

# Table of Contents

- 1 Introduction ..... 13
  - 1.1 Welcome to Automatic PDF Processor 2 ..... 13
  - 1.2 Starting the Program ..... 14
  - 1.3 Main Window Overview ..... 14
  - 1.4 Basic Principle: Profiles ..... 14
  - 1.5 Getting Started ..... 15
  - 1.6 Processing Workflow ..... 15
  - 1.7 System Requirements ..... 16
  - 1.8 Help and Support ..... 16
- 2 Main Menu ..... 16
  - 2.1 Overview ..... 16
  - 2.2 Options ..... 16
  - 2.3 Enter License Key ..... 16
  - 2.4 Help ..... 17
  - 2.5 About ..... 17
  - 2.6 More ..... 17
  - 2.7 Exit ..... 19
  - 2.8 Additional Buttons in the Menu Bar ..... 19
- 3 Profile Toolbar ..... 19
  - 3.1 Overview ..... 19
  - 3.2 New Profile ..... 19
  - 3.3 Edit ..... 20
  - 3.4 Duplicate ..... 20
  - 3.5 Activate ..... 20
  - 3.6 Catch Up ..... 21
  - 3.7 Deactivate ..... 21
  - 3.8 Remove ..... 21
  - 3.9 More ..... 22
  - 3.10 Filter Area ..... 23
  - 3.11 Keyboard Shortcuts ..... 23
- 4 Profile List ..... 23

4.1 Overview.....	23
4.2 Columns.....	24
4.3 Adjusting Column Width .....	25
4.4 Sorting.....	25
4.5 Context Menu.....	25
4.6 Profile Colors .....	25
4.7 Multiple Selection .....	26
4.8 Tips for Profile Management.....	26
4.9 Central Profile Management.....	27
5 Log Toolbar .....	27
5.1 Overview.....	27
5.2 Filter Log Entries.....	27
5.3 Create Excel Report.....	28
5.4 Hide Log Entries .....	28
5.5 Restore Log Entries.....	29
5.6 Minimize/Maximize Log Area .....	29
5.7 Tips for Log Usage .....	29
6 Log List .....	30
6.1 Overview.....	30
6.2 Tabs .....	30
6.3 Context Menu.....	32
6.4 Color Coding.....	32
6.5 Tips for Log Analysis .....	33
7 Status Bar .....	33
7.1 Overview.....	33
7.2 Display Areas .....	33
7.3 Updates.....	34
7.4 Tips for Status Bar Usage .....	35
8 Profile Settings .....	35
8.1 Overview.....	35
8.2 General .....	36
8.3 Monitored Folders .....	38
8.4 Filter.....	39
8.5 Example Files.....	39

8.6 File Grouping.....	40
8.7 Data Extraction .....	41
8.8 Tasks.....	41
8.9 Notifications .....	42
8.10 Follow-up Operations .....	42
8.4 Filter .....	43
8.4.1 Overview and Basic Principle.....	43
8.4.2 Tab: File Properties .....	43
8.4.3 Tab: PDF Data .....	44
8.4.4 Tab: Results Preview .....	45
8.4.5 Tab: Overlap Check .....	46
8.4.6 AND/OR Logic .....	46
8.4.7 Regular Expressions .....	47
8.4.8 Number Ranges.....	48
8.4.9 Dynamic Filter Lists .....	49
8.4.10 Practical Tips.....	49
9 Extraction Rules .....	50
9.1 Overview.....	50
9.2 General .....	51
9.3 Data Source Document Text or Barcode - Determination: Position.....	52
9.4 Data Source Document Text or Barcode - Determination: Keyword .....	52
9.5 Data Source Document Text - Determination: Text of Page(s) .....	54
9.6 Data Types .....	54
9.7 Data Source: Form Field .....	55
9.8 Data Source: Sequential Number .....	55
9.9 Clean-up .....	56
9.10 Verification: Text.....	57
9.11 Verification: Date.....	58
9.12 Verification: Number.....	58
9.13 Verification: Query.....	59
9.14 Formatting .....	59
9.15 Using Placeholders .....	59
10 Tasks - Overview .....	60
10.1 What are Tasks? .....	60

10.1 Task List and Order .....	60
10.2 Common Concepts .....	62
10.2 Tips for Task Configuration .....	65
11 Rename File .....	66
11.1 Description .....	66
11.2 General Settings .....	66
11.3 Filename.....	66
11.4 Name Collisions .....	67
11.5 File Date .....	67
11.6 Afterwards .....	68
11.7 Example: Name Invoices by Date and Number.....	68
11.5 Tips and Notes .....	69
12 Copy File .....	69
12.1 Description .....	70
12.2 General Settings .....	70
12.3 Storage Location .....	70
12.4 Copy Additional Files .....	71
12.5 File Date .....	71
12.6 Afterwards .....	72
12.7 Example: Invoice Archive with Year/Month Structure.....	72
12.8 Creating Multiple Copies.....	73
12.6 Tips and Notes .....	73
13 Move File.....	74
13.1 Description .....	74
13.2 General Settings .....	74
13.3 Storage Location .....	74
13.4 Move Additional Files .....	75
13.5 File Date .....	76
13.6 Afterwards .....	76
13.7 Example: Sort Invoices by Supplier .....	77
13.8 Difference from Copy.....	77
13.6 Tips and Notes .....	78
14 Delete File .....	78
14.1 Description .....	78

14.2 General Settings .....	79
14.3 Delete Behavior .....	79
14.4 Delete Parent Folder .....	79
14.5 Example: Empty Inbox Folder After Processing .....	79
14.4 Tips and Notes .....	80
15 Print.....	80
15.1 Description .....	80
15.2 General Settings .....	81
15.3 Printer.....	81
15.4 Copies .....	81
15.5 Page Range .....	82
15.6 Duplex (Double-Sided Printing).....	82
15.7 Color .....	82
15.8 Paper.....	82
15.9 Scaling and Orientation.....	83
15.10 Example: Automatically Print Invoices.....	83
15.11 Print Presets.....	84
15.12 Batch Printing for Large Documents .....	84
15.13 Print to PDF File .....	85
15.14 Tips and Notes.....	86
16 Send Email .....	87
16.1 Description .....	87
16.2 General Settings .....	87
16.3 Recipients.....	87
16.4 Email Content .....	88
16.5 Attachments .....	89
16.6 Sending Options .....	90
16.7 Attachment Encryption .....	91
16.8 Example: Send Invoices to Customers .....	91
16.9 Sending Methods.....	92
16.7 Tips and Notes .....	92
17 Merge PDFs .....	93
17.1 Description .....	93
17.2 General Settings .....	94

17.3 Files to Merge .....	94
17.4 Storage Location .....	95
17.5 File Date .....	96
17.6 Afterwards .....	96
17.7 Example: Merge Invoice with Delivery Note.....	96
17.8 Example: Merge All Scans from a Folder .....	97
17.6 Tips and Notes .....	97
18 Split PDFs.....	98
18.1 Description .....	98
18.2 General Settings .....	99
18.3 Splitting Method .....	99
18.4 By Page Count .....	99
18.5 By File Size .....	99
18.6 By Top-Level Bookmarks .....	100
18.7 By Keywords.....	100
18.8 By Barcode.....	100
18.9 By Placeholder (Value Change) .....	101
18.10 By Separator Page with Text.....	102
18.11 By Blank Pages .....	102
18.12 Page Exclusion .....	102
18.13 Source for New Documents.....	102
18.14 Storage Location.....	103
18.15 File Date .....	104
18.16 Afterwards .....	104
18.17 Example: Split Combined Invoice by Customer .....	104
18.18 Example: Split Manual by Chapters .....	105
18.5 Tips and Notes .....	105
19 OCR (Text Recognition) .....	106
19.1 Description .....	106
19.2 General Settings .....	106
19.3 Language .....	107
19.4 DPI Setting (Resolution) .....	107
19.5 Image Optimization .....	108
19.6 Handling Files That Already Contain Text .....	108

19.7 Character Restriction.....	109
19.8 Processing .....	109
19.9 Storage Location .....	109
19.10 File Date .....	110
19.11 Afterwards .....	110
19.12 Example: Make Scanned Invoices Searchable .....	111
19.13 Example: Process Multilingual Documents .....	111
19.7 Tips and Notes .....	112
20 Save Attachments.....	112
20.1 Description .....	113
20.2 General Settings .....	113
20.3 Attachment Filter.....	113
20.4 Storage Location .....	114
20.5 File Date .....	114
20.6 Afterwards .....	115
20.7 Example: Extract ZUGFeRD XML .....	115
20.8 Example: Extract All Attachments from Document Collection.....	116
20.5 Tips and Notes .....	116
21 Save Extractions .....	117
21.1 Description .....	117
21.2 General Settings .....	117
21.3 Rules to Export .....	117
21.4 CSV Settings .....	118
21.5 Character Encoding.....	119
21.6 Storage Location .....	119
21.7 File Date .....	120
21.8 Afterwards .....	120
21.9 Example: Export Invoice Data for Accounting.....	121
21.10 Example: Individual CSV per PDF .....	121
21.6 Tips and Notes .....	122
22 Use Template .....	123
22.1 Description .....	123
22.2 General Settings .....	123
22.3 Template.....	123

22.4 Replacements .....	124
22.5 Embedding (optional).....	124
22.6 Storage Location .....	125
22.7 File Date .....	125
22.8 Afterwards .....	125
22.9 Example: XML File for Accounting System .....	125
22.6 Tips and Notes .....	126
23 Change File Protection .....	127
23.1 Description .....	127
23.2 General Settings .....	127
23.3 Remove Password .....	127
23.4 User Password .....	128
23.5 Permission Password .....	128
23.6 Permissions .....	129
23.7 File Date .....	129
23.8 Afterwards .....	129
23.9 Example: Protect Documents with Customer Number .....	129
23.10 Example: Allow Printing, Forbid Editing.....	130
23.5 Tips and Notes .....	130
24 Compress PDF .....	131
24.1 Description .....	131
24.2 General Settings .....	132
24.3 Image Compression.....	132
24.4 Remove Elements.....	133
24.5 Additional Options.....	133
24.6 Storage Location .....	134
24.7 File Date .....	134
24.8 Afterwards .....	134
24.9 Example: Optimize Scan Files for Archiving .....	134
24.5 Tips and Notes .....	135
25 Embed Attachments .....	135
25.1 Description .....	136
25.2 General Settings .....	136
25.3 Files to Embed.....	136

25.4 Duplicate Handling .....	136
25.5 Storage Location .....	137
25.6 File Date .....	137
25.7 Afterwards .....	137
25.8 Example: Add Excel Raw Data to Invoice .....	137
25.9 Example: Embed Multiple Supporting Files .....	138
25.5 Tips and Notes .....	138
26 Create ZUGFeRD/Factur-X .....	139
26.1 Description .....	139
26.2 General Settings .....	139
26.3 ZUGFeRD Version .....	140
26.4 Conformance Level .....	140
26.5 XML File .....	140
26.6 Storage Location .....	141
26.7 File Date .....	141
26.8 Afterwards .....	141
26.9 Example: Invoices for Government (XRechnung) .....	141
26.10 Example: Factur-X for French Customers .....	142
26.5 Tips and Notes .....	142
27 Convert to PDF/A .....	143
27.1 Description .....	143
27.2 General Settings .....	143
27.3 PDF/A Conformance Level .....	144
27.4 Storage Location .....	144
27.5 File Date .....	145
27.6 Afterwards .....	145
27.7 Example: Invoice Archiving .....	145
27.8 Example: Preparation for DMS .....	146
27.4 Tips and Notes .....	146
28 Remove Pages .....	147
28.1 Description .....	147
28.2 General Settings .....	147
28.3 Removal Criteria .....	147
28.4 Storage Location .....	148

28.5 File Date .....	149
28.6 Afterwards .....	149
28.7 Example: Remove Cover Pages from Scans .....	149
28.8 Example: Remove Blank Pages from Scans .....	150
28.9 Example: Remove Separator Pages from Batch Scan .....	150
28.5 Tips and Notes .....	151
29 Add Stamp/Watermark.....	151
29.1 Description .....	151
29.2 General Settings .....	152
29.3 Annotation Type .....	152
29.4 Stamp Settings .....	152
29.5 Watermark Settings .....	153
29.6 Page Selection .....	155
29.7 Embedding Mode.....	155
29.8 Storage Location .....	156
29.9 File Date .....	156
29.10 Afterwards .....	156
29.11 Example: Paid Stamp on Invoices .....	157
29.12 Example: Watermark for Confidential Documents .....	157
29.13 Example: Tiled Copyright Watermark.....	158
29.6 Tips and Notes .....	158
30 Insert Barcode .....	159
30.1 Description .....	159
30.2 General Settings .....	159
30.3 Barcode Format.....	160
30.4 Barcode Content .....	160
30.5 Position .....	161
30.6 Size .....	162
30.7 Page Selection .....	162
30.8 Storage Location .....	163
30.9 File Date .....	164
30.10 Afterwards .....	164
30.11 Example: QR Code with Archive Link .....	164
30.12 Example: Barcode with Invoice Number .....	164

30.13 Example: Tracking Code on All Pages .....	165
30.6 Tips and Notes .....	166
31 Replace Content .....	166
31.1 Description .....	167
31.2 General Settings .....	167
31.3 Configure Replacements.....	167
31.4 Placeholder (Extraction Rule) .....	168
31.5 Replacement Sources.....	168
31.6 Scope .....	170
31.7 Replace Images with Placeholder .....	171
31.8 Preview .....	171
31.9 Storage Location .....	172
31.10 File Date .....	172
31.11 Afterwards .....	172
31.12 Example: Anonymize Customer Data .....	173
31.13 Example: Replace Header on All Pages .....	173
31.14 Example: Update Date .....	174
31.15 Tips and Notes.....	174
32 Flatten Pages.....	175
32.1 Description .....	175
32.2 General Settings .....	176
32.3 Settings .....	176
32.4 Storage Location .....	176
32.5 File Date .....	177
32.6 Afterwards .....	177
32.7 Example: Seal Filled-In Forms .....	177
32.8 Example: Fix Annotations Before Distribution .....	178
32.5 Tips and Notes .....	178
33 Auto-rotate Pages .....	179
33.1 Description .....	179
33.2 General Settings .....	179
33.3 Settings .....	179
33.4 Storage Location .....	180
33.5 File Date .....	180

33.6 Afterwards .....	181
33.7 Example: Correct Scanned Documents.....	181
33.8 Example: Landscape Pages in Mixed Documents.....	181
33.5 Tips and Notes .....	182
34 Redact Text .....	182
34.1 Description .....	182
34.2 General Settings .....	183
34.3 Redactions.....	183
34.4 Storage Location .....	184
34.5 File Date .....	185
34.6 Afterwards .....	185
34.7 Example: Anonymize Personnel Data .....	185
34.8 Example: Remove Confidential Reference Numbers .....	186
34.5 Tips and Notes .....	186
35 Program Options.....	187
35.1 Overview.....	187
35.2 General.....	187
35.3 Processing .....	189
35.4 Email Account for Sending .....	191
35.5 Excel Report via Email.....	193
35.6 Log .....	193
35.7 Processing Log .....	194
35.8 Backups.....	195
35.9 Print Settings .....	195
35.10 Dynamic Lists.....	196
35.11 Password List .....	197
35.12 Other .....	197
35.13 Tips and Notes.....	198
36 Advanced Features .....	199
36.1 Command Line Parameters.....	199
36.2 Headless Mode / Service Operation .....	201
36.3 Processing Time Diagnostics .....	203
36.4 Catch-Up Processing.....	204
36.5 Retry on Errors.....	205

36.6 Tips and Notes .....	207
37 Language Installation .....	207
37.1 Overview .....	207
37.2 Available Languages .....	208
37.3 Installation .....	208
37.4 Uninstallation .....	208
37.5 Usage in Profiles .....	208
37.6 Tips and Notes .....	209
38 Catch-Up Processing .....	209
38.1 Overview .....	209
38.2 Procedure .....	210
38.3 Processing .....	210
38.4 Log .....	211
38.5 Use Cases .....	211
38.6 Tips and Notes .....	211
39 Processing Time Diagnostics .....	212
39.1 Overview .....	212
39.2 Tab: Runtime .....	212
39.3 Tab: Global Processing Times .....	212
39.4 Tab: Profile-Specific Processing Times .....	213
39.5 Tab: Waiting Files .....	213
39.6 Schedule Types .....	214
39.7 Troubleshooting Tips .....	214

# 1 Introduction

## 1.1 Welcome to Automatic PDF Processor 2

Automatic PDF Processor is a powerful Windows application for automated processing of PDF documents. The program monitors folders you specify and automatically performs the desired tasks on incoming PDF files.

### Typical Use Cases

- Automatically rename incoming invoices by date and invoice number
- Sort receipts into a structured folder hierarchy

- Extract invoice data (date, amount, invoice number) and export as CSV
  - Make scanned documents searchable using OCR
  - Automatically forward PDF files via email to defined recipients
  - Compress large PDF files to save storage space
  - Split multi-page PDFs according to specific criteria
  - Merge individual PDFs into a single document
  - Create ZUGFeRD-compliant electronic invoices
  - Add watermarks or stamps to documents
- 

## 1.2 Starting the Program

After installation, start Automatic PDF Processor from the Start menu or the desktop shortcut.

**Tip:** You can configure the program to start automatically with Windows. You can find this setting under *Extras > Program Options > General*.

---

## 1.3 Main Window Overview

The main window of Automatic PDF Processor is divided into several areas:

1. **Menu Bar** - Access to all program functions
  2. **Profile Toolbar** - Quick access for profile management
  3. **Profile List** - Overview of all created profiles
  4. **Log Toolbar** - Control of the processing display
  5. **Log List** - Log of processed files
  6. **Status Bar** - Current status information
- 

## 1.4 Basic Principle: Profiles

Working with Automatic PDF Processor is based on **profiles**. A profile defines:

- **Which folders** should be monitored
- **Which PDF files** should be processed (filter)
- **Which data** should be extracted from the PDFs (extraction rules)
- **Which tasks** should be performed (up to 20 different types)

You can create as many profiles as you need. Each profile works independently and can monitor different folders and perform different tasks.

**Example: Simple Archiving Profile**

A typical profile for automatic invoice archiving might look like this:

1. **Monitored Folder:** C:\Inbox\Invoices
  2. **Filter:** Document text contains “Invoice”
  3. **Data Extraction:** Extract invoice date and invoice number from text
  4. **Task:** Rename file using pattern <InvoiceDate>\_<InvoiceNumber>.pdf
  5. **Task:** Move file to D:\Archive\Invoices\<Year>\<Month>\
- 

## 1.5 Getting Started

### 1. Create a New Profile

Click **New Profile...** in the Profile Toolbar or select *File > New Profile...* from the menu.

### 2. Set the Monitored Folder

Switch to the **Monitored Folders** tab and add the folder that should be monitored.

### 3. Set Filters (optional)

If not all PDFs in the monitored folder should be processed, define appropriate criteria on the **Filter** tab.

### 4. Configure Tasks

Switch to the **Tasks** tab and activate the desired tasks. Configure each task according to your requirements.

### 5. Save and Activate Profile

Click **OK** to save the profile. The profile will immediately begin monitoring the specified folder.

**Important:** Files already present in the folder are initially only indexed (marked as “known”) and not automatically processed. To process existing files retroactively, use the **Catch Up** button in the Profile Toolbar. Newly added files, however, are processed automatically.

---

## 1.6 Processing Workflow

When a new PDF file appears in a monitored folder, the following process occurs:

1. **Detection** - The program detects the new file
2. **Filter Check** - It is verified whether the file matches the filter criteria
3. **Data Extraction** - Defined data is extracted from the PDF
4. **Task Execution** - All activated tasks are executed in sequence
5. **Logging** - The result is displayed in the Log List

**Tip:** In the Log List, you can view the status of each processed file. Double-clicking an entry shows details about the processing.

---

## 1.7 System Requirements

- Windows 10 or Windows 11
  - .NET 8.0 Runtime (automatically installed if needed)
  - At least 4 GB RAM recommended
  - For OCR: Additional storage space for language files
- 

## 1.8 Help and Support

The following resources are available for questions or issues:

- **Online Help:** Select *Menu > Online Help*
  - **Tutorials:** Step-by-step tutorials on our website
  - **Support:** [info@gillmeister-software.de](mailto:info@gillmeister-software.de)
- 

# 2 Main Menu

## 2.1 Overview

The main menu is located in the upper left corner of the main window and is accessed via the **Menu** button. It contains basic program functions and settings.

---

## 2.2 Options

Opens the **Program Options** window where you can configure basic settings for the program.

**Keyboard Shortcut:** None

**More Information:** See Chapter 31 (Program Options)

---

## 2.3 Enter License Key

Opens the window for entering the license key. After purchase, you will receive a license key that you can enter here to unlock the full version.

**Demo Version**

Without a valid license key, the program runs in demo mode for 30 days. During this period, all features are available. However, the maximum number of profiles in the demo version is limited to 2 (unlimited in the purchased version).

## Licensing

After entering the license key, the license is validated. The license status is displayed in the title bar of the main window.

---

## 2.4 Help

Opens the online help in the default web browser. The online help contains detailed documentation of all program functions.

**Keyboard Shortcut:** F1

---

## 2.5 About

Displays information about the program:

Information	Description
<b>Version</b>	Current program version
<b>Build</b>	Build number
<b>License</b>	Current license status
<b>Support</b>	Contact information

---

## 2.6 More

The **More** submenu contains additional functions for maintenance and diagnostics.

### 2.6.1 Generate Error Report

Creates an error report as a text file on the desktop. This file contains:

- Program version and system information
- Error log from recent days
- Configuration information

**Usage:** Send this file to support when you need help with problems.

### 2.6.2 Clear Error Logs

Deletes all saved error logs. Use this function after resolving issues to clean up the log files.

### 2.6.3 Backup Application Data

Creates a ZIP file with all application data:

- Profiles and settings
- Log database
- Configuration files

**Tip:** Perform regular manual backups, especially before making major changes to profiles.

### 2.6.4 Backup Application Data (for new PC)

Creates a backup specifically for transfer to another computer. The backup contains all necessary data to set up the program on a new PC with identical settings.

### 2.6.5 Restore Application Data

Restores application data from a previously created backup. After restoration, the program will restart.

**Warning:** All current settings will be replaced by the data from the backup.

### 2.6.6 Open Application Data Folder

Opens the application data folder in Windows Explorer:

%APPDATA%\Gillmeister Software\Automatic PDF Processor 2\

#### Folder Contents:

File/Folder	Description
Profiles.json	Profiles and settings
Options.json	Program options
PDFInfos\	Processing logs
tessdata\	OCR language files

### 2.6.7 Delete Log Database

Deletes the entire log database with all processing logs. This function is helpful when the database has become very large and is slowing down the program.

**Warning:** All log data will be permanently deleted.

### 2.6.8 Processing Time Diagnostics

Opens the **Processing Time Diagnostics** window, which provides an overview of program activity and configured processing times.

**More Information:** See Chapter 39 (Processing Time Diagnostics)

---

## 2.7 Exit

Completely exits the program. All running processes are completed before the program exits.

**Note:** If you close the window using the X button, the program is minimized to the notification area (System Tray) and continues running in the background. To completely exit the program, use **Menu > Exit** or the context menu of the tray icon.

---

## 2.8 Additional Buttons in the Menu Bar

In addition to the main menu, there are additional buttons in the menu bar:

### 2.8.1 Step-by-Step Instructions

Opens the website with video tutorials and instructions in the web browser. The tutorials explain typical use cases step by step.

**Tip:** This button can be hidden in the program options under **Other**.

### 2.8.2 Order Now

This button only appears in demo mode and opens the order page in the web browser.

### 2.8.3 Update Available

This button only appears when a newer program version is available. Clicking it opens the download page. The new version can simply be installed over the existing installation without losing settings.

---

# 3 Profile Toolbar

## 3.1 Overview

The Profile Toolbar is located above the Profile List and contains the most important functions for profile management. The toolbar is divided into two areas: profile actions on the left, filter options on the right.

---

## 3.2 New Profile

Creates a new profile and opens the **Profile Settings** window.

**More Information:** See Chapter 8 (Profile Settings)

## Procedure

1. Click **New Profile...**
2. The Profile Settings window opens with default values
3. Configure the profile according to your requirements
4. Click **OK** to save the profile

**Tip:** New profiles are automatically activated and immediately begin folder monitoring. However, only files that are placed in the monitored folder after creating or modifying a profile are automatically processed.

---

## 3.3 Edit

Opens the **Profile Settings** window for the currently selected profile.

**Prerequisite:** A profile must be selected in the Profile List.

**Keyboard Shortcut:** Double-click on a profile in the list or ENTER

---

## 3.4 Duplicate

Creates a copy of the selected profile. The new profile receives the original's name with the suffix "(Copy)".

### Options (via the dropdown arrow)

Option	Description
<b>and select new profile</b>	The duplicated profile is selected in the list
<b>and open new profile</b>	The duplicated profile is immediately opened for editing

### Use Case

Duplicating is useful when you need a similar profile with minor variations. Example: You have a profile for invoices from Supplier A and want to create a similar one for Supplier B.

---

## 3.5 Activate

Activates the selected profile. Activated profiles:

- Are marked with a checkmark in the **Status** column
- Actively monitor their configured folders
- Automatically process incoming PDF files

**Multiple Selection:** You can activate multiple profiles simultaneously by selecting them while holding the Ctrl key.

---

## 3.6 Catch Up

Opens the **Catch-Up Processing** window for the selected profile. This allows you to:

- Process files already present in a monitored folder
- Process files from another folder with the selected profile
- Reprocess files after a profile change

**More Information:** See Chapter 38 (Catch-Up Processing)

### Typical Use Cases

Scenario	Description
<b>Initial Setup</b>	Process existing files after profile creation
<b>Profile Correction</b>	Reprocess files after changes
<b>Migration</b>	Process files from another folder

---

## 3.7 Deactivate

Deactivates the selected profile. Deactivated profiles:

- Are shown without a checkmark in the **Status** column
- Do not monitor any folders
- Do not process any files

**Multiple Selection:** You can deactivate multiple profiles simultaneously.

**Tip:** Temporarily deactivate profiles when you want to make changes without files being processed during editing.

---

## 3.8 Remove

Deletes the selected profile after a confirmation prompt.

**Warning:** Deleting a profile cannot be undone. Before deleting important profiles, create a backup via **Menu > More > Backup Application Data**.

**Multiple Selection:** You can delete multiple profiles simultaneously.

---

## 3.9 More

The **More** dropdown menu contains additional functions for profile management.

### 3.9.1 Apply All Profiles Subsequently

Performs catch-up processing for all active profiles. Useful when you want to process all monitored folders at once.

### 3.9.2 Select All Profiles

Selects all profiles in the list. Corresponds to the keyboard shortcut **Ctrl+A**.

### 3.9.3 Activate All Profiles

Activates all profiles with one click.

### 3.9.4 Deactivate All Profiles

Deactivates all profiles with one click. Useful for maintenance work or when you want to temporarily stop processing.

### 3.9.5 Remove All Profiles

Deletes all profiles after a confirmation prompt.

**Warning:** This action cannot be undone.

### 3.9.6 Import Profiles

Imports profiles from a previously exported JSON file.

**Conflict Handling:** If an imported profile has the same name as an existing one:

Option	Description
<b>Overwrite</b>	The existing profile is replaced
<b>Add as Copy</b>	The imported profile receives a new name
<b>Skip</b>	The imported profile is not imported

### 3.9.7 Export Profiles

Exports profiles to a JSON file. The submenu offers three options:

Option	Description
<b>All...</b>	Exports all profiles
<b>Active...</b>	Exports only activated profiles
<b>Selected...</b>	Exports only the currently selected profiles

**Tip:** Use export to transfer profiles between different computers or to create backup copies of individual profiles.

---

## 3.10 Filter Area

In the right area of the toolbar, you will find an option to filter the Profile List.

### 3.10.1 Search Field

Enter a search term to filter the Profile List. The search searches:

- Profile name
- Comment
- Monitored folders

**Tip:** The search begins immediately while typing (live search).

### 3.10.2 Show Exclusively

When enabled, only profiles matching the search term are displayed. Non-matching profiles are hidden.

### 3.10.3 Highlight

When enabled, profiles matching the search term are highlighted in the list, but all profiles remain visible.

---

## 3.11 Keyboard Shortcuts

Action	Keyboard Shortcut
Edit profile	Double-click or Enter
Select multiple profiles	Ctrl + Click
Select range	Shift + Click
Select all	Ctrl + A

---

## 4 Profile List

### 4.1 Overview

The Profile List displays all created profiles in a clear table view. Here you can see the status of all profiles at a glance and manage them.

---

## 4.2 Columns

The Profile List contains the following columns:

### 4.2.1 Status

Shows whether the profile is active:

Symbol	Meaning
<input checked="" type="checkbox"/> (Checkmark)	Profile is activated and monitoring folders
<input type="checkbox"/> (Empty)	Profile is deactivated

**Tip:** Click on the checkmark to quickly activate or deactivate a profile.

### 4.2.2 Name

The name of the profile. Names can be changed in the Profile Settings.

**Sorting:** Click on the column header to sort the list alphabetically.

### 4.2.3 Created/Modified On

Shows when the profile was created or last modified.

### 4.2.4 Last Application (regular)

Date and time of the last regular processing by this profile. “Regular” means that a new file was automatically detected and processed.

**Display:** If no processing has taken place yet, the field remains empty.

### 4.2.5 Last Application (subsequently)

Date and time of the last catch-up processing by this profile. This refers to processing triggered via the **Catch Up** function.

### 4.2.6 Monitored Folders

Shows the configured monitored folders of the profile. When there are multiple folders, they are displayed separated by semicolons.

**Tip:** Move the mouse pointer over the cell to see the full path in the tooltip.

### 4.2.7 Comment

Shows the optional comment of the profile. Comments can be entered in the Profile Settings under **General**.

---

## 4.3 Adjusting Column Width

You can individually adjust the width of each column:

1. Move the mouse pointer to the dividing line between two column headers
2. The cursor changes to a double arrow
3. Drag the dividing line left or right

**Tip:** Double-click on the dividing line to automatically adjust the column to its content.

---

## 4.4 Sorting

Click on a column header to sort the list by that column:

- **First click:** Sort ascending (A-Z, oldest first)
- **Second click:** Sort descending (Z-A, newest first)

A small arrow in the column header shows the current sort direction.

**Note:** Regardless of the sorting set here, profiles are always processed in the order of their names.

---

## 4.5 Context Menu

Right-click on the Profile List to open the context menu:

Menu Item	Description
<b>Select All</b>	Selects all profiles
<b>Select None</b>	Deselects all profiles
<b>Invert Selection</b>	Inverts the current selection
<b>Export Profile...</b>	Exports the selected profile

---

## 4.6 Profile Colors

Profiles can be color-coded. The color is displayed as a colored bar on the left edge of the log rows. The color is set in the Profile Settings under **General**.

### Use Cases for Profile Colors

Color	Example Use
Green	Production profiles
Yellow	Test profiles

Color	Example Use
Red	Profiles requiring special attention
Blue	Archiving profiles

---



---

## 4.7 Multiple Selection

You can select multiple profiles simultaneously:

Method	Description
<b>Ctrl + Click</b>	Adds individual profiles to the selection or removes them
<b>Shift + Click</b>	Selects a range of profiles
<b>Ctrl + A</b>	Selects all profiles

### Actions for Multiple Profiles

The following actions can be applied to multiple selected profiles:

- Activate
- Deactivate
- Remove
- Export

---



---

## 4.8 Tips for Profile Management

### Naming Conventions

Use meaningful names that describe the purpose of the profile:

Good	Bad
“Archive Invoices Supplier ABC”	“Profile 1”
“Order Confirmations to Purchasing”	“New Profile”

### Use Comments

Use the comment field for additional information:

- Contact person for this profile
- Special notes about processing
- Date of last review

### Group Profiles

Although there are no actual groups, you can group profiles through naming conventions:

[Accounting] 1 - Incoming Invoices  
[Accounting] 2 - Outgoing Invoices  
[Sales] 1 - Archive Quotes  
[Sales] 2 - Orders to ERP

---

## 4.9 Central Profile Management

When central profile management is enabled (configured in Program Options), profiles can be centrally managed on a network drive. In this case:

- Profiles are loaded from the central file at startup
- Changes are written to the central file
- Multiple installations can use the same profiles

**Note:** If editing of centrally managed profiles is disabled, the buttons for creating, editing, and deleting are grayed out.

---

## 5 Log Toolbar

### 5.1 Overview

The Log Toolbar is located above the Log List and provides functions for controlling and analyzing processing logs.

---

### 5.2 Filter Log Entries

Opens the **Log Filter** window, which allows you to filter the displayed log entries according to various criteria.

#### Filter Criteria

Criterion	Description
<b>Time Period</b>	Display entries from a specific period (Today, Yesterday, Last X Days, From-To)
<b>Profile</b>	Display only entries from a specific profile
<b>Status</b>	Filter by processing result (Success, Error, No Match, etc.)
<b>File Name</b>	Search by file name (partial terms also work)
<b>Use Case</b>	

Filters are particularly useful when you: - Are looking for errors in a specific profile - Want to review the processing of a specific day - Are searching for specific files in the logs

---

## 5.3 Create Excel Report

Creates a detailed Excel report about the processing history.

### Report Content

The Excel report contains several worksheets:

Worksheet	Content
<b>Overview</b>	Overall statistics (number of processes, success rate)
<b>Successful Processing</b>	All successfully processed files
<b>Errors</b>	Files with processing errors
<b>Not Overwritten</b>	Files that were not overwritten
<b>No Match</b>	Files that did not match any filter
<b>No Text</b>	PDF files without recognizable text
<b>Failed Retries</b>	Files with failed retry attempts

### Report Period

When creating the report, you can choose the time period:

Option	Description
<b>Today</b>	Only entries from today
<b>Yesterday</b>	Only entries from yesterday
<b>Last X Days</b>	Entries from the last X days (default: 7)
<b>From-To</b>	Custom time period
<b>Unlimited</b>	All available entries

### Automatic Sending

The report can also be sent automatically via email. Configure this function in **Program Options** under **Excel Report via Email**.

---

## 5.4 Hide Log Entries

Hides all current log entries from the display. The entries are not deleted, only hidden.

### Difference from “Delete”

Action	Effect
<b>Hide</b>	Entries are removed from the view but remain in the database
<b>Delete</b> (via menu)	Entries are permanently removed from the database

**Tip:** Use “Hide” to keep the log view clear without losing log data.

---

## 5.5 Restore Log Entries

Restores previously hidden log entries. All hidden entries are displayed again in the Log List.

---

## 5.6 Minimize/Maximize Log Area

Minimizes or maximizes the log area in the main window.

State	Description
<b>Maximized</b>	Log List is fully visible
<b>Minimized</b>	Log area is hidden, only the Profile List is visible

**Tip:** Minimize the log area when you are mainly working with profiles and need more space for the Profile List.

### Keyboard Shortcut

The size of the log area can also be continuously adjusted by dragging the separator between the Profile List and Log List.

---

## 5.7 Tips for Log Usage

### Regular Review

Regularly review the Log List to: - Detect errors early - Monitor processing - Identify unexpected results

### Error Analysis

When there are errors in processing:

1. Switch to the **Errors** tab in the Log List
2. Double-click on the entry for details
3. The error message provides hints about the cause
4. Check the Profile Settings

## Report Creation for Audits

Use the Excel report function to: - Document processing statistics - Create evidence for audits - Analyze trends over longer periods

---

# 6 Log List

## 6.1 Overview

The Log List displays all processed files in various categories. The display is divided into tabs representing different processing results.

---

## 6.2 Tabs

### 6.2.1 Successful Processing

Shows all files that were successfully processed. An entry appears here when all configured tasks were executed without errors.

#### Columns:

Column	Description
<b>Date</b>	Time of processing
<b>Profile</b>	Name of the applied profile
<b>PDF Modification Date</b>	Date of last modification in the PDF file
<b>File Creation Date</b>	Date of creation
<b>File Name</b>	Original file name
<b>Source Folder</b>	Original folder path
<b>Attachments</b>	Names of contained attachments
<b>File Size</b>	File size of the PDF file
<b>Task</b>	Name of the respective task
<b>Task Details</b>	Detailed information about the tasks

### 6.2.2 Errors

Shows files where errors occurred during processing.

#### Typical Error Causes:

Error	Possible Cause
File not found	File was moved or deleted

Error	Possible Cause
Access denied	File is locked or permissions are missing
Target folder does not exist	Configured target folder not present
PDF error	Corrupted or password-protected PDF
Email sending failed	Network or configuration problem

### 6.2.3 No Match

Shows files that did not match the filter criteria of any active profile.

#### Typical Reasons:

- No profile monitors the folder
- The file does not match the filter criteria
- All matching profiles are deactivated

**Tip:** Check the filter settings of the profile if files appear here that should actually be processed. Add the files to the profile's example files and check the filter results preview.

### 6.2.4 No Text

Shows PDF files that contain no extractable text. This typically applies to:

- Scanned documents without OCR
- Images saved as PDF
- PDF files with embedded graphics instead of text

**Solution:** Use the OCR task to process these files with text recognition.

### 6.2.5 Locked Files

Shows files that could not be accessed during processing because they were locked by another program.

#### Columns:

Column	Description
<b>Detected Since</b>	Time since the file was detected as locked
<b>File Path</b>	Full path to the locked file

#### Typical Causes:

- File is open in a PDF viewer
- Scanner software has not yet released the file
- Antivirus program is checking the file

**Note:** The program periodically retries accessing locked files. Once the lock is released, processing continues.

### 6.2.6 Retry

Shows files where retry attempts after errors are configured.

#### Additional Columns:

Column	Description
<b>Next Attempt</b>	Time of the next retry attempt
<b>Remaining Attempts</b>	Number of attempts still possible
<b>Initial Error</b>	The original error

**More Information:** See Chapter 36.5 (Retries on Errors)

### 6.2.7 Processing Details

Shows real-time information about current processing. This tab is particularly useful during active processing.

#### Display:

- Currently processed file
- Current processing step
- Processing progress

---

---

## 6.3 Context Menu

Right-click on a log entry to open the context menu:

Menu Item	Description
<b>Open File</b>	Opens the processed file (target file)
<b>Show File in Explorer</b>	Opens the file's folder in Explorer
<b>Reprocess File</b>	Processes the source file again
<b>Add file to profile's example files</b>	Adds the source file to the example files of the associated profile (only visible for errors, no matches, or retries)
<b>Copy Entry to Clipboard</b>	Copies the row information as text

---

---

## 6.4 Color Coding

Log entries are color-coded by text:

Color	Meaning
<b>Green</b>	Successfully processed
<b>Red</b>	Error during processing

Color	Meaning
<b>Orange</b>	No match (no matching profile)
<b>Gray</b>	No text / Locked file
<b>Blue</b>	Retry scheduled

---

## 6.5 Tips for Log Analysis

### Recognizing Error Patterns

If multiple files show the same error, there may be a systematic problem:

- Check the profile settings
- Verify network connections
- Check permissions

### Monitoring Processing Speed

The **Processing Details** tab shows how quickly files are being processed. Slow processing may indicate:

- Overloaded network connections
- Slow target folders (network drives)
- Complex OCR processing
- Large email attachment sending

### Regular Cleanup

Regularly hide old log entries to maintain clarity. The data remains in the database and can be restored if needed.

---

## 7 Status Bar

### 7.1 Overview

The Status Bar is located at the bottom of the main window and displays important status information at a glance.

---

### 7.2 Display Areas

#### 7.2.1 Profiles

Shows the number of active and total profiles.

**Format:** Profiles: X/Y active

Value	Meaning
X	Number of activated profiles
Y	Total number of all profiles

**Example:** Profiles: 5/8 active means that 5 of 8 profiles are activated.

### 7.2.2 Log Filter

Shows the number of log entries restricted by filters.

**Format:** Log Filter: X

Value	Meaning
0	No filter active, all entries are displayed
>0	A filter is active, only X entries are displayed

**Tip:** If you expect to see more entries, check whether a filter is active.

### 7.2.3 Log Entries

Shows the total number of log entries across all categories.

**Format:** Log Entries: X

This number includes all entries across all tabs of the Log List: - Successful Processing - Errors - No Match - No Text - Locked Files - Retries

### 7.2.4 Status Information

Shows dynamic status information about current program activity.

**Possible Displays:**

Status	Meaning
<b>Ready</b>	Program is ready and monitoring folders
<b>Processing...</b>	A file is currently being processed
<b>X files found</b>	Number of newly detected files
<b>Checking folders...</b>	Folders are being checked for new files
<b>Waiting for time window</b>	Processing outside the configured time window

---

## 7.3 Updates

The Status Bar is automatically updated:

- **Profiles:** With every change in profile activation
  - **Log Filter:** When applying or removing filters
  - **Log Entries:** After each processing
  - **Status Information:** In real-time during processing
- 

## 7.4 Tips for Status Bar Usage

### Quick Overview

The Status Bar provides a quick overview without navigating into details:

- **All profiles deactivated?** → Check profiles display
- **No new log entries?** → Check status information
- **Expected entries missing?** → Check log filter

### Monitoring Processing Progress

During processing of many files, you can track progress in the status information without having to open the **Processing Details** tab.

### Identifying Problems

Observation	Possible Problem
Profiles: 0/X active	No profile is activated
Long “Processing...” display	File may be large or OCR is taking time
“Waiting for time window”	Processing restricted by time window

---

## 8 Profile Settings

### 8.1 Overview

The **Profile Settings** window is the central configuration window for profiles. Here you define which folders are monitored, which files should be processed, which data is extracted, and which tasks are executed.

**Opening:** Click **New Profile...** or **Edit** in the Profile Toolbar.

### Window Structure

The window is divided into two areas:

Area	Description
<b>Navigation (left)</b>	List of all categories showing configured elements
<b>Content (right)</b>	Settings for the selected category

## Categories

Category	Description
<b>General</b>	Name, activation status, color, processing times
<b>Monitored Folders</b>	Folders monitored for new PDF files
<b>Filter</b>	Criteria for which PDFs are processed
<b>Example Files</b>	Test PDFs for rule preview
<b>Data Extraction</b>	Rules for extracting data from PDFs
<b>Tasks</b>	Actions applied to matching PDFs
<b>File Grouping</b>	Grouping of related files
<b>Subsequently</b>	Notifications and follow-up operations

## Buttons

Button	Description
<b>OK</b>	Saves changes and closes the window
<b>Cancel</b>	Discards all changes and closes the window
<b>Apply</b>	Saves changes, keeps the window open

---

## 8.2 General

The **General** tab contains basic profile settings.

### 8.2.1 Name

The name of the profile. Use meaningful names that describe the purpose of the profile.

**Examples:** - “Archive Incoming Invoices” - “Order Confirmations to Purchasing” - “Process Scanned Receipts with OCR”

### 8.2.2 Comment

Optional free text field for additional information about the profile. The comment is displayed in the Profile List.

### 8.2.3 Active

Activates or deactivates the profile. Only activated profiles monitor folders and process files.

### 8.2.4 Profile Color

Assigns a color to the profile, which is displayed as a colored bar in the Log List. Useful for visual grouping of profiles.

### 8.2.5 Processing Times

Allows restricting processing to specific time windows.

### Use Profile-Specific Processing Times

When enabled, this profile uses its own processing times instead of global settings.

### Schedule Configuration

Setting	Description
<b>Repeat</b>	Every day, Every X days, On specific weekdays
<b>Weekdays</b>	Active weekdays (Mon-Sun)
<b>Time Type</b>	“At” (point in time) or “From-To” (time range)
<b>Time</b>	Start time and optionally end time

#### Time Type “At” (Point in Time)

Processing occurs once at the specified time. Files detected later wait until the next time window.

**Catch-Up Function:** When enabled, missed processing is caught up if the program starts within the catch-up time.

**Example:** Time 09:00, Catch-up 30 minutes - Program starts at 09:15 → Processing is performed - Program starts at 09:35 → Processing is missed, next time point is used

#### Time Type “From-To” (Time Range)

Processing occurs continuously during the specified time range. New files are processed immediately.

### 8.2.6 Retries on Errors

Configures automatic retry attempts on errors.

Setting	Description
<b>Active</b>	Enable retry function
<b>Maximum Attempts</b>	Number of retry attempts (default: 3)
<b>Time Period</b>	Period over which attempts are distributed

**More Information:** See Chapter 36.5 (Retries on Errors)

### 8.2.7 File Group Wait Mode

Determines how the profile interacts with file groupings from other profiles.

Mode	Description
<b>Wait for all</b>	Waits for all profiles with file grouping (default)
<b>Wait for selected</b>	Waits only for specific file group profiles

Mode	Description
<b>Wait for none</b>	Processes immediately without waiting for file groups

### 8.2.8 Recurring Files

Option	Description
<b>Reprocess recurring files</b>	When enabled, files that reappear at the same location (after moving/deleting) are processed again

### 8.2.9 Delay Before Processing

Option	Default	Description
<b>Delay before processing (seconds)</b>	0	Wait time in seconds before the tasks of this profile are executed. Useful when files come from slow sources (e.g., scanners, network drives) and you need to ensure they are completely written. 0 = no delay.

### 8.2.10 Revalidate Filter Before Each Task

Option	Default	Description
<b>Revalidate filter before each task</b>	No	When enabled, the filter is not only checked once at the start of the profile, but is re-evaluated before each individual task. This is useful when a task modifies the file in a way that could change the filter result for subsequent tasks (e.g., when the “Remove Pages” task removes pages containing text that is used as a filter criterion).

**Note:** For existing profiles (before version 2.0.43), this option is enabled by default to maintain the previous behavior. For new profiles, it is disabled.

## 8.3 Monitored Folders

The **Monitored Folders** tab defines which folders are monitored for new PDF files.

### 8.3.1 Folder List

Shows all configured monitored folders.

#### Buttons:

Button	Description
<b>Add</b>	Adds a new folder
<b>Remove</b>	Removes the selected folder

### 8.3.2 Adding Folders

When adding a folder, you can: - Select local folders - Enter network paths (UNC paths: \\Server\Share\Folder) - Use environment variables (e.g., %USERPROFILE%\Documents)

### 8.3.3 Include Subfolders

When enabled, files in subfolders of the monitored folder are also processed.

### 8.3.4 Ignore Documents in Source Folder

This option prevents infinite loops when OCR, compression, splitting, or merging creates new documents in the same monitored folder.

When enabled, such files are added to the list of known files and ignored in the next scan.

**Recommendation:** This option should remain enabled (default) unless you intentionally want to reprocess the generated files.

## 8.4 Filter

The **Filter** tab defines criteria for which PDF files should be processed.

**Detailed Documentation:** See Chapter 8.4 (Filter) in the separate file.

## 8.5 Example Files

The **Example Files** tab allows adding test PDFs for previewing extraction rules.

### 8.5.1 Purpose

Example files serve to: - Test extraction rules in the rule preview - Verify results before production use - Identify problems with data extraction

### 8.5.2 Adding Files

Button	Description
<b>Add</b>	Selects one or more PDF files
<b>Remove</b>	Removes the selected example file
<b>Remove All</b>	Removes all example files

**Tip:** Add representative examples that cover different variants of your documents to be processed.

### 8.5.3 Preview

The example files are used in the data extraction rules to show keyword locations, etc. You can switch between example files to test the rules with different documents.

---

## 8.6 File Grouping

The **File Grouping** tab enables processing related PDF files as a group.

### 8.6.1 Use Cases

- Process invoice with associated delivery note
- Merge main document with attachments
- Combine multiple pages of a scanned document

### 8.6.2 Primary File Group

The primary file group contains the main files to be processed.

Option	Description
<b>Subgroup can contain multiple files</b>	Allows grouping multiple PDFs with identical extraction value and fulfillment of filter criteria
<b>Process files individually</b>	Each file is processed separately
<b>Subsequent tasks apply to all files in primary subgroup</b>	All files in the group receive the listed tasks

### 8.6.3 Secondary Subgroups

Secondary subgroups define additional files belonging to a main file.

Setting	Description
<b>Existence Type</b>	Must exist / Can exist
<b>Multiple Matches</b>	Process all / Newest / Oldest / Report error
<b>Validations</b>	Optional comparisons of placeholder values between different groups (same customer number in primary and secondary document)

### 8.6.4 Maximum Wait Time

The time in seconds to wait for additional files of a group (default: 20 seconds).

**Tip:** Increase the wait time if files arrive with a delay (e.g., with a slow scanner).

---

## 8.7 Data Extraction

The **Data Extraction** tab shows the configured extraction rules.

### 8.7.1 Rule List

Shows all defined rules with the following information:

Column	Description
<b>Name</b>	Name of the rule
<b>Source</b>	Data source (document text, barcode, etc.)
<b>Data Type</b>	Text, date, number, or query

### 8.7.2 Buttons

Button	Description
<b>Add</b>	Creates a new rule
<b>Edit</b>	Opens the selected rule for editing
<b>Duplicate</b>	Creates a copy of the selected rule
<b>Remove</b>	Deletes the selected rule
<b>Move Up/Down</b>	Changes the order of rules

### 8.7.3 Rule Order

The order is important when rules reference each other (e.g., through placeholders). Rules being referenced must be defined before the referencing rules.

**More Information:** See Chapter 9 (Extraction Rules)

---

## 8.8 Tasks

The **Tasks** tab shows all available task types and enables their configuration.

### 8.8.1 Task List

Tasks are displayed as an expandable list. Each task type can be added multiple times.

### 8.8.2 Task Order

The order of tasks determines the execution order. Tasks are processed from top to bottom.

**Tip:** Place critical tasks (e.g., email sending) at the end so they are only executed when all previous tasks were successful.

**Detailed Documentation:** See Chapters 10-30 for individual task types.

---

## 8.9 Notifications

The **Subsequently** tab contains notification settings among other things.

### 8.9.1 Notification Types

Event	Description
<b>On Success</b>	Notification on successful processing
<b>On Error</b>	Notification on processing errors
<b>On No Match</b>	Notification when no filter matches
<b>On Exhausted Retries</b>	Notification when all retries have failed

### 8.9.2 Email Notification

For each notification type, you can: - Specify email recipients (multiple separated by semicolon) - Configure subject and text of the email - Use placeholders for dynamic content

### 8.9.3 Sound Notification

Enable sound notification to receive an audible alert for certain events.

Setting	Description
<b>Enable Sound</b>	Enables sound playback
<b>WAV File</b>	Path to the sound file to play

---

---

## 8.10 Follow-up Operations

The **Subsequently** tab also contains settings for actions on errors or non-matches.

### 8.10.1 On Errors

Option	Description
<b>Move Files</b>	Moves files with errors to a separate folder
<b>Target Folder</b>	Folder for failed files

### 8.10.2 On No Match

Option	Description
<b>Move Files</b>	Moves files without matching filter to a separate folder
<b>Target Folder</b>	Folder for unassigned files

**Note:** These profile-specific settings take precedence over global settings in the Program Options.

---

---

## 8.4 Filter

### 8.4.1 Overview and Basic Principle

Filters allow you to precisely define which PDF files should be processed by a profile. If no filter criteria are defined, all PDF files in the monitored folders will be processed.

#### How Filters Work

For each new PDF file, the program checks: 1. Does the file meet all defined filter criteria? 2. If yes: The file is processed 3. If no: The file is skipped (status: “No Match” if not processed by any profile due to filters)

#### The Filter Tabs

The filter settings are divided into the following tabs:

Tab	Description
<b>File Properties</b>	Filters by file name, path, size, date
<b>PDF Data</b>	Filters by PDF metadata, content, attachments, and barcodes
<b>Results Preview</b>	Shows filter results for example files
<b>Overlap Check</b>	Checks for conflicts with other profiles

### 8.4.2 Tab: File Properties

Here you filter by properties of the file itself.

#### Path Contains / Does Not Contain

Filters by the storage location of the file.

**Example:** To process only files from the “Invoices” subfolder: - Path contains: Invoices

#### File Name Contains / Does Not Contain

Filters by the name of the PDF file.

**Examples:** - File name contains: Invoice - Processes all files with “Invoice” in the name - File name does not contain: DRAFT - Ignores files with “DRAFT” in the name

#### File Size

Filters by the size of the file. Available comparison operators: - Less than - Greater than - Between

**Units:** Bytes (B), Kilobytes (KB), Megabytes (MB)

**Use Case:** Route very large files (e.g., > 50 MB) to a separate profile for compression.

## Creation Date / Modification Date

Filters by the file date. Available options: - **Between:** Specify two fixed dates - **Older than:** X days/weeks/months/years - **Newer than:** X days/weeks/months/years

**Use Case:** Only process files created in the last 7 days.

---

## 8.4.3 Tab: PDF Data

Here you filter by contents and properties of the PDF file itself.

### PDF Metadata

PDF files can contain metadata set by the creator program.

Filter	Description
<b>Author contains / does not contain</b>	The author of the document
<b>Title contains / does not contain</b>	The document title
<b>Subject contains / does not contain</b>	The subject
<b>Keywords contain / do not contain</b>	Stored keywords
<b>Creator contains / does not contain</b>	The creation program
<b>Producer contains / does not contain</b>	The conversion program

**Tip:** You can view the metadata of a PDF file in the document properties (right-click > Properties in many PDF viewers).

**Use Case:** A scanner stores its name as “Creator”. This allows you to process files from different scanners differently.

### Document Text Contains / Does Not Contain

Searches the entire text of the PDF document.

**Examples:** - Document text contains: Invoice - Only PDFs with the word “Invoice” - Document text contains: Mustermann GmbH - Only PDFs from this sender

**Page Range:** Optionally, you can restrict the search area to specific pages: - All pages (default) - First page only - Last page only - Page range (e.g., “1-3”)

### Page Count

Filters by the number of pages. Comparison operators: - Less than - Greater than - Between

**Use Case:** Process single-page documents differently than multi-page ones.

### Character Count

Filters by the number of characters in the document.

**Use Case:** Route PDFs without text (character count = 0) to OCR processing.

### PDF Attachments

Some PDF files contain embedded attachments.

**Attachment Count:** Filters by the number of embedded attachments. Comparison operators: - Less than - Greater than - Between

**Attachment Name Contains / Does Not Contain:** Filters by the name of embedded attachments.

**Example:** ZUGFeRD invoices often contain an attachment named factur-x.xml: - Attachment name contains: factur-x.xml

**Use Case:** Process PDFs with attachments (e.g., ZUGFeRD invoices) separately.

### Barcode

Filters by barcode contents in the PDF.

Option	Description
<b>PDF must contain barcode</b>	Only PDFs with at least one barcode
<b>PDF must not contain barcode</b>	Only PDFs without barcodes
<b>Barcode value contains</b>	Filters by barcode content
<b>Barcode value does not contain</b>	Excludes certain barcode values

**Supported Barcode Types:** - **1D Codes:** Code128, Code39, EAN-13, EAN-8, UPC-A, ITF, Codabar - **2D Codes:** QR Code, DataMatrix, PDF417, Aztec

**Tip:** Barcode detection uses machine learning for higher accuracy. You can enable or disable this option in Program Options under *Processing*.

---

## 8.4.4 Tab: Results Preview

The results preview shows you how your filter settings affect the example files.

### Prerequisite

Add at least 5 representative PDF files on the **Example Files** tab that correspond to the typical documents to be processed.

### Display

For each example file, the following is shown: - **Yes** - The file meets all filter criteria - **No** - The file does not meet at least one filter criterion (with indication of the unmet criterion)

### How to Use the Preview Effectively

1. Add files that should be processed (expected result: “Yes”)
2. Add files that should NOT be processed (expected result: “No”)
3. Check whether the results match your expectations
4. Adjust filters as needed

**Tip:** The results preview is particularly important for complex filters with AND/OR combinations or regular expressions.

---

## 8.4.5 Tab: Overlap Check

This tab shows potential conflicts with other profiles.

### How It Works

The program checks whether other profiles: - Use the same monitored folders - Have similar or overlapping filter criteria

### Displayed Information

Column	Description
<b>Profile</b>	Name of the possibly overlapping profile
<b>Monitored Folders</b>	Common monitored folders
<b>Filter Overlap</b>	Type of possible overlap

### Why Is This Important?

If multiple profiles could process the same files: - The file may be processed multiple times  
- The processing order could be unpredictable - Conflicts with file operations may occur

**Recommendation:** Ensure your filters are unambiguous or enable “Stop processing after applying” in the first matching profile.

---

## 8.4.6 AND/OR Logic

When entering multiple terms in a filter field, you can combine them with logical operators.

### AND Combination

All terms must be present.

**Syntax:** <AND> or <UND>

**Example:** Invoice<AND>Mustermann - Matches: “Invoice to Mustermann GmbH” - Does not match: “Invoice to Schmidt GmbH”

### OR Combination

At least one of the terms must be present.

**Syntax:** <OR> or <ODER>

**Example:** Invoice<OR>Bill<OR>Receipt - Matches any PDF containing “Invoice”, “Bill”, or “Receipt”

### Combinations

You can combine AND and OR. The rule is: **AND separates expressions, OR is evaluated within segments.**

**Example:** Invoice<AND>2024<OR>Bill<AND>2024

For clarification, think of <AND> as a line break:

```
Invoice
<AND>
2024<OR>Bill
<AND>
2024
```

This means: (Invoice) AND (2024 OR Bill) AND (2024)

**Result:** - Matches if: Text contains “Invoice” AND (“2024” OR “Bill”) AND “2024” - Matches “Invoice dated 15.12.2024” (contains Invoice, 2024, 2024) - Matches “Invoice Bill 2024” (contains Invoice, Bill, 2024) - Does not match “Invoice dated 15.12.2023” (does not contain “2024”)

**Tip:** Always test complex filters with the results preview to ensure the desired result is achieved.

---

## 8.4.7 Regular Expressions

For advanced filtering, regular expressions (Regex) are available.

### Syntax

Enclose the regular expression with: <BeginOfRegex>PATTERN<EndOfRegex>

### Examples

Regex	Description	Matches
<BeginOfRegex>INV - \d{5}<EndOfRegex>	Invoice number with 5 digits	INV-12345, INV-00001
<BeginOfRegex>^In voice<EndOfRegex>	Starts with “Invoice”	“Invoice No. 123”
<BeginOfRegex>\d{ 2}\.\d{2}\.\d{4}<	Date in format DD.MM.YYYY	15.12.2024

Regex	Description	Matches
EndOfRegex>		
<BeginOfRegex>€\s *\d+[,\.\.]\d{2}<EndOfRegex>	Euro amount	€ 123.45 or €99.00

### Commonly Used Regex Elements

Element	Meaning
\d	A digit (0-9)
\d{5}	Exactly 5 digits
\d+	One or more digits
\s	A whitespace
\s*	Any number of whitespaces (including none)
^	Beginning of line/text
\$	End of line/text
.	Any character
.*	Any number of any characters
[A-Z]	An uppercase letter
[a-zA-Z]	A letter (upper or lowercase)

**Tip:** Test your regular expressions on websites like [regex101.com](https://regex101.com) before using them in the filter.

## 8.4.8 Number Ranges

With the syntax <NumberRange{MIN,MAX}> you can filter by number ranges.

### Syntax

<NumberRange{Minimum,Maximum}>

### Examples

Filter	Description
<NumberRange{1,99}>	Numbers from 1 to 99
<NumberRange{2020,2025}>	Years from 2020 to 2025
<NumberRange{100,999}>	Three-digit numbers

### Use Case

Filter for documents with specific customer numbers: - Document text contains: Customer number: <NumberRange{5100000,5200000}>

This matches all PDFs containing a customer number between 5100000 and 5200000.

---

## 8.4.9 Dynamic Filter Lists

With dynamic lists, you can make filters flexible without having to change the profile.

### How It Works

1. Create a new list under *Extras > Program Options > Dynamic Lists*
2. Add the desired entries (e.g., client names, project numbers)
3. Use the list in the filter with the syntax: `<EntryFromList{ListName}>`

### Syntax

```
<EntryFromList{Name of List}>
```

### Example

You have a list “Important Customers” with the following entries: - Mustermann GmbH - Schmidt AG - Meyer & Co

**Filter:** Document text contains: `<EntryFromList{Important Customers}>`

The program automatically checks whether any of the list entries appears in the document.

### Advantages

- **Flexibility:** New entries can be added to the list at any time
- **Central Management:** One list can be used in multiple profiles
- **Clarity:** Long lists don’t need to be in the filter itself

**Tip:** Lists can also be imported from external sources such as Excel files or databases.

---

## 8.4.10 Practical Tips

### Build Filters Step by Step

Start with a simple filter and expand it gradually: 1. First filter only by file name 2. Then add text filter 3. Add more complex conditions

### Don’t Filter Too Restrictively

Overly strict filters can cause files to be missed: - Watch for different spellings (Invoice/INVOICE) - Consider typos in source documents - Use OR combinations for variants

### Case Sensitivity

Normal text search does **not** distinguish between uppercase and lowercase - “Invoice”, “INVOICE”, and “invoice” are treated equally.

**Note:** Regular expressions by default distinguish between uppercase and lowercase. Use the flag (?i) for case-insensitive regex search:

<BeginOfRegex>(?*i*)invoice<EndOfRegex> finds “Invoice”, “INVOICE”, “invoice”, etc.

### Multiple Profiles Instead of Complex Filters

Sometimes it’s easier to create multiple profiles with simple filters than one profile with very complex filters.

**Example:** Instead of a complex filter for different document types: - Profile 1: Invoices (Filter: Document text contains “Invoice”) - Profile 2: Delivery Notes (Filter: Document text contains “Delivery Note”) - Profile 3: Orders (Filter: Document text contains “Order”)

---

## 9 Extraction Rules

### 9.1 Overview

Extraction rules enable automatic extraction of information from PDF documents. The extracted data can be used as placeholders in file names, email texts, target folders, and many other contexts.

**Opening:** Click **Add** or **Edit** in the Profile Settings under **Data Extraction**.

#### Typical Applications

Application	Example
File names	<InvoiceDate>_<InvoiceNumber>.pdf
Target folders	D:\Archive\<Year>\<Month>\
Email subject	Invoice <InvoiceNumber> dated <InvoiceDate>
CSV export	All extracted values in a table

#### Structure of a Rule

Each rule consists of several components:

Component	Description
<b>General</b>	Name, source, data type
<b>Determination</b>	How the value is found
<b>Clean-up</b>	Preprocessing of the raw value
<b>Verification</b>	Checking the found value

Component	Description
<b>Format</b>	Post-formatting of the value

## 9.2 General

The **General** tab contains basic settings of the rule.

### 9.2.1 Name

The name of the rule. This name is used for placeholders.

**Format:** <RuleId:N(RuleName)> The rule ID N is automatically determined and used.

**Tip:** Use meaningful names without special characters, for example CustomerNumber or InvoiceDate.

**Note:** If you create multiple rules with the same name, it is sufficient if one of these rules achieves a valid result. The program automatically uses the first successful result. This is useful for fallback scenarios, e.g., when one rule fails for certain document types.

### 9.2.2 Comment

Optional field for notes about the rule.

### 9.2.3 Data Source

Determines where the data is extracted from:

Source	Description
<b>Document Text</b>	Text of the PDF document
<b>Barcode</b>	Content of a barcode in the PDF
<b>PDF Property</b>	PDF metadata (title, author, etc.)
<b>File Property</b>	File properties (name, path, date)
<b>Custom Text</b>	Fixed or calculated value, for example n/a
<b>Placeholder Value</b>	Reference to another rule located above the current rule
<b>Form Field</b>	Value of a PDF form field
<b>Sequential Number</b>	Automatically incrementing number (e.g., for document IDs)

### 9.2.4 Data Type

The required type of the extracted value:

Data Type	Description
<b>Text</b>	Any text

Data Type	Description
<b>Date</b>	Date values with automatic detection
<b>Number</b>	Numeric values
<b>Query</b>	Conditional value selection
<b>Query (with List)</b>	Value from a static or file-based list

## 9.3 Data Source Document Text or Barcode - Determination: Position

With position-based determination, a resizable selection rectangle marks the desired area on the page.

### 9.3.1 Determine Page

Option	Description
<b>Specify page number</b>	Selection rectangle is always positioned on the specified page number
<b>Find page with keyword</b>	Selection rectangle is always positioned on the page with the specified keyword

### 9.3.2 Mark in PDF Viewer

Mark the desired area directly in the page preview: 1. Click **Change position** and adjust the position and size of the selection rectangle to define the desired area 2. Click **Fix position**

## 9.4 Data Source Document Text or Barcode - Determination: Keyword

With keyword determination, a value is extracted relative to a search term (keyword).

### 9.3.1 Determine Page

Option	Description
<b>No determination necessary</b>	The page is defined by the keyword specified in Define Data Area
<b>Specify page number</b>	The page is defined by a specified page number
<b>Find page with keyword</b>	The page is defined by the keyword specified here

### 9.4.1 Define Data Area

#### 9.4.1.1 Keyword

The text searched for in the document.

**Example:** Invoice number: to find the number to the right of it.

### 9.4.1.2 Search Options

Option	Description
<b>Case sensitive</b>	Respects capitalization
<b>Regular expression</b>	Interpret keyword as regex
<b>On multiple occurrences</b>	A specific occurrence, should normally be the first occurrence

### 9.4.1.3 Data Position (Position Relative to Keyword)

Position	Description
<b>Right</b>	Text to the right of the keyword
<b>Left</b>	Text to the left of the keyword
<b>Above</b>	Text above
<b>Below</b>	Text below
<b>Found location area</b>	The searched keyword (optimal for adjusting the data area by extending it to define the desired area)

### 9.4.2 Extend Data Area

Allows relocating and/or extending the area found via the keyword from which data is extracted:

Setting	Description
<b>To the left</b>	Relocates the left edge of the data area by a positive or negative value
<b>To the right</b>	Relocates the right edge of the data area by a positive or negative value
<b>Upward</b>	Relocates the upper edge of the data area by a positive or negative value
<b>Downward</b>	Relocates the lower edge of the data area by a positive or negative value

### 9.4.3 Adjust Data Area Extension

If a keyword was referenced in the previous data area extension, you can fine-tune here

### 9.4.4 Visualization in PDF Viewer

The PDF viewer displays: - **Red:** The found keyword - **Green:** The data area - **Blue:** The extracted value

---

## 9.5 Data Source Document Text - Determination: Text of Page(s)

With this determination, the entire text of one or more pages is used as a basis.

### 9.5.1 Data Determination (Page Text)

#### 9.5.1.1 Determine Page

Option	Description
<b>No determination necessary</b>	Uses the text of all pages
<b>Specify page number</b>	Uses the text of the page with the specified page number
<b>Find page with keyword</b>	Uses the text of the page with the specified keyword

#### 9.5.1.2 Combination with Clean-up

Data determination using page text often yields a lot of text. Use clean-up to extract the relevant part.

---

## 9.6 Data Types

### 9.6.1 Text

For extraction, verification, and formatting of text

For most cases, the data type Text is the right choice.

### 9.6.2 Date

For extraction and verification of a date

With the data type Date, all dates in the text are automatically evaluated. If you don't specify a keyword, the first found date is used. When using this data type, all date components are separately available when using the placeholder for path or file name. For example, you can use only the four-digit year and the month name.

### 9.6.2 Number

For extraction and verification of a number

### 9.6.1 Simple Query

With queries, a value is determined based on conditions.

Defines conditions and associated return values:

Document text contains: "X<OR>Y<OR>Z", then use as result "Delivery Note", else ""

## 9.6.2 Query (with List)

You can use the data type “Query (with List)” to search for the occurrence of a term and use the associated value as the result, e.g., an email address or a folder name.

**List Format:** Search term and result value are separated by semicolon.

**Example 1:** Assign email addresses based on customer numbers:

```
Customer number : 19006;x@y.de  
Customer number : 1900;a@b.de  
Customer number : 18765;c@d.de
```

If the PDF contains “Customer number : 19006”, “x@y.de” is used as the result.

**Example 2:** Search IBAN, use company name as result:

```
DE02120300000000202051<OR>DE02 1203 0000 0000 2020 51;Mustermann GmbH  
DE02500105170137075030;Musterfrau GmbH
```

Here the IBAN (with or without spaces) is searched and the associated company name is returned.

---

## 9.7 Data Source: Form Field

Extracts values from PDF form fields.

### 9.7.1 Field Selection

Shows all form fields present in the PDF:

Field Type	Description
<b>TextBox</b>	Text input field
<b>CheckBox</b>	Selection field (Yes/No)
<b>RadioButton</b>	Option button
<b>ComboBox</b>	Dropdown list
<b>ListBox</b>	Selection list

### 9.7.2 Form Field (Field Name)

Select the form field by its name. The name is defined in the PDF form settings.

---

## 9.8 Data Source: Sequential Number

Generates an automatically incrementing number saved per profile and rule. Unlike other data sources, no data is extracted from the PDF — instead, a new value is generated.

**Typical applications:** Document IDs, barcode content, invoice numbers, reference numbers.

### 9.8.1 Settings

Setting	Description
<b>Start Value</b>	The first number (e.g., 1 or 1000)
<b>Increment</b>	Increase per document (e.g., 1, 10, 100)
<b>Format</b>	Output format for leading zeros (e.g., 000000 → 000001)
<b>Auto-reset after days</b>	Counter is deleted after X days (0 = never)

### 9.8.2 Reset Value

To reset the current counter to the start value, enable the **Reset current value to start value** checkbox and save the profile.

### 9.8.3 Example

#### Document ID with QR code:

1. Create extraction rule “DocNr” with data source **Sequential Number**
  - o Start value: 1, Format: 000000
2. Barcode task with content: DOC-<RuleId:1(DocNr)>
3. Result: DOC-000001, DOC-000002, DOC-000003, ...

**Note:** The sequential number is also available in the **Replace Content** task as a standalone replacement source. There, start value, format, and increment are configurable directly in the replacement settings.

---

## 9.9 Clean-up

Clean-up enables preprocessing of the extracted raw value.

### 9.9.1 Available Clean-up Tasks

#### Replace Operations

Task	Description
<b>Replace text</b>	Replaces one text with another
<b>Replace text before marker</b>	Replaces everything before a marker
<b>Replace text after marker</b>	Replaces everything after a marker
<b>Replace regex result</b>	Replaces regex matches
<b>Replace line breaks</b>	Replaces line breaks with text
<b>Replace with Excel file</b>	Replaces based on Excel mapping

## Insert Operations

Task	Description
<b>Insert before marker</b>	Inserts text before a marker
<b>Insert after marker</b>	Inserts text after a marker
<b>Insert at position</b>	Inserts text at a specific position

## Remove Operations

Task	Description
<b>Remove text</b>	Removes a specific text
<b>Remove text before marker</b>	Removes everything before a marker
<b>Remove first/last characters</b>	Removes X characters at the beginning/end
<b>Remove regex result</b>	Removes regex matches
<b>Remove blank lines</b>	Removes all blank lines
<b>Remove lines with regex</b>	Removes lines matching a pattern

## Line Operations

Task	Description
<b>Extract line X</b>	Extracts only a specific line
<b>Move line X</b>	Moves a line to another position
<b>Move lines with text</b>	Moves lines containing specific text

### 9.9.2 Clean-up Order

Multiple clean-up tasks are executed in the defined order. Use the arrow buttons to adjust the order.

---

## 9.10 Verification: Text

Text verifications check the extracted value for certain conditions.

### 9.10.1 Available Checks

Check	Description
<b>Text equals</b>	Exact match
<b>Text does not equal</b>	No match
<b>Text contains</b>	Contains the search term
<b>Text does not contain</b>	Does not contain the search term

Check	Description
<b>Text starts with</b>	Starts with the search term
<b>Text ends with</b>	Ends with the search term
<b>Text matches regex</b>	Matches the regular expression
<b>Text does not match regex</b>	Does not match the expression
<b>Extracted text is empty</b>	No value extracted
<b>Number of characters</b>	Checks text length
<b>Number of lines</b>	Checks line count

### 9.10.2 Character Verification

Checks individual characters at specific positions:

Check	Description
<b>Is digit</b>	Character is 0-9
<b>Is letter</b>	Character is A-Z or a-z
<b>Is uppercase</b>	Character is A-Z
<b>Is lowercase</b>	Character is a-z
<b>Is alphanumeric</b>	Character is letter or digit
<b>Matches regex</b>	Character matches a pattern

## 9.11 Verification: Date

Date verifications check whether the extracted value is a valid date.

### 9.11.1 Available Checks

Check	Description
<b>Date is valid</b>	Value is a recognizable date
<b>Date is between</b>	Date is within the specified period

### 9.11.2 Date Formats

The system automatically recognizes various date formats: - 01.12.2024 (German) - 12/01/2024 (American) - 2024-12-01 (ISO) - December 1, 2024 (with month name)

## 9.12 Verification: Number

Number verifications check numeric values.

### 9.12.1 Available Checks

Check	Description
<b>Number is valid</b>	Value is a recognizable number
<b>Number is between</b>	Value is within range

### 9.12.2 Number Formats

Recognized formats: - 1234 (integer) - 1.234,56 (German) - 1,234.56 (English) - -123.45 (negative)

## 9.13 Verification: Query

Query verifications check values based on conditions.

### 9.13.1 Available Checks

Check	Description
<b>Query returns result</b>	The query returns a value

## 9.14 Formatting

Formatting enables post-processing of the verified value. The main difference from clean-up is that here the search word must be present for replacements.

## 9.15 Using Placeholders

Extracted values can be used as placeholders in many contexts.

### 9.15.1 Placeholder Syntax

Syntax	Description
<RuleName>	Simple placeholder
<RuleId:1(RuleName)>	Complete syntax with ID
<RuleName{DatePart}>	Extract date part

### 9.15.2 Date Parts

DatePart	Description	Example
Year4	Four-digit year	2024
Year2	Two-digit year	24
Month	Month (two digits)	12
MonthName	Month name	December
MonthNameAbbreviated	Abbreviated month name	Dec
Day	Day (two digits)	15

**Example:** <InvoiceDate{Year4}>-<InvoiceDate{Month}> yields “2024-12”

### 9.15.3 Fallback Rules

If multiple rules have the same name, the first successful rule is used. This enables fallback values:

1. Rule “Date” - Attempts extraction from document text
  2. Rule “Date” - If failed: Uses file date
- 

## 10 Tasks - Overview

### 10.1 What are Tasks?

Tasks are the actions that Automatic PDF Processor performs on recognized PDF files. Each profile can contain any number of tasks that are executed in an order you define.

---

### 10.1 Task List and Order

#### Available Task Types (24)

No.	Task	Description
1	Rename file	Renames the PDF file according to a defined pattern
2	Copy file	Copies the PDF file to a target folder
3	Move file	Moves the PDF file to a target folder
4	Delete file	Deletes the PDF file (to Recycle Bin)
5	Print	Prints the PDF file on a printer
6	Send email	Sends the PDF file via email
7	Merge PDFs	Combines multiple PDFs into one file
8	Split PDFs	Splits a PDF into multiple files
9	OCR (Text recognition)	Makes scanned PDFs searchable
10	Save attachments	Extracts embedded attachments from the PDF
11	Save extractions	Exports extracted data as CSV
12	Use template	Fills templates with extracted data
13	Change file protection	Sets or removes password protection
14	Compress PDF	Reduces file size
15	Embed attachments	Embeds files as attachments in the PDF

No.	Task	Description
16	Create ZUGFeRD/Factur-X	Creates electronic invoices according to ZUGFeRD standard
17	Convert to PDF/A	Converts to the archiving format PDF/A
18	Remove pages	Removes specific pages from the PDF
19	Stamp/Watermark	Adds stamps or watermarks
20	Insert barcode	Inserts QR codes or other barcodes
21	Replace content	Replaces text content in the PDF
22	Flatten pages	Permanently burns annotations and form fields into the page
23	Auto-rotate pages	Detects text orientation and corrects misrotated pages
24	Redact text	Permanently redacts extracted text areas in the PDF

### Adding and Managing Tasks

In the profile settings, you will find a list of all available task types in the **Tasks** category. To configure a task:

1. Select the desired task type in the list
2. Click **Add** or double-click the entry
3. Configure the task in the dialog that appears
4. Click **OK** to add the task to the profile's task list

### Setting the Order

The order of tasks is important as they are executed sequentially. Use the **Move Up** and **Move Down** buttons to adjust the order.

**Typical Order:** 1. Rename file 2. Copy/Move file 3. Send email (if desired) 4. Delete file (if desired)

**Tip:** Plan the order carefully. A “Delete file” task, for example, should always be at the end, after all other tasks are completed.

### Multiple Tasks of the Same Type

You can add multiple tasks of the same type. This is useful when you want to, for example:

- Copy a file to multiple folders
  - Send an email to different recipients with different content
  - Place different stamps on different pages
-

## 10.2 Common Concepts

Many tasks share common settings areas. These are explained here once and apply to all relevant tasks.

### General Settings

Each task has a **General Settings** area with the following options:

- **Enabled:** Turns the task on or off. Disabled tasks are skipped.
- **Name:** An optional name for better identification of the task in the list.
- **Comment:** Space for your notes about the task.

### Storage Location

Many tasks (Rename, Copy, Move, OCR, etc.) require specifying a storage location. This consists of:

#### Directory

The base directory where the file should be stored. You can: - Enter a fixed path (e.g., D:\Archive\Invoices) - Select the folder via **Browse...** - Use placeholders (e.g., D:\Archive\

#### Subfolder

Optionally, you can define subfolders that are created dynamically. Use placeholders to create a folder structure based on extracted data:

- <Year4>\<Month> creates e.g., 2024\12
- <RuleId:1(Client)> creates a folder with the extracted client name

#### Filename

The name of the target file. You can: - Enter a fixed name - Keep the original name with <FileName> - Compose a dynamic name from placeholders

**Example:** <RuleId:1(Date)>\_<RuleId:2(InvoiceNo)>\_<FileName> Creates: 2024-12-15\_INV-12345\_Invoice.pdf

#### Name Collisions

For cases where a file with the defined name already exists, the following options are available:

Option	Description
<b>Overwrite</b>	The existing file is replaced
<b>Append number</b>	Adds a number (e.g., File (1).pdf, File (2).pdf)
<b>Append date</b>	Adds date and time

Option	Description
<b>Cancel operation</b>	The task is not executed, but processing still counts as completed, not as an error

## Placeholders

Placeholders are a central concept in Automatic PDF Processor. They enable dynamic filenames and folder paths.

### File Properties

Placeholder	Description	Example
<FileName>	Filename (without extension)	Invoice_Company
<FileNamePartL:N>	Left part of filename (N characters)	with N=8: Invoice_
<FileNamePartR:N>	Right part of filename (N characters)	with N=7: Company
<ParentFolderNameOfLevel:N>	Name of parent folder (level N)	Inbox

### Date Values (File System)

Placeholder	Description	Example
<FileCreatedYear4>	Creation year (4-digit)	2024
<FileCreatedYear2>	Creation year (2-digit)	24
<FileCreatedMonth>	Creation month (2-digit)	12
<FileCreatedMonthName>	Creation month (name)	December
<FileCreatedDay>	Creation day	15
<FileModified...>	Modification date (same variants)	-

### Current Date

Placeholder	Description	Example
<TodaysYear4>	Current year (4-digit)	2024
<TodaysMonth>	Current month	12
<TodaysDay>	Current day	15

### PDF Properties

Placeholder	Description
<Author>	PDF author
<Title>	PDF title
<Subject>	PDF subject
<Keywords>	PDF keywords

Placeholder	Description
<PageCount>	Page count
<PDFCreated...>	PDF creation date (same variants as file system)
<PDFModified...>	PDF modification date

### Prior Variants (Previous Month/Day/Year)

Placeholder	Description
<FileCreatedPriorYear4>	Previous year of creation date
<FileCreatedPriorMonth>	Previous month of creation date
<FileCreatedPriorDay>	Previous day of creation date
<TodaysPriorYear4>	Previous year
<TodaysPriorMonth>	Previous month
<TodaysPriorDay>	Previous day

These are particularly useful for monthly statements that arrive in the following month and should be labeled with the previous month's name.

### Extracted Data

Values from your extraction rules can be used as placeholders:

<RuleId:N(RuleName)> where N is the automatically determined ID of the rule.

**Example:** <RuleId:1(InvoiceNumber)> inserts the invoice number extracted with rule 1.

**Tip:** Click the **Placeholders** link above an input field to see and select all available placeholders in the "Extracted Data" menu.

### File Date

Many tasks allow you to adjust the file date of the target file. You can choose between:

Option	Description
<b>Do not change</b>	The file date remains unchanged
<b>Creation date of original file</b>	Uses the original creation date
<b>Modification date of original file</b>	Uses the modification date
<b>PDF creation date</b>	Date from the PDF metadata
<b>Extracted date</b>	A date extracted with a rule
<b>Current date</b>	Sets today's date

**Use Case:** When archiving invoices by invoice date, you can set the file date to the extracted invoice date. This way, Windows Explorer displays the correct date.

### Afterwards (External Program)

After completing a task, an external program can be called automatically. This enables integration with other applications.

**Configuration:** - **Program:** Path to the executable file - **Parameters:** Command line parameters to be passed

**Available Placeholders for Parameters:** - <DestinationPath> - Full path of the processed file - <DestinationFolder> - Folder of the processed file - <DestinationFileName> - Name of the processed file

**Example:** Opening the processed file in a PDF viewer: - Program: C:\Program Files\Adobe\Acrobat\Acrobat.exe - Parameters: "<DestinationPath>"

---

## 10.2 Tips for Task Configuration

### Test with Example Files

Before putting a profile into production: 1. Add representative PDFs on the **Example Files** tab 2. Use the **Result Preview** to check the outcome 3. Test the profile with a single file

### Follow Logical Order

- Rename before Copy if the new name is used in the target path
- Delete always at the end

**Note about OCR:** The OCR task creates a new file in the configured target folder. This new file must be processed further with a separate profile that monitors the OCR output folder if needed.

### Plan Error Handling

Consider what should happen when a task fails: - Under *Subsequent Operations*, you can move failed files to a special folder - The *Notifications* can inform you of errors via email

### Combine Tasks Effectively

**Scenario: Archiving with Backup** 1. Rename file (with placeholders) 2. Copy file (to archive folder) 3. Copy file (to backup folder on network drive) 4. Delete original file

### Scenario: OCR Workflow with Two Profiles

*Profile 1 - Perform OCR:* 1. Perform OCR (output to C:\OCR-Output) 2. Delete original file (optional)

*Profile 2 - Process OCR Result:* - Monitors: C:\OCR-Output 1. Rename file 2. Move file (to archive)

---

# 11 Rename File

Task: Rename File

*Task: Rename File*

## 11.1 Description

The **Rename File** task changes the name of a PDF file according to a pattern you define. The original storage location remains unchanged - the file is only renamed, not moved.

### Typical Use Cases

- **Invoice Archiving:** Scan001.pdf becomes 2024-12-15\_Invoice\_INV-12345.pdf
  - **Document Management:** Consistent naming convention for all documents
  - **Sorting:** Filenames begin with date for chronological sorting
- 

## 11.2 General Settings

### Enabled

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

## 11.3 Filename

### New Filename

Enter the desired naming pattern here. You can: - Enter fixed text - Use placeholders for dynamic values - Combine both

### Examples:

Input	Result
Invoice	Invoice.pdf
<Year4>-<Month>-<Day>_Document	2024-12-15_Document.pdf
<RuleId:1(Date)>_<RuleId:2(InvoiceNo)> <FileName>_archived	2024-12-15_INV-12345.pdf Scan001_archived.pdf

### Using Placeholders

Click the **Placeholders** link above the input field to open a menu with all available placeholders.

### Commonly Used Placeholders:

Placeholder	Description	Example Value
<FileName>	Filename (without extension)	Scan001
<Year4>	Current year (4-digit)	2024
<Month>	Current month	12
<Day>	Current day	15
<FileCreatedYear4>	File creation year	2024
<RuleId:N(RuleName)>	Extracted value from rule N	(variable)

**Tip:** For archiving, use the extracted document date instead of the current date, so invoices from December are still named correctly in January.

### Invalid Characters

Windows does not allow certain characters in filenames: \ / : \* ? " < > |

If extracted values contain such characters, they are automatically replaced with underscores.

## 11.4 Name Collisions

### Behavior When File Exists

Choose what should happen if a file with the new name already exists:

Option	Description
<b>Overwrite</b>	The existing file is replaced. <b>Caution:</b> Data may be lost!
<b>Append number</b>	Adds a number: Invoice.pdf, Invoice (1).pdf, Invoice (2).pdf
<b>Append date</b>	Adds date and time: Invoice_2024-12-15_14-30-25.pdf
<b>Cancel operation</b>	The rename is not performed, no error is logged

**Recommendation:** For most use cases, **Append number** is the safest option.

## 11.5 File Date

### Adjust Creation and Modification Date

Optionally, you can change the file date of the renamed file:

Option	Description
<b>Do not change</b>	File dates remain unchanged

Option	Description
<b>Creation date of original file</b>	Uses the original creation date
<b>Modification date of original file</b>	Uses the modification date
<b>PDF creation date</b>	Date from PDF metadata
<b>Extracted date</b>	A date obtained with an extraction rule
<b>Current date</b>	Sets today's date and current time

**Use Case:** When archiving invoices by invoice date, it makes sense to set the file date to the extracted invoice date. This way, Windows Explorer displays the correct date, and files can be sorted chronologically.

---

## 11.6 Afterwards

### Call External Program

After renaming, an external program can be started automatically.

**Program:** Path to the executable file (e.g., C:\Program Files\...\program.exe)

**Parameters:** Command line parameters. Available placeholders: -

<PathIncludingFilename> - Full path of the renamed file - <ParentDirectory> - Path of the parent folder - <Filename> - New filename

**Example:** Open renamed file in default PDF viewer: - Program: cmd.exe - Parameters: /c start "" "<PathIncludingFilename>"

---

## 11.7 Example: Name Invoices by Date and Number

### Initial Situation

You receive scanned invoices with meaningless names like Scan001.pdf, Scan002.pdf. These should be automatically renamed according to the pattern YYYY-MM-DD\_InvoiceNumber\_Supplier.pdf.

### Prerequisites

You have created extraction rules for: - Rule 1: "InvoiceDate" (extracts date in format DD.MM.YYYY) - Rule 2: "InvoiceNumber" (extracts invoice number) - Rule 3: "Supplier" (extracts company name)

### Configuration

1. **Enabled:** Yes

2. **Filename:** <RuleId:1(InvoiceDate){Year4}>-<RuleId:1(InvoiceDate){Month}>-<RuleId:1(InvoiceDate){Day}>\_<RuleId:2(InvoiceNumber)>\_<RuleId:3(Supplier)>
3. **On name collision:** Append number
4. **File date:** Extracted date (Rule 1: InvoiceDate)

## Result

Original File	New Name
Scan001.pdf	2024-12-15_INV-12345_Sample Company Inc.pdf
Scan002.pdf	2024-12-10_INV-12340_Smith Corp.pdf

## 11.5 Tips and Notes

### Make Filenames Unique

Use multiple placeholders to create unique filenames. A name consisting only of the date quickly leads to collisions.

**Good:** <RuleId:1(Date)>\_<RuleId:2(InvoiceNo)>\_<RuleId:3(Customer)> **Problematic:** <RuleId:1(Date)> (multiple invoices on the same day)

### Order with Other Tasks

The “Rename File” task should be placed **before** “Move File” or “Copy File” if the new name should be used in the target path.

### Keep Original Name

If you want to keep the original name as part of the new name, use <FileName>:

<Year4>-<Month>-<Day>\_<FileName> creates 2024-12-15\_Scan001.pdf

### Filename Length

Windows limits path length to 260 characters. Long paths combined with long filenames can cause problems. Therefore, keep filenames as short and concise as possible.

### Special Characters in Extracted Values

Extracted values may contain special characters that are not allowed in filenames. These are automatically replaced with underscores. If this leads to unsightly results, clean up the values in the extraction rules (Cleanup tab).

## 12 Copy File

Task: Copy File

Task: Copy File

## 12.1 Description

The **Copy File** task creates a copy of the PDF file at a destination you specify. The original file remains unchanged at its original location.

### Typical Use Cases

- **Archiving:** Copy invoices to a structured archive directory
  - **Backup:** Create copy to a network drive or backup folder
  - **Distribution:** Copy documents to multiple department folders
  - **Workflow:** Pass documents to a processing folder for other applications
- 

## 12.2 General Settings

### Enabled

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

## 12.3 Storage Location

### Directory

Specify the target directory for the copy. You can: - Enter a fixed path (e.g., D:\Archive\Invoices) - Select the folder via **Browse...** - Use placeholders for dynamic folder paths

### Examples with Placeholders:

Input	Result
D:\Archive\<TodaysYear4>\<TodaysMonth>	D:\Archive\2024\12
D:\Customers\<RuleId:1(Customer)>\Server\Backup\<FileCreatedYear4>	D:\Customers\Sample Company Inc \Server\Backup\2024

**Note:** If you don't specify a directory, the file is saved in the same folder as the original. However, it's recommended to use a separate folder for each processing step to ensure clear separation.

### Filename

Set the name for the target file. You can: - Leave the field empty (original name is used) - Enter a fixed name - Use placeholders for dynamic names

## Examples:

Input	Result
(empty)	Scan001.pdf (original name)
<RuleId:1(Date)>_<RuleId:2(InvoiceNo)>	2024-12-15_INV-12345.pdf
Archive_<FileName>	Archive_Scan001.pdf

## Name Collisions

Choose what should happen if a file with the target name already exists:

Option	Description
<b>Overwrite</b>	The existing file is replaced. <b>Caution:</b> Data may be lost!
<b>Append number</b>	Adds a number: File.pdf, File(1).pdf, File(2).pdf
<b>Append date</b>	Adds processing date: File(2024-12-15).pdf
<b>Append date and time</b>	Adds date and time: File(2024-12-15_14-30-25).pdf
<b>Cancel operation</b>	The copy is not created, processing is considered completed

**Recommendation:** For most use cases, **Append number** is the safest option.

---

## 12.4 Copy Additional Files

### Copy Remaining Files in Folder

Enable this option to copy all other files from the source folder to the target folder in addition to the PDF file.

**Use Case:** You receive documents together with accompanying files (e.g., a PDF invoice with associated image files). All files should be copied together to the archive.

### Wait Time

When you enable this option, you can set a wait time in seconds. The program waits the specified time before copying the additional files. This is useful when accompanying files arrive in the source folder with a delay.

---

## 12.5 File Date

### Adjust Creation and Modification Date

Optionally, you can change the file date of the copied file:

Option	Description
<b>Do not change</b>	The copy automatically receives the current date
<b>Creation date of original file</b>	Uses the original creation date
<b>Modification date of original file</b>	Uses the modification date
<b>PDF creation date</b>	Date from PDF metadata
<b>Extracted date</b>	A date obtained with an extraction rule
<b>Current date</b>	Sets today's date and current time

**Use Case:** When archiving by invoice date, you can set the file date to the extracted invoice date. This way, Windows Explorer displays the correct date.

## 12.6 Afterwards

### Call External Program

After copying, an external program can be started automatically.

**Program:** Path to the executable file (e.g., C:\Program Files\...\program.exe)

**Parameters:** Command line parameters. Available placeholders: -

<PathIncludingFilename> - Full path of the copied file - <ParentDirectory> - Path of the parent folder - <Filename> - Filename of the copy

**Example:** Automatically open copied file: - Program: cmd.exe - Parameters: /c start "" "<PathIncludingFilename>"

## 12.7 Example: Invoice Archive with Year/Month Structure

### Initial Situation

Incoming invoices should be automatically copied to a structured folder hierarchy: Archive\Year\Month\Supplier\.

### Prerequisites

You have created extraction rules for: - Rule 1: "InvoiceDate" (extracts date) - Rule 2: "Supplier" (extracts company name)

### Configuration

1. **Enabled:** Yes
2. **Directory:**  
D:\Archive\<RuleId:1(InvoiceDate){Year4}>\<RuleId:1(InvoiceDate){Month}>\<RuleId:2(Supplier)>
3. **Filename:** (empty - keep original name)

4. **On name collision:** Append number
5. **File date:** Extracted date (Rule 1: InvoiceDate)

## Result

Original File	Destination
C:\Inbox\Scan001.pdf	D:\Archive\2024\12\Sample Company Inc\Scan001.pdf
C:\Inbox\Scan002.pdf	D:\Archive\2024\12\Smith Corp\Scan002.pdf

## 12.8 Creating Multiple Copies

You can add multiple “Copy File” tasks to the same profile to copy a file to different destinations.

### Example: Archive and Backup

1. **Task 1:** Copy to local archive (D:\Archive\...)
2. **Task 2:** Copy to backup server (\\Server\Backup\...)
3. **Task 3:** Copy to department folder (\\Server\Department\...)

All three copies are created sequentially.

## 12.6 Tips and Notes

### Using Network Paths

For network paths, use UNC notation: - \\ServerName\Share\Folder

Make sure the program or user running the program has write permissions on the network drive.

### Combination with Other Tasks

The “Copy File” task combines well with other tasks:

Scenario	Task Order
Archive and Backup	1. Rename, 2. Copy (Archive), 3. Copy (Backup)
Process and Keep Original	1. Copy (Backup), 2. OCR, 3. Move
Distribute to Departments	1. Copy (Accounting), 2. Copy (Management), 3. Delete Original

### Keep Original

Unlike the “Move File” task, “Copy File” preserves the original file. Use “Copy” when: - You need a backup of the original file - The file should be distributed to multiple locations - The original file must remain available for further processing

### **Avoid Endless Loops**

If the target folder for the copy is also a monitored folder, this can lead to an endless loop. In this case, enable the option “Ignore documents created by copying to source folder” in the monitored folder settings.

---

## 13 Move File

Task: Move File

*Task: Move File*

### 13.1 Description

The **Move File** task moves the PDF file from its original location to a destination you specify. Unlike copying, no file remains at the original location.

#### **Typical Use Cases**

- **Archiving:** Move processed invoices to an archive directory
  - **Sorting:** Sort documents into different folders by type or date
  - **Cleanup:** Automatically empty inbox folders after processing
  - **Workflow:** Pass documents to the next processing step
- 

### 13.2 General Settings

#### **Enabled**

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

### 13.3 Storage Location

#### **Directory**

Specify the target directory. You can: - Enter a fixed path (e.g., D:\Archive\Invoices) - Select the folder via **Browse...** - Use placeholders for dynamic folder paths

#### **Examples with Placeholders:**

Input

Result

---

Input	Result
D:\Archive\<<TodaysYear4>\<TodaysMonth>	D:\Archive\2024\12
D:\Customers\<<RuleId:1(Customer)>\Server\Archive\<<FileCreatedYear4>	D:\Customers\Sample Company Inc \Server\Archive\2024

**Note:** It's recommended to use a separate folder for each processing step to ensure clear separation.

### Filename

Set the name for the target file. You can: - Leave the field empty (original name is used) - Enter a fixed name - Use placeholders for dynamic names

### Examples:

Input	Result
(empty)	Scan001.pdf (original name)
<RuleId:1(Date)>_<RuleId:2(InvoiceNo)>	2024-12-15_INV-12345.pdf
Archive_<FileName>	Archive_Scan001.pdf

### Name Collisions

Choose what should happen if a file with the target name already exists:

Option	Description
<b>Overwrite</b>	The existing file is replaced. <b>Caution:</b> Data may be lost!
<b>Append number</b>	Adds a number: File.pdf, File(1).pdf, File(2).pdf
<b>Append date</b>	Adds processing date: File(2024-12-15).pdf
<b>Append date and time</b>	Adds date and time: File(2024-12-15_14-30-25).pdf
<b>Cancel operation</b>	The file is not moved, processing is considered completed

**Recommendation:** For most use cases, **Append number** is the safest option.

## 13.4 Move Additional Files

### Move Remaining Files in Folder

Enable this option to move all other files from the source folder to the target folder in addition to the PDF file.

**Use Case:** You receive documents together with accompanying files (e.g., a PDF invoice with associated image files). All files should be moved together to the archive.

## Wait Time

When you enable this option, you can set a wait time in seconds. The program waits the specified time before moving the additional files. This is useful when accompanying files arrive in the source folder with a delay.

## Delete Parent Folder If Empty

Enable this option to automatically delete the parent folder when it's empty after moving.

**Use Case:** PDFs are placed in subfolders (e.g., Inbox\2024-12-15\Invoice.pdf). After moving the invoice, the then-empty folder 2024-12-15 should be automatically deleted.

---

## 13.5 File Date

### Adjust Creation and Modification Date

Optionally, you can change the file date of the moved file:

Option	Description
<b>Do not change</b>	File date remains unchanged
<b>Creation date of original file</b>	Uses the original creation date
<b>Modification date of original file</b>	Uses the modification date
<b>PDF creation date</b>	Date from PDF metadata
<b>Extracted date</b>	A date obtained with an extraction rule
<b>Current date</b>	Sets today's date and current time

**Use Case:** When archiving by invoice date, you can set the file date to the extracted invoice date. This way, Windows Explorer displays the correct date.

---

## 13.6 Afterwards

### Call External Program

After moving, an external program can be started automatically.

**Program:** Path to the executable file (e.g., C:\Program Files\...\program.exe)

**Parameters:** Command line parameters. Available placeholders: - <PathIncludingFilename> - Full path of the moved file - <ParentDirectory> - Path of the parent folder - <Filename> - Filename at new location

**Example:** Automatically open moved file: - Program: cmd.exe - Parameters: /c start "" "<PathIncludingFilename>"

---

## 13.7 Example: Sort Invoices by Supplier

### Initial Situation

Incoming invoices from the scanner folder should be automatically moved to supplier-specific folders.

### Prerequisites

You have created an extraction rule for: - Rule 1: "Supplier" (extracts company name from invoice)

### Configuration

1. **Enabled:** Yes
2. **Directory:** D:\Archive\Invoices\<<RuleId:1(Supplier)>
3. **Filename:** (empty - keep original name)
4. **On name collision:** Append number
5. **Delete parent folder:** No

### Result

Original File	Destination
C:\Scanner\Scan001.pdf (Sample Company Inc)	D:\Archive\Invoices\Sample Company Inc\Scan001.pdf
C:\Scanner\Scan002.pdf (Smith Corp)	D:\Archive\Invoices\Smith Corp\Scan002.pdf

The scanner folder is automatically emptied.

---

## 13.8 Difference from Copy

Aspect	Move	Copy
Original is preserved	No	Yes
Inbox folder is emptied	Yes	No
Multiple destinations possible	No (only one destination)	Yes (multiple tasks)
Suitable for archiving	Yes	Yes (with backup function)
Suitable for distribution	No	Yes

**When to use Move:** - The inbox folder should be automatically cleaned up - The file should only exist in one location - The original file is no longer needed after processing

**When to use Copy:** - The original file should be preserved - The file should be distributed to multiple locations - A backup should be created before further processing

---

## 13.6 Tips and Notes

### Consider Order

Place the “Move File” task typically at the end of the task list. After moving, the file is no longer available at the original location.

**Recommended Order:** 1. Rename file (if desired) 2. Copy file (for backup, if desired) 3. Send email (if desired) 4. Move file (as last task)

### Network Paths

For network paths, use UNC notation: - \\ServerName\Share\Folder

When moving to a network drive, the file is first copied and then deleted at the original location. This may take longer for large files or slow network connections.

### Don't Combine with Delete

The “Delete File” task after “Move File” is not meaningful since the file no longer exists at the original location. If you want to delete the file after moving, use “Copy File” followed by “Delete File” instead.

---

## 14 Delete File

Task: Delete File

*Task: Delete File*

### 14.1 Description

The **Delete File** task removes the PDF file from its original location. The file is moved to the Windows Recycle Bin and can be restored if needed.

### Typical Use Cases

- **Cleanup:** Automatically empty inbox folders after successful processing
  - **Workflow:** Remove original file after copying/moving
  - **Storage Space:** Automatically dispose of files no longer needed
-

## 14.2 General Settings

### Enabled

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

## 14.3 Delete Behavior

### Delete Settings

#### *Delete Settings*

### Move to Recycle Bin

The file is moved to the Windows Recycle Bin and can be restored if needed.

**Advantages:** - Safety in case of accidental deletion - File can be restored from Recycle Bin

**Note:** For security reasons, only moving to the Recycle Bin is possible. Permanent deletion without Recycle Bin is not supported to prevent accidental data loss.

---

## 14.4 Delete Parent Folder

### Delete Parent Folder If Empty

Enable this option to automatically delete the parent folder when it's empty after deleting the PDF file.

**Use Case:** PDFs are placed in day-specific subfolders (e.g., Inbox\2024-12-15\Invoice.pdf). After deleting the last file in folder 2024-12-15, this empty folder should also be removed.

**Note:** Only the immediate parent folder is deleted, not multiple levels.

---

## 14.5 Example: Empty Inbox Folder After Processing

### Initial Situation

Scanned documents are first copied to the archive. Afterwards, the original files in the scanner folder should be deleted.

### Configuration

**Profile Tasks:** 1. **Copy File:** To D:\Archive\... 2. **Delete File:** Move to Recycle Bin

**Delete Settings:** - **Delete option:** Move to Recycle Bin - **Delete parent folder:** No

## Result

After processing, the copy is in the archive, and the original file is in the Recycle Bin (if accidentally deleted, it can be restored).

---

## 14.4 Tips and Notes

### Always as Last Task

The “Delete File” task should **always be at the end** of the task list. After deletion, the file is no longer available at the original location, and subsequent tasks cannot access it.

**Recommended Order:** 1. Rename file 2. Copy file (Archive) 3. Copy file (Backup) 4. Send email 5. **Delete file** (last task)

### Create Safety Copy

Before using “Delete File”, ensure that a copy of the file exists at another location. Add a “Copy File” task before the delete task.

### Empty Recycle Bin Regularly

If you use the Recycle Bin, remember to empty it regularly to free up storage space. Windows can be configured to automatically empty the Recycle Bin.

### Don't Combine with Move

After the “Move File” task, the original file no longer exists at the original location. A subsequent “Delete File” task would fail. Instead use: - “Move File” without subsequent deletion, or - “Copy File” followed by “Delete File”

### Restore from Recycle Bin

If a file was accidentally deleted: 1. Open the Windows Recycle Bin (double-click the Recycle Bin icon) 2. Find the deleted file 3. Right-click → “Restore”

The file is restored to its original location.

---

## 15 Print

Task: Print

*Task: Print*

### 15.1 Description

The **Print** task automatically sends the PDF file to a printer. You can select any installed printer and configure various print settings.

## Typical Use Cases

- **Automatic Printing:** Print incoming invoices directly
  - **Filing:** Automatically print documents for paper filing
  - **Archiving:** Create paper printout in addition to digital archiving
  - **Labels:** Send PDF labels to a label printer
- 

## 15.2 General Settings

### Enabled

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

## 15.3 Printer

### Print Settings

#### *Print Settings*

### Printer Selection

Select the target printer from the list of installed printers. All locally installed and network printers are displayed.

**Note:** Make sure the selected printer is turned on and ready when files are processed.

### Use Default Printer Settings

Enable this option to use the default settings of the selected printer. All other print settings are then ignored.

---

## 15.4 Copies

### Number of Copies

Specify how many copies should be printed (1-999).

### Collate

For multiple copies of multi-page documents: - **Enabled:** Prints complete documents sequentially (1-2-3, 1-2-3, 1-2-3) - **Disabled:** Prints all copies of one page sequentially (1-1-1, 2-2-2, 3-3-3)

---

## 15.5 Page Range

### All Pages

Prints the entire document.

### Pages From/To

Prints only a specific page range: - **From:** First page to print - **To:** Last page to print

**Example:** From 1 to 1 prints only the first page.

---

## 15.6 Duplex (Double-Sided Printing)

### Available Options

Option	Description
<b>Simplex</b>	Single-sided printing
<b>Vertical</b>	Double-sided, sheet is flipped on long edge (like a book)
<b>Horizontal</b>	Double-sided, sheet is flipped on short edge (like a notepad)

**Note:** This option is only available if the printer supports duplex printing.

---

## 15.7 Color

### Color Printing

- **Enabled:** Prints in color (if printer supports it)
- **Disabled:** Prints in black and white or grayscale

**Tip:** Black and white printing is often faster and more cost-effective.

---

## 15.8 Paper

### Paper Size

Select the paper format: - A4 (210 × 297 mm) - Standard in Europe - A3 (297 × 420 mm) - Letter (8.5 × 11 in) - US Standard - Legal (8.5 × 14 in) - Additional printer-specific formats

### Paper Source

Select the paper tray: - Automatic - Tray 1, Tray 2, etc. - Manual feed - Additional printer-specific options

---

## 15.9 Scaling and Orientation

### Size Mode

Option	Description
<b>Fit to page</b>	Scales PDF to paper size (enlarges or reduces)
<b>Actual size</b>	Prints at original size (100%)
<b>Custom scale</b>	Scales to a fixed percentage value
<b>Use printer default</b>	Uses printer's default setting

### Scale in Percent

For "Custom scale": Enter the desired scaling factor (1-999%). - 100% = Original size - 50% = Halved - 200% = Doubled

### Auto-Rotate

Option	Description
<b>Auto-rotate</b>	Automatically rotates page for optimal fit
<b>Do not auto-rotate</b>	Keeps original orientation
<b>Use printer default</b>	Uses printer's default setting

### Auto-Center

Option	Description
<b>Auto-center</b>	Centers content on page
<b>Do not auto-center</b>	Prints from top left corner
<b>Use printer default</b>	Uses printer's default setting

---

## 15.10 Example: Automatically Print Invoices

### Initial Situation

All incoming invoices should be automatically printed on the office printer.

### Configuration

1. **Enabled:** Yes
2. **Printer:** "HP LaserJet Pro" (or your printer name)
3. **Copies:** 1

4. **Page range:** All pages
5. **Duplex:** Simplex
6. **Color:** No (Black and White)
7. **Paper size:** A4
8. **Size mode:** Fit to page

## Result

Each processed PDF file is automatically printed in one copy on the configured printer.

---

## 15.11 Print Presets

### Save Presets

In program options under **Print Settings**, you can save frequently used print configurations as presets. These are then available for quick selection in all profiles.

### Advantages of Presets

- Quick configuration of new profiles
  - Consistent print settings across multiple profiles
  - Changes to a preset affect all profiles that use it
- 

## 15.12 Batch Printing for Large Documents

### Problem

When printing very large PDF files (1000+ pages), the following problems may occur: - Very long spooling times - Extremely large spool files (several GB) - Possible system instability due to memory usage - Printer not responding for long periods

### Solution: Automatic Batch Splitting

The batch printing feature automatically splits large print jobs into smaller batches to relieve the printer spooler.

### Settings

Option	Description	Default
<b>Automatically split into batches</b>	Enables batch printing	Disabled
<b>From page count</b>	Minimum pages for activation (50-10000)	200
<b>Pages per batch</b>	Number of pages per print job (10-1000)	100

## How It Works

1. The program checks the total page count of the print range
2. If page count  $\geq$  "From page count", printing is split into batches
3. Each batch is sent separately to the printer
4. A short pause (500ms) is inserted between batches
5. All pages are printed completely without gaps

## Example

**Document:** 850 pages **Settings:** Activated from 200 pages, 100 pages per batch

**Result:** 9 print jobs - Batch 1: Pages 1-100 - Batch 2: Pages 101-200 - Batch 3: Pages 201-300 - ... - Batch 9: Pages 801-850

## When to Use?

**Enable batch printing when:** - You regularly print PDFs with more than 500 pages - The printer takes long or doesn't respond with large documents - The printer spooler shows error messages - You observe memory problems when printing

**Not necessary when:** - You mostly print smaller documents ( $< 100$  pages) - Your printer handles large jobs without problems

---

---

## 15.13 Print to PDF File

### Overview

When you select a PDF printer (e.g., "Microsoft Print to PDF", "Adobe PDF", "Foxit PDF Printer", "doPDF", "CutePDF Writer", "Bullzip PDF Printer", or "PDFCreator"), an additional configuration section for the output path automatically appears.

**Note:** The program automatically detects PDF printers by name - all printers with "PDF" in their name are treated as PDF printers.

### Output Path Configuration

When selecting a PDF printer, you can configure the following settings:

Setting	Description
<b>Target directory</b>	The folder where the PDF file will be saved
<b>Filename</b>	The name of the output file (supports placeholders)
<b>If file exists</b>	Action when file already exists

### Placeholders in Filename

You can use placeholders to create dynamic filenames:

Placeholder	Description
<FileName>	Filename of source file (without extension)
<ParentDirectory>	Parent folder path
<TodaysYear4>	Current year (4 digits)
<TodaysMonth>	Current month (2 digits)
<TodaysDay>	Current day (2 digits)

**Example:** <FileName>\_printed.pdf → Invoice\_printed.pdf

### If File Exists

Option	Description
<b>Add numbering</b>	Adds a number (e.g., File (1).pdf)
<b>Overwrite</b>	Replaces the existing file
<b>Cancel</b>	Cancel the operation

### Example: Create PDF and Archive

**Initial situation:** Incoming documents should be saved as PDF in an archive folder.

**Configuration:** 1. **Printer:** “Microsoft Print to PDF” 2. **Target directory:**

C:\Archive\PDFs\<TodaysYear4> 3. **Filename:** <FileName>.pdf 4. **If file exists:** Add numbering

**Result:** The file Invoice.pdf is printed as PDF and saved to C:\Archive\PDFs\2026\Invoice.pdf. If the file already exists, it will be saved as Invoice (1).pdf.

## 15.14 Tips and Notes

### Check Print Queue

When many files are processed simultaneously, print jobs may queue up. Check the Windows print queue if problems occur.

### Network Printers

For network printers, ensure that: - The printer is reachable on the network - The Windows user running the program has print permissions - No print dialogs (e.g., PIN entry) are required

### Multiple Printers

You can add multiple “Print” tasks to the same profile to send a file to different printers: 1.

**Task 1:** Print on main printer 2. **Task 2:** Print on archive printer 3. **Task 3:** Print on label printer

## Print Quality

Print quality (DPI) can be set via printer resolution. Higher resolutions lead to better quality but longer print times.

---

# 16 Send Email

Task: Send Email

*Task: Send Email*

## 16.1 Description

The **Send Email** task automatically sends the PDF file via email to one or more recipients. You can specify fixed recipients, determine recipients from extracted data, or read them from a CSV file.

### Typical Use Cases

- **Invoice Delivery:** Automatically send invoices to customers
  - **Document Distribution:** Forward reports to departments
  - **Notification:** Forward incoming documents for information
  - **Archiving:** Send copy via email to an email archive
- 

## 16.2 General Settings

### Enabled

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

## 16.3 Recipients

### Recipient Type

You can choose between three methods for determining recipients:

#### Fixed Email Address

Enter one or more email addresses directly. Separate multiple addresses with semicolon (;).

**Example:** accounting@company.com; archive@company.com

#### Placeholder

The email address is determined from extracted data.

**Example:** <RuleId:1(CustomerEmail)>

The extraction rule must return a valid email address.

### CSV File

The email address is determined from an external CSV file based on an extracted key value.

**Configuration:** - **CSV File:** Path to the mapping file - **Delimiter:** Column separator (default: semicolon) - **Key Value Column:** Column with the search value (e.g., customer number) - **Email Address Column:** Column with the email address

### Example CSV File:

```
CustomerNo;CompanyName;Email
C001;Sample Company Inc;invoice@sample.com
C002;Smith Corp;accounting@smith.com
```

The program searches for the extracted key value in the first column and uses the email address from the third column.

### CC (Copy)

Additional recipients who receive a copy of the email. These addresses are visible to all recipients.

### BCC (Blind Copy)

Additional recipients who receive a copy of the email. These addresses are not visible to other recipients.

---

## 16.4 Email Content

### Subject

The subject line of the email. You can: - Enter fixed text - Use placeholders

**Example:** Invoice <RuleId:1(InvoiceNo)> dated <RuleId:2(Date)>

Result: "Invoice INV-12345 dated 12/15/2024"

### Message

The body text of the email. You can: - Enter fixed text - Use placeholders for dynamic content - Use line breaks

### Example:

Dear Sir or Madam,

please find attached invoice <RuleId:1(InvoiceNo)> for <RuleId:2(Amount)> EUR.

Best regards  
Sample Company Inc

### HTML Message

Enable this option to format the message as HTML. This allows: - Formatting (bold, italic, underlined) - Colors and fonts - Tables and lists - Embedded images

### Example HTML:

```
<html>  
<body>  
<p>Dear Sir or Madam,</p>  
<p>please find attached <b>invoice <RuleId:1(InvoiceNo)></b>.</p>  
<p>Best regards<br>  
<i>Sample Company Inc</i></p>  
</body>  
</html>
```

---

## 16.5 Attachments

### Attach PDF File

By default, the processed PDF file is sent as an attachment. You can disable this option if you only want to send other attachments.

### Additional Attachments

Add static files that should always be sent: - Terms.pdf - PriceList.pdf - Logo.png

### Dynamic Attachments

Files determined from a folder based on a search pattern: - **Source Folder:** Folder to search in (placeholders possible) - **Filename Contains:** Search pattern for the filename - **Existence:** Must exist / Can exist - **Wait Time:** Seconds to wait for the file

**Use Case:** Attach related delivery note PDF from another folder to the invoice.

### File Group Attachments

If you use file groups, you can add files from secondary subgroups as attachments.

### Replace Special Characters in Attachment Names

Enable this option to replace special characters (umlauts, special characters) in attachment names with ASCII characters. This can avoid compatibility issues with some email systems.

---

## 16.6 Sending Options

### Save as Draft

The email is not sent immediately but saved as a draft. This option is available for all sending methods:

Sending Method	Behavior
<b>Outlook</b>	The draft is saved in the Outlook Drafts folder
<b>SMTP</b>	The draft is saved locally as an .eml file in the application data folder
<b>Exchange Online</b>	The draft is saved locally as an .eml file in the application data folder

For SMTP and Exchange Online, drafts are managed via the **Drafts window** (see section 16.6.1).

### Open Draft (Outlook only)

When “Save as Draft” is enabled, the draft is automatically opened in Outlook. This option is only available when sending via Outlook.

#### 16.6.1 Managing Email Drafts

When drafts have been saved for SMTP or Exchange Online, the “**Email Drafts (N)**” button appears in the main toolbar, where N indicates the number of saved drafts.

Clicking this button opens the **Drafts window**:

Email Drafts

*Email Drafts*

#### Features:

**Left: Draft List** - Shows all saved drafts with recipient and subject - Enable/disable individual drafts via checkbox for sending

#### Right: Message Preview

Tab **Message**: - Sender selection: Switch between saved and current send settings - Recipients (To, CC, BCC) and subject — directly editable - Message text as preview - “Edit message...” to modify the message text

Tab **Attachments**: - List of all attachments with file size - preview for the selected attachment - Add, open, or remove attachments

**Sending:** The “Send selected messages...” button opens the **Send dialog**:

### Email Sending

#### *Email Sending*

- List of messages to be sent with status display
- **Pause after sending each message** — Prevents messages from being rejected by the email provider due to mass sending (default: 10 seconds)
- Progress bar during sending

**Note:** After successful sending, the status in the processing log is automatically updated from “Draft saved for” to “File sent to”.

#### **Task-Specific Send Settings**

Enable this option to use different email settings for this task than the global program options. This allows e.g.: - Sending via a different email account - Different SMTP settings - Different Outlook configuration

---

## 16.7 Attachment Encryption

### **Encrypt Attachments**

Enable this option to protect PDF attachments with a password before sending.

#### **Password**

Option	Description
<b>Static Password</b>	Fixed password for all emails
<b>Dynamic Password</b>	Password from extracted data (e.g., customer number)

---

## 16.8 Example: Send Invoices to Customers

### **Initial Situation**

Created invoices should be automatically