



AUTOMATIC PDF PROCESSOR 1

PROGRAM HELP

GILLMEISTER SOFTWARE

WWW.GILLMEISTER-SOFTWARE.COM

1 TABLE OF CONTENTS

1	Table of contents.....	1
2	Start	3
3	Main menu – menu entries of the group Menu.....	3
3.1	Options... ..	3
3.2	Enter license key.....	3
3.3	Help	3
3.4	About... ..	3
3.5	Other --> Create error report	3
3.6	Other --> Clear error logs	3
3.7	Other --> Backup application data.....	4
3.8	Other --> Restore application data.....	4
3.9	Exit	4
4	Buttons of the profile toolbar	4
4.1	New profile... ..	4
4.2	Edit.....	4
4.3	Duplicate.....	4
4.4	Catch up.....	4
4.5	Activate.....	5
4.6	Deactivate.....	5
4.7	Delete	5
4.8	Other	5
5	Profile list.....	5
5.1	Status.....	6
5.2	Name	6
5.3	Last use (regularly)	6
5.4	Last use (subsequently).....	6
5.5	Monitored folders	6
5.6	Comment.....	6
5.7	Context menu of the profile list	6
6	Profile settings.....	6
6.1	General	6
6.2	Monitored folders	7
6.3	filter	7
6.4	Example files.....	8
6.5	Data extraction.....	8

6.6	Extraction rules.....	8
6.7	Task type: File operations (rename, move, copy ...)	11
6.8	Task type: Print.....	12
6.9	Notifications	13
7	Toolbar of the log area	13
7.1	Filter log entries.....	13
7.2	Create Excel report.....	14
8	Log List.....	14
9	Status bar.....	14
9.1	Profiles.....	14
9.2	Log filter.....	15
9.3	Status.....	15
10	Program Options	15
10.1	General	15
10.2	Processing.....	15
10.3	Email account for sending	16
10.4	Excel report by email.....	16
10.5	Backups.....	16
10.6	Print settings.....	16
10.7	Other	16
11	EULA	16

2 START

With Automatic PDF Processor, PDF files can be processed automatically. Any number of folders can be monitored to automatically print, rename, or move incoming PDF files to dynamically named folders. Numerous metadata of the PDF document can be used in addition to the content, for example to include invoice information in the file path. The PDF files to be processed can be narrowed down with various profile-specific filters - the document text, PDF metadata and general file information are available as filters.

The program only processes PDF files. If the program options have been set to move or send PDF files without text (e. g., scanned documents), this is done immediately after the files have been checked and no profiles will be applied. With a solution like OmniPage Ultimate, PDF image files can also be automatically converted into searchable PDF files in-place and then processed by the Automatic PDF Processor.

The help describes the program functions and provides instructions for the use of Automatic PDF Processor.

3 MAIN MENU – MENU ENTRIES OF THE GROUP MENU

3.1 OPTIONS...

Use this menu item to open the dialog box for customizing the program options.

3.2 ENTER LICENSE KEY...

This menu entry opens the registration dialog to enter the license key respectively to unlock the full version of the program.

3.3 HELP

Clicking on this menu item opens the online help in the default browser. To use the online help, an active Internet connection is required.

3.4 ABOUT...

Using this menu item, the dialog window to display the program version of the license state will be opened. This dialog window contains also links to contact the technical support, the product's web page etc.

3.5 OTHER --> CREATE ERROR REPORT

In case of technical problems, you can create a bug report by using this menu item. The created file is named "Automatic PDF Processor - error report" and located on the Desktop. You can then send us the error report attached to an email with a short description of the problem.

3.6 OTHER --> CLEAR ERROR LOGS

This menu item allows the error log files to be emptied manually. However, the program automatically removes entries that are older than 7 days during every saving process.

3.7 OTHER --> BACKUP APPLICATION DATA...

This menu item allows you to store the entire application data (profiles, log, etc.) as ZIP archive in a directory to be selected. After saving the ZIP archive is highlighted in the Explorer.

