nuance may do so. You may be required to reactivate the Software if you modify your computer hardware or the Software.

NOTIFICATIONS. According to Nuance's discretion, the Software may contain a component that will automatically activate your Internet browser and attempt to initiate a connection through the Internet to a website maintained by Nuance that contains notification information related to the Software. This connection may be made using the Internet connections and telephone lines under your control. From time to time, you may receive notices about the Software or other information through this Internet connection. By installing the Software on your computer, you hereby consent to have the Software initiate a connection through the Internet to Nuance's website, to use your resources to connect to such website, and to receive notices about the Software and other information through this Internet connection.

RESTRICTIONS. Except as otherwise expressly permitted in this Agreement, Licensee may not: (a) reproduce or copy any of the Software; (b) modify or create any derivative works of the Software, including translation or localization; (c) decompile, disassemble, reverse engineer, or otherwise attempt to derive the source code for the Software; (d) redistribute, encumber, sell, rent, lease, sublicense, or otherwise transfer rights to the Software; (e) remove or alter any trademark, logo, copyright or other proprietary notices, legends, symbols or labels in the Software; (f) without Nuance's prior written consent (which may be given or withheld in Nuance's sole discretion) either (1) provide service bureau services using the Software, or (2) otherwise enter into an agreement with a third party to use the Software on such third party's behalf for the primary purpose of obviating the third party's need to license the Software itself; or (g) copy the printed materials accompanying the Software. As between Licensee and Nuance, any changes to, modifications to, or derivative works of the Software shall become the exclusive property of Nuance.

TERMINATION. Without prejudice to any other rights, Nuance may terminate this Agreement if Licensee breaches any of its terms and conditions. Upon termination, Licensee shall destroy all copies of the Software.

PROPRIETARY RIGHTS. Title, ownership rights, and intellectual property rights in the Software shall remain in Nuance and/or its suppliers or licensors. Licensee acknowledges such ownership and intellectual property rights and will not take any action to jeopardize, limit or interfere in any manner with Nuance's or its suppliers' or licensors' ownership of or rights with respect to the Software. The Software is protected by copyright and other intellectual property laws and by international treaties.

DISCLAIMER OF WARRANTY. THE SOFTWARE (INCLUDING WITHOUT LIMITATION THE RELATED DOCUMENTATION) IS PROVIDED ON AN "AS IS" BASIS, WITHOUT WARRANTY OF ANY KIND, INCLUDING WITHOUT LIMITATION THE WARRANTIES THAT IT IS FREE OF DEFECTS, MERCHANTABLE, FIT FOR A PARTICULAR PURPOSE OR NON-INFRINGEMENT. THE ENTIRE RISK AS TO THE QUALITY AND PERFORMANCE OF THE SOFTWARE IS BORNE BY LICENSEE. SHOULD THE SOFTWARE PROVE DEFECTIVE IN ANY RESPECT, LICENSEE AND NOT NUANCE OR ITS SUPPLIERS, LICENSORS OR RESELLERS ASSUMES THE ENTIRE COST OF ANY SERVICE AND REPAIR. THIS DISCLAIMER OF WARRANTY CONSTITUTES AN ESSENTIAL PART OF THIS AGREEMENT. NO USE OF THE SOFTWARE IS AUTHORIZED HEREUNDER EXCEPT UNDER THIS DISCLAIMER. SOME STATES OR JURISDICTIONS DO NOT ALLOW THE EXCLUSION OF IMPLIED WARRANTIES OR LIMITATIONS ON HOW LONG AN IMPLIED WARRANTY MAY LAST, SO THE ABOVE LIMITATIONS MAY NOT APPLY TO YOU.



LIMITATION OF LIABILITY. TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, IN NO EVENT WILL NUANCE OR ITS SUPPLIERS OR LICENSORS BE LIABLE FOR ANY INDIRECT, SPECIAL, INCIDENTAL, ECONOMIC OR CONSEQUENTIAL DAMAGES ARISING OUT OF THE USE OF OR INABILITY TO USE THE SOFTWARE PACKAGE OR DOCUMENTATION, INCLUDING, WITHOUT LIMITATION, DAMAGES FOR LOSS OF GOODWILL, WORK STOPPAGE, COMPUTER FAILURE OR MALFUNCTION, OR ANY AND ALL OTHER COMMERCIAL DAMAGES OR LOSSES, EVEN IF ADVISED OF THE POSSIBILITY THEREOF, AND REGARDLESS OF THE LEGAL OR EQUITABLE THEORY (CONTRACT, TORT OR OTHERWISE) UPON WHICH THE CLAIM IS BASED, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. IN ANY CASE, NUANCE'S AND ITS SUPPLIERS' AND LICENSORS' ENTIRE LIABILITY UNDER ANY PROVISION OF THIS AGREEMENT SHALL NOT EXCEED IN THE AGGREGATE THE SUM OF THE FEES LICENSEE PAID FOR THE SOFTWARE PACKAGE (IF ANY), SOME STATES OR JURISDICTIONS DO NOT ALLOW THE EXCLUSION OR LIMITATION OF INCIDENTAL OR CONSEQUENTIAL DAMAGES, SO IN THOSE STATES OR JURISDICTIONS, THE ABOVE LIMITATION OR EXCLUSION MAY NOT BE APPLICABLE. AS A CONDITION OF YOUR USE OF THE SOFTWARE PACKAGE, YOU AGREE TO INDEMNIFY NUANCE FOR ALL CLAIMS RELATING TO YOUR USE, REPRODUCTION AND/OR RECEIPT OF CONTENT THROUGH USE OF THE SOFTWARE PACKAGE.

EXPORT CONTROL. Licensee agrees to comply with all applicable export laws and restrictions and regulations of the United States of America or foreign agencies or authorities, and not to export or re-export the Software in violation of any such restrictions, laws or regulations, or without all necessary approvals.

U.S. GOVERNMENT END USERS. This Section applies to all acquisitions of the Software by or for the government of the United States of America ("government") or by any prime contractor or subcontractor (at any tier) under any contract, grant, cooperative agreement, or other activity with the government. By accepting delivery of the Software, the government hereby agrees that this software qualifies as "commercial" computer software within the meaning of the acquisition regulation(s) applicable to the procurement. The terms and conditions of this Agreement shall pertain to the government's use and disclosure of the Software and shall supersede any conflicting contractual terms or conditions. If this Agreement fails to meet the government's needs or is inconsistent in any respect with the federal law of the United States of America, the government agrees to return the Software, unused, to Nuance. The following additional statement applies only to acquisitions governed by DFARS Subpart 227.4 (October 1988): "Restricted Rights—Use, duplication, and disclosure by the Government is subject to restrictions as set forth in subparagraph (c)(1)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.227-7013 (Oct. 1988)." In the event any of the above referenced agency regulations is amended or replaced, the equivalent successor regulation shall apply instead.

GENERAL. This Agreement constitutes the entire agreement between the parties concerning the subject matter hereof. This Agreement may be amended only by a writing signed by both parties. The terms and conditions as set forth in any purchase order which differ from, conflict with, or are not included in this Agreement shall not become part of this Agreement unless specifically accepted by Nuance in writing. Except to the extent, if any, applicable law requires otherwise, this Agreement shall be governed by the laws of the Commonwealth of Massachusetts, United States of America, excluding its conflict of law provisions. Unless otherwise agreed in writing, all disputes relating to this Agreement (excepting any dispute relating to Nuance's or its licensors' or suppliers' intellectual property rights) shall be subject to final and binding arbitration in Boston, Massachusetts, United States of America under the auspices of a single arbitrator pursuant to the commercial arbitration rules of the American Arbitration Association then in effect, with the losing party paying all costs of arbitration. This Agreement shall not be governed by the United Nations Convention on Contracts



for the International Sale of Goods. If any provision in this Agreement should be held illegal or unenforceable by a court having jurisdiction, such provision shall be modified to the extent necessary to render it enforceable without losing its intent, or severed from this Agreement if no such modification is possible, and the other provisions of this Agreement shall remain in full force and effect. The controlling language of this Agreement is English. The Licensee agrees to bear any and all costs of interpreters if necessary. If Licensee has received a translation into another language, it has been provided for Licensee's convenience only. A waiver by either party of any term or condition of this Agreement or any breach thereof, in any one instance, shall not waive such term or condition or any subsequent breach thereof. The provisions of this Agreement that require or contemplate performance after the expiration or termination of this Agreement shall be enforceable notwithstanding such expiration or termination. Licensee may not assign or otherwise transfer by operation of law or otherwise this Agreement or any rights or obligations herein except to an acquirer of Licensee's business in the case of a merger or the sale of all or substantially all of Licensee's assets to such acquirer. This Agreement shall be binding upon and shall inure to the benefit of the parties, their successors and permitted assigns. The relationship between Nuance and Licensee is that of independent contractors and neither Licensee nor its agents shall have any authority to bind Nuance in any way. If any dispute arises under this Agreement, the prevailing party shall be reimbursed by the other party for any and all legal fees and costs associated therewith. The headings to the sections of this Agreement are used for convenience only and shall have no substantive meaning. Nuance may use Licensee's name in any customer reference list or in any press release issued by Nuance regarding the licensing of the Software.

LICENSEE OUTSIDE THE U.S. If Licensee is located outside the U.S., then the provisions of this Section shall apply. (a) Les parties aux présentes confirment leur volonté que cette convention de même que tous les documents y compris tout avis qui s'y rattache, soient rédigés en langue anglaise. (Translation: "The parties confirm that this Agreement and all related documentation is and will be in the English language.") (b) Licensee is responsible for complying with any local laws in its jurisdiction which might affect its right to import, export or use the Software, and Licensee represents that it has complied with any regulations or registration procedures required by applicable law to make this license enforceable.

COUNTRY UNIQUE TERMS. For licenses granted outside of the United States or Canada, the following terms are in addition to, or replace or modify the referenced terms above. All terms above that are not changed by these amendments remain unchanged and in effect.

Governing Law. The phrase "this Agreement shall be governed by the laws of the Commonwealth of Massachusetts." is replaced as follows "This Agreement shall be governed by the laws of the Republic of Ireland."

© 2013 Nuance Communications. All Rights Reserved. Nuance and the Nuance logo are trademarks or registered trademarks of Nuance Communications, Inc. or its affiliates in the United States and/or other countries.

10. Duration

The agreement runs for an indefinite time. The right of the license holder to use the Foxit Europe Software Products are terminated automatically, without notice from Foxit Europe, if the licensee fails to comply with any term(s) of this license. On termination of the license the license holder is obliged destroy the original version of the product, together with any copies and documentation.

11. Warranty



The license issuer gives a warranty only in within the scope of the law. The Foxit Europe Software Products are under warranty for twelve months from the date of delivery. The warranty covers only product redelivery or product improvement. If these remedies are unsuccessful then the license fee may be reduced or the contract cancelled. Altered, modified, changed or extended Foxit Europe Software Products are not covered by the warranty.

12. Disclaimer of Warranty

Foxit Europe calls your attention to the fact that you are liable for all damage to Foxit Europe arising from copyright violation. Foxit Europe shall not be liable for any damage incurred by the use of Foxit Europe Software Products including loss of profit, loss of information, loss of data or any financial loss. No damages shall be paid by Foxit Europe unless instructed to by law.

13. Export

If you wish to export the Foxit Europe Software Product then you are responsible for obeying all export regulations for the country of purchase.

14. Controlling Law and Severability

If there is a local subsidiary of Foxit Europe in the country in which the Foxit Europe Software Product was purchased, then the local law in which the subsidiary sits shall govern this license. Otherwise, the license shall be governed by the German Law. If for any reason a court of competent jurisdiction finds any provision, or portion thereof, to be unenforceable, the remainder of this license shall continue in full force and effect.

15. Complete Agreement

This license constitutes the entire agreement between the parties with respect to the use of the Foxit Europe Software Product and supersedes all prior or contemporaneous understandings regarding such subject matter. No amendment to or modification of this license will be binding unless in writing and signed by Foxit Europe.

16. Additional Items

Only the product sold under the name "C-SDK incl. Runtime-License" can be acquired as a flat-rate package. It consists of a LuraWave JP2 C-SDK and a runtime license which is valid no matter how many copies you make of the target product or how long you keep the license. However, the runtime license is limited to the use of just a single target product. Only for the above-named product no additional runtime licensing agreement is necessary if in your order you indicate a clearly decipherable product name for the target product into which you will integrate the LuraWave JP2 with the C-SDK.

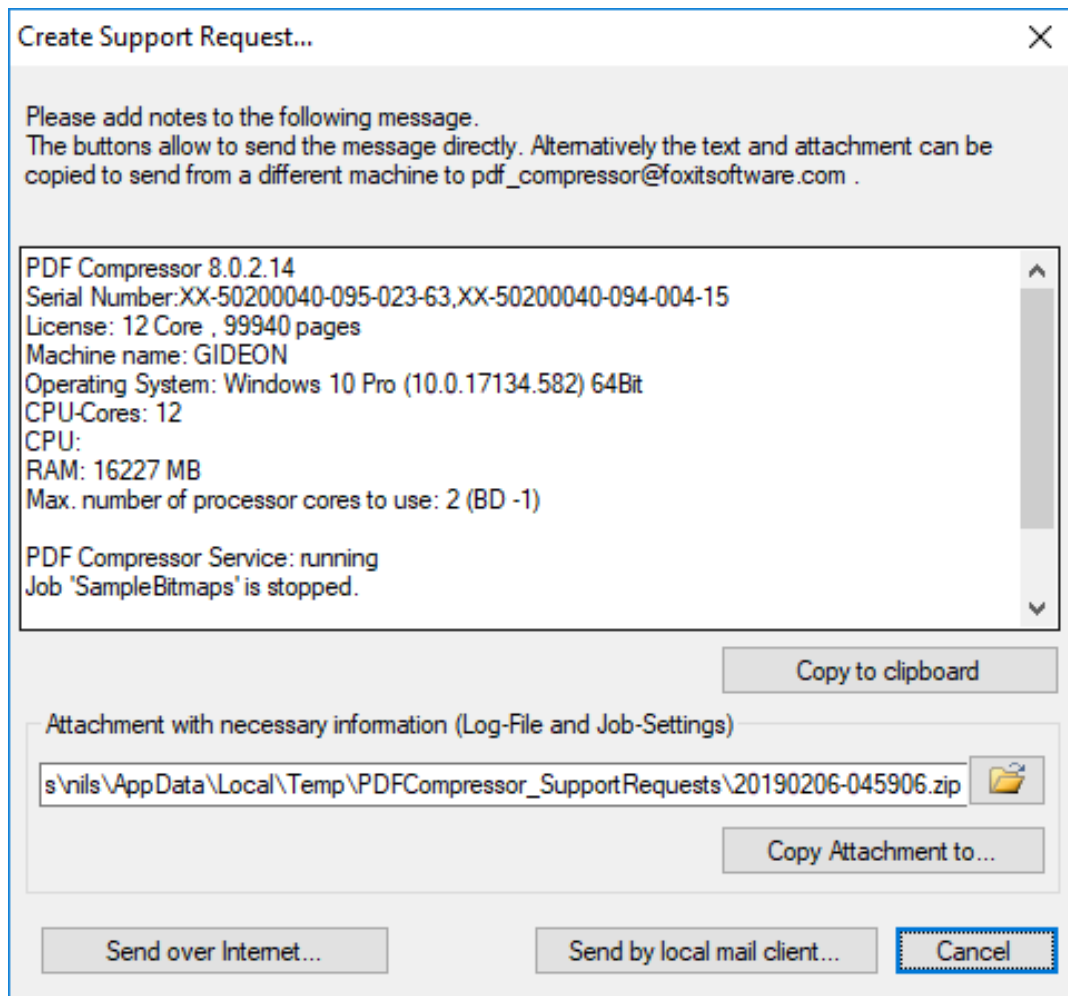
© 2019

Foxit Europe GmbH
Neue Kantstr. 14
14057 Berlin
GERMANY



Chapter 13. Support and Contact

If you incur any problems installing or using PDF Compressor you can contact the Foxit Europe support via e-mail. Please use the option *Create Support Request...* in menu *Help* or in the error dialog.



Please add some notes regarding the error or problem in the text field. By opening the support request dialog a zip file is created including necessary files like job settings and log. The dialog displays the zip file path in the attachment field. In case the machine is connected to the internet, you can send the message (including attachment) directly to Foxit by clicking on the *Send over Internet...* button. After clicking on *Send over Internet...*, you will also be asked to enter some sender information.

A local e-mail client can be used for sending using the *Send by local mail client...* button. In case of failure or missing internet connection please copy the message and attachment to a machine with internet connection and send it from there.

If you don't use the option to send a support request, please include the following information:

1. The version number of PDF Compressor, e.g. 8.0.2.14. You find this number in the *About...* dialog, which you open either per *Help* → *About PDF Compressor...* or by clicking the button in the toolbar.



2. The operating system on your computer, e.g. Windows 8.1 or Windows 10 v1803.
3. A description of the error and the logfile as specified under *File* → *Options* → *Log file*.

Please send all support requests to: pdf_compressor@foxitsoftware.com

You can find further information and contact data on our web page: <http://www.foxitsoftware.com>

Version 8.1

Date 2019-07-11

The specifications contained herein are subject to change without notice. Foxit Europe shall not be liable for errors contained herein or for incidental or consequential damages (including lost profits) in connection with the furnishing, performance or use of this material whether based on warranty, contract, or other legal theory.

© 2019 Foxit Europe GmbH
Neue Kantstr. 14
14057 Berlin
Germany