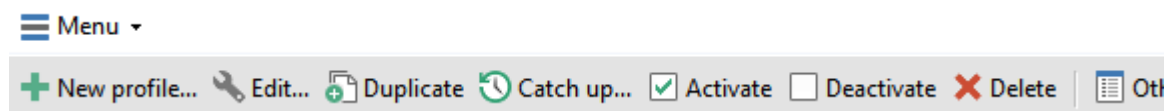
3.8 OTHER --> RESTORE APPLICATION DATA...

This menu entry can be used to restore previously saved application data.

3.9 EXIT

Use this menu item to exit the program.

4 BUTTONS OF THE PROFILE TOOLBAR



4.1 NEW PROFILE...

A click on this button opens a window for creating a new profile.

4.2 EDIT...

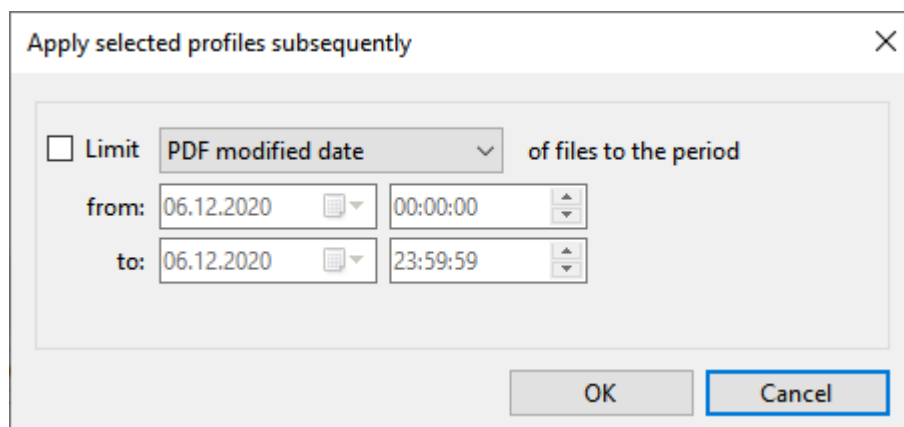
This button opens the window for editing the settings of the profile that is currently selected in the list.

4.3 DUPLICATE

Creates a copy of the selected profile.

4.4 CATCH UP...

This button allows you to apply the selected profiles retrospectively to all PDF files that meet the respective profile-specific filter criteria. Optionally, the search process can be limited to a certain period.



4.5 ACTIVATE

A click on this button activates the selected entries of the profile list, i.e., the profiles get the status Active and are applied to newly incoming PDF documents.

4.6 DEACTIVATE

This button sets the selected profiles to the status Inactive, i.e., the profiles are ignored when processing newly incoming PDF documents.

4.7 DELETE

By clicking the button, you can delete the selected profiles after a confirmation prompt.

4.8 OTHER

Contains the following entries:

4.8.1 Apply all profiles subsequently...

This menu entry opens a new window in which you can optionally specify the period of a file date. After confirmation, all PDF files of the monitored folders will be processed subsequently using all active profiles.

4.8.2 Activate all profiles

With this button, you can set all profiles to the status Active, i.e., the profiles are applied to newly incoming PDF documents.

4.8.3 Deactivate all profiles

This button sets all profiles to the status Inactive, i.e., the profiles are ignored when processing newly incoming PDF documents.

4.8.4 Delete all profiles

Clicking this button deletes all profiles after a confirmation prompt.

4.8.5 Import profiles...

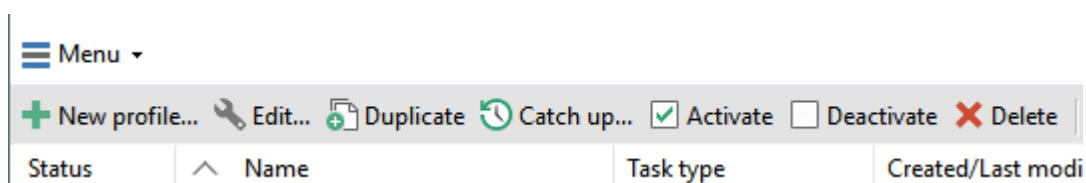
This menu item is used to import profiles from a JSON file.

4.8.6 Export profiles

Here you can export profiles, for example to transfer them to another computer. Choose from the following options:

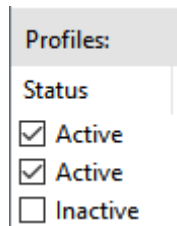
- Export all profiles
- Export active profiles
- Export selected profiles

5 PROFILE LIST



5.1 STATUS

Newly created profiles are given the status Active. All profiles with this status will be applied to incoming files. To disable a profile, click the check box at the beginning of each row. The status then changes to Inactive.



A status change causes the profile change date to be altered. A profile is only automatically applied to files whose detection date lies after the profile creation date and the profile change date.

5.2 NAME

The name of the profile.

5.3 LAST USE (REGULARLY)

Shows the date and time of the last successful regular respectively automatic application of the profile.

5.4 LAST USE (SUBSEQUENTLY)

Shows the date and time of the last successful application of a profile using the Catch-Up function.

5.5 MONITORED FOLDERS

A list of all folders the profile will be applied to.

5.6 COMMENT

An optional comment to the profile.

5.7 CONTEXT MENU OF THE PROFILE LIST

The context menu of the profile list contains the following entries:

- Select all
- Select none
- Invert selection
- Export profile

6 PROFILE SETTINGS

6.1 GENERAL

6.1.1 Name

Here you can give the profile a meaningful name.

6.1.2 Comment

Optionally, you can enter a comment that will be displayed in the profile list.

6.1.3 Color highlighting in the log list

Here you have the option to determine a color in which the profile will be highlighted in the log list.

6.1.4 Password for PDF files (if needed)

If necessary, enter the required password for the PDF files here. The program tries to open the files without a password at first and only uses the password stored here if necessary.

6.2 MONITORED FOLDERS

Here you can add one or more folders. The profile is applied to all added folders if the specified filter criteria are met. When the option Including Sub-Folders is activated, additionally, all PDF files from folders which are subordinate to the added folders are processed.

6.3 FILTER

Here you specify the (optional) conditions that must be met by the file properties. Only if all filter criteria are met, the PDF document is processed. You can use logical comparison operators to set several conditions for a file property. In this case, the AND operator takes precedence over the OR operator. The filter is case insensitive - no distinction is made between uppercase and lowercase letters of the entered terms. Wildcards as the asterisk are not supported. Use the placeholder Regex for this. Please note that a Regex placeholder must either be the only filter value or must be surrounded by logical operators only. Example:

<BeginOfRegex>^Offer<EndOfRegex><AND>Company XYZ

The screenshot shows a web-based configuration interface for filters. At the top, there's a 'Filter' header. Below it, two tabs are visible: 'File properties (1)' and 'PDF data'. The 'File name' section has a 'contains:' label and a text input field containing 'Invoice<AND>'. To the right of this field is a link 'Insert operator/placeholder'. Below the input field, there are two buttons: 'Logical AND' and 'Logical OR'. A dropdown menu is open, showing 'Placeholders' and 'Logical operators' with right-pointing arrows. Below the 'File name' section, there's a 'File size' section with a dropdown menu currently set to 'not evaluated'.

Filter

File properties (1) PDF data (1)

Text of document

contains: [Insert operator/placeholder](#)

Invoice

does not contain: [Insert operator/placeholder](#)

Created with

contains: [Insert operator/placeholder](#)

6.4 EXAMPLE FILES

Specify here 5 or more PDF files that correspond to the ones you want to process. In the profile settings, you will get a preview of the extracted data, verification results etc. based on these sample files. If the PDF files are only to be printed, this specification is not necessary.

6.5 DATA EXTRACTION

In this category, you can create or change rules for extracting data. For already created rules, you can preview their following properties here:

- Name of the rule
- Comment
- Data source
- Data type
- Verification result
- Determined data

By double-clicking on the name of the rule, you can navigate directly to the properties of the rule to adjust them if necessary.

Values to be determined

Create/Edit rules... Example file for preview: < R20201204_1234.pdf > Open selected example file

Name of the rule	Comment	Data source	Data type	Verification result	Determined data
Invoice number		Document text (keyword)	Text	Verification successful	1234
Invoice date		Document text (keyword)	Date	Verification successful	04.12.2020
IBAN		Document text (position)	Text	Verification successful	DE05 1005 0000 0111 2222 33

6.6 EXTRACTION RULES

6.6.1 General

6.6.1.1 Name

Here you can give the extraction rule a meaningful name. The name given here is used in several places in the program. For example, it is listed in the list of available placeholders. When exporting as

a CSV file, it serves as an identifier for the respective column. It should therefore be short and concise. If you create several rules with the same name, it is sufficient if one of these rules produces a valid result.

6.6.1.2 Comment

Optionally, you can enter a comment that will be included in the rule overview.

6.6.1.3 Data source

The following sources are available for the extraction of data:

- Determine data from document text
 - by using a keyword (variable position of data, for invoices etc.)
 - by specifying the position (fixed position of the data, for forms etc.)
- Using document metadata (title, author, subject, keywords, created with, program name, date created, date modified, version number, number of pages)
- Use file information (file name, folder name, path, path including file name, creation date, modification date)
- Custom Text

Depending on the selection, different configuration options or tabs are available.

6.6.1.4 Data type

The following data types can be assigned to the extracted values:

- Text
- Date
- Number
- Query

Depending on the selection, among others different verification options are available. For the datatype Date, all components of the date are provided as placeholders for use as file names etc. The type Query allows to assign a certain value in case of existence/non-existence of a search term.

6.6.2 Data determination (Keyword)

6.6.2.1 Select file

Here, you can navigate through the individual sample files (specified in 6.4) to check the extraction result and to ensure that the expected extraction result is determined in all sample files. This is also possible in the rules overview - there the result of all established rules for the respective sample file is displayed.

6.6.2.2 Determine page

This step is optional here and only necessary if the document contains several pages and if the keyword can occur several times at unpredictable positions. If the occurrence of the keyword specified in the next step (for example "Invoice date:") always precedes the value to be determined, the default setting No Determination Necessary can be left here. Otherwise, either the corresponding page number or a unique keyword (including occurrence) can be entered to determine the page.

Note: Depending on the structure of the page, the first visible occurrence is not necessarily the first occurrence in the document structure. In this case, the occurrence number can be specified accordingly.

6.6.2.3 Set data area

Enter here the search term (for example "Invoice date:") and its occurrence (*a .NET-compatible regular expression can also be entered as a search term*). Then, you specify the data position, for example "To the right of the found location" if the value to be extracted directly follows the search term. There are two options for specifying the data range:

- Text block
- First character

The default setting Text Block covers all subsequent characters of the text block adjacent to the search word and is sufficient in most cases. However, if the text block overlaps into an unwanted adjacent data range, you must switch to the First Character setting. With this setting, only the first visible character of the text block is used as the extraction result - the data range must therefore be extended in almost all cases.

The preview located below the configuration area shows the currently extracted value.

6.6.2.4 Extend data area/Adjust data area extension

Here the data range can be extended or (for example by specifying negative values) reduced. The following options are available:

- by X mm (or point)
- up to before keyword
- up to behind keyword
- up to page margin

6.6.3 Data determination (Position)

This data retrieval type is only suitable for PDF documents in which the data is always found in the same position (mainly forms or for documents with fixed data in the footer).

Unlike keyword data retrieval, the page of the document to be used must be specified. The data to be extracted is defined here using a resizable selection rectangle. The selection rectangle can be unlocked or locked, i.e., protected against unintentional changes, by clicking the Change/Lock Position button.

The preview located below the configuration area shows the currently extracted value.

6.6.4 Clean-up

This tab provides numerous functions for cleaning up the text. For example, surrounding spaces or additionally captured text lines can be removed. Furthermore, among other things, an addition to the text is possible. However, the main purpose of the cleanup is to prepare the extracted text for verification. In order to keep it as simple as possible, it is better to keep the text short here and format it only in a further step.

6.6.5 Verification

Here you can set up criteria which must be met by the extracted and cleaned text. For example, the number of extracted characters or the occurrence of certain terms can be verified. Failed verifications result in a processing error and are listed in the details in the log area on the Errors tab. For example, a different number of characters would be visible there as "Error: Number of characters".

6.6.6 Formatting

This tab is only available for the data type Text and provides functions for final editing of the extracted text.

6.7 TASK TYPE: FILE OPERATIONS (RENAME, MOVE, COPY ...)

6.7.1 General settings

Here you can define whether the task is assigned the status Activate and will therefore be executed when there are matching PDF files, or whether it is assigned the status Inactive. Activate tasks are given a check mark in the task name.

6.7.2 Storage location

6.7.2.1 Folder

Select the base directory where the attachments should be stored by clicking on the button with the label "...".

6.7.2.2 Subfolders

Use the Link Menu above the input field to select file properties that should be used to create an optional subfolder structure within the base directory.

Storage location

Base directory:
D:\Invoices

Subfolders (dynamically generated from ...)
<RuleId:2(Invoice date){Year4}>

File name (dynamically generated from ...)

Preview:
D:\Invoices\2020\R20201204_1234.pdf

If a file with that name already exists:
Add numbering

Invoice number
Invoice date

Extracted data
File properties
PDF properties
Current date
Path separator

6.7.2.3 File name

Use the Link Menu above the input field to select file properties from which the file name should be generated. If you leave this field empty, the original name of the PDF file is used.

6.7.2.4 Name collisions

If a file with the distinguished name already exists, the collision rule determined here will be applied. Decide whether the program should: overwrite the file, add a number or the processing date to the name, or cancel the operation.

6.7.2.5 Execute program

For automatic further processing, the path of the stored file can be passed to a program as parameter. Here, you can optionally enter the path of a program that should be executed after storing the file.

6.7.2.6 Parameter

Optionally enter one or more parameters that should be passed to the program to be executed by using the link menu.

6.8 TASK TYPE: PRINT

6.8.1 General settings

Here you can define whether the task is assigned the status Activate and will therefore be executed when there are matching PDF files, or whether it is assigned the status Inactive. Activated tasks are given a check mark in the task name.

6.8.2 Print settings

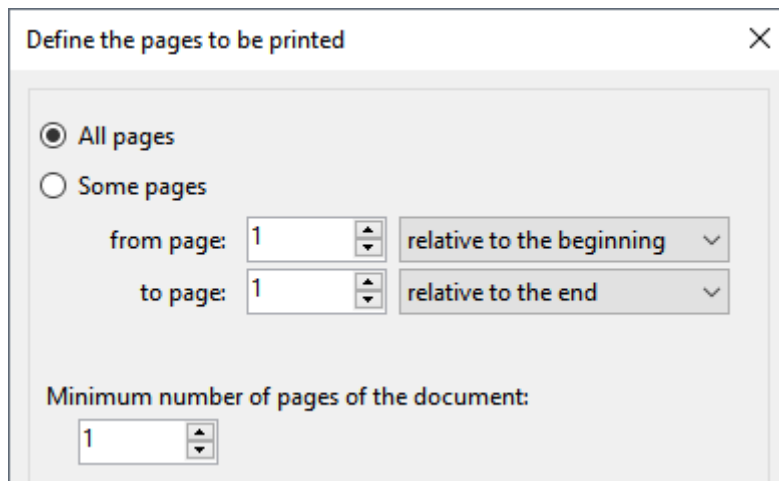
Here you determine whether matching files should be printed immediately after receipt. Activate the respective check box, and then make the appropriate settings on the printer selection dialog. Confirm these in the printer selection dialog by clicking the Print button.

The screenshot shows a 'Print file' dialog box with three tabs: 'Printout #1 (active)', 'Printout #2', and 'Printout #3'. The 'Printout #1 (active)' tab is selected. Below the tabs, there are two sections: 'Task status' and 'Print settings'. In the 'Task status' section, there are two radio buttons: 'Active' (which is selected) and 'Inactive'. In the 'Print settings' section, there are three rows of controls. The first row has a 'Print settings...' button and the text 'Selected printer: HP OfficeJet Pro 8730'. The second row has a 'Page range...' button and the text 'Pages: All pages'. The third row has a dropdown menu showing 'Auto-rotate pages: Preset' and the text 'Program option preset => Yes'.

6.8.2.1 Page range

By using the Page Range dialog, you can specify which pages to print. For example, to skip the first page, select the type "Some pages", enter the value 2 in the field "from page", and select "relative to beginning". In the field "To page", enter the value 1 and select "Relative to end".

With the "Minimum number of pages of the document", you can ensure that only documents corresponding to this setting are printed.



6.8.2.2 Automatically rotate pages

Determines whether the pages to be printed should be rotated automatically. This is necessary, for example, if some documents/pages to be processed are in portrait and some in landscape format. The default setting is set in the program options.

6.9 NOTIFICATIONS

Here you can make settings that are used to tell the processing status of a file.

6.9.1 Status notification via email

Select the case in which a notification is to be made:

- on success
- on errors
- if filter criteria are not fulfilled

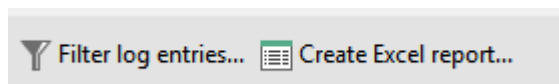
and enter one or more (comma-separated) recipients of the status message.

6.9.2 Status notification by an acoustic signal

Here you can optionally select a wave file that you want to play after successful processing.

7 TOOLBAR OF THE LOG AREA

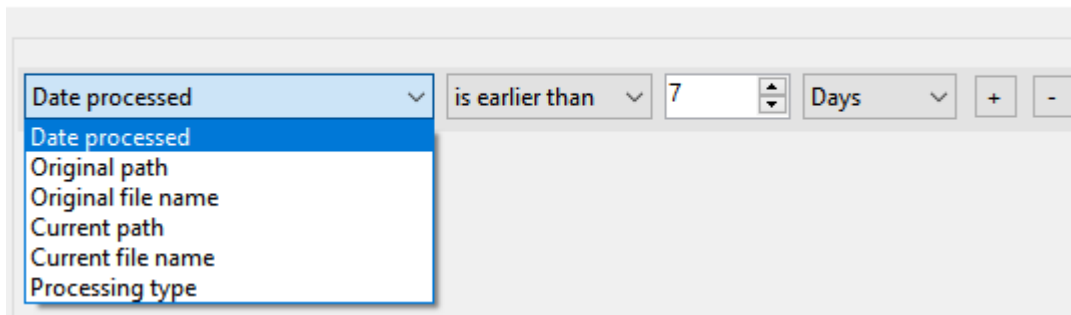
The toolbar contains the following buttons:



7.1 FILTER LOG ENTRIES

By using the log filter, log entries can be limited. The filter criteria may be, for example, the date of receipt, which is to be limited to a period within the last week. The comparison operator is context-sensitive, that is, he always fits the selected filter criteria. By clicking on the button with the plus or minus symbol, further filters can be added (linked with a logical AND) or the respective filter will be removed.

Log filter



7.2 CREATE EXCEL REPORT

This button is used to generate an overview of the processings as an Excel report. This contains the following sheets:

- Overview
- Successful processing
- Errors
- No match
- No text

and allows filtering and custom sorting of processing data. In the program options, you can configure a periodic sending of the current report.

8 LOG LIST

The log list contains information about all processed files and is divided into the four tabs: "Successful processed", "Error", "No match", "No text".

Depending on the active tab page, the context menu contains the following entries:

- Open saved file
(opens a saved file with the associated program)
- Select saved file in Explorer
(opens the Explorer and selects the saved file)
- Copy log entry to clipboard
(copies relevant data of the selected entry to the clipboard)
- Clear list
(removes all entries from the log list)

9 STATUS BAR

9.1 PROFILES

Shows the total number of profiles and the number of active profiles.

9.2 LOG FILTER

Shows the number of entries that do not fulfill the conditions of the log filter, for example, entries that are too old.

9.3 STATUS

Displays the following status information:

- the number of seconds until the next check for new PDF files
- information about the current check (the number of files already checked)
- information on current processing (the number of files already processed)
- other information about the current process

10 PROGRAM OPTIONS

10.1 GENERAL

Here you can define settings such as the language of the program interface and the startup behavior as well as a centrally managed profile file.

As soon as a centrally managed profile file has been specified, no further changes can be made to the individual profiles - even deletion and recreation are blocked. Changes can or should only be made by the administrator. The administrator edits the profile file locally (i.e., using the application data directory "%AppData%\Automatic PDF Processor") and then copies it to the designated network folder (provided that he does not use a centrally managed profile file himself).

The centrally managed profile file is reloaded by the individual processes before each run and the profile list in the main window is updated at regular intervals (every 5 minutes). Updating the profile list in the main window can be forced by minimizing the application to the notification area of the taskbar and then restoring it.

10.2 PROCESSING

10.2.1 Interval of the check for new files to be processed

Determines at what interval (in seconds) "Automatic PDF Processor" checks whether new files have been stored.

10.2.2 Pause between processing multiple new files

Here you can set whether the program should pause for a certain time after processing a file. This option is relevant for the further processing of saved files.

10.2.3 Log errors for the generation of error reports

Determines whether errors are recorded. If problems occur, this option must be activated because detailed information about the error is necessary to generate an error report. Otherwise, this option can be disabled to reduce write accesses to the hard disk.

10.2.4 Report general errors by email

Determines whether general errors, such as missing write permissions, are reported by email.

10.2.5 Move files with faulty processing to

Enables the specification of a general directory for files that could not be processed properly and therefore have to be checked or processed manually.

10.2.6 Move files, which do not fulfill filter criteria, to

Allows to specify a common directory for files that do not meet all filter criteria.

10.2.7 Move files without text to

Allows you to specify a common directory for files that do not contain text (e.g., scanned files without OCR processing). Be sure to read section 2.0 (2nd paragraph) for more information.

10.2.8 Send files without text by email

Allows you to specify an email box for documents that do not contain text (e.g., scanned files without OCR processing). Be sure to read section 2.0 (2nd paragraph) for more information.

10.2.9 Process files only in the following time

This option allows files to be processed automatically only at the times specified here.

10.3 EMAIL ACCOUNT FOR SENDING

Here, the email account for sending notifications is set. You can use the default Outlook email account or provide custom settings. The account data specified here will be used to send status notifications to the recipients listed in the respective profile settings.

10.4 EXCEL REPORT BY EMAIL

This option allows you to define an interval at which an automatically generated Excel report about the processings of a specific period is sent to the specified email address.

10.5 BACKUPS

If this option is enabled, the program periodically creates a complete backup of the application data as ZIP archive in the specified folder. The backups can be used to restore application data to an earlier state via Menu → Other → Restore application data.

10.6 PRINT SETTINGS

Here you can define as a default setting whether pages to be printed should be rotated automatically. This setting can be overwritten in the profile settings.

10.7 OTHER

10.7.1 Maximum number of entries in the log area

Defines the maximum number of entries available in the log area. The default setting is 25,000 entries.

11 EULA

The license agreement can be found at the following address: <http://gillmeister-software.com/products/automatic-pdf-processor/download.aspx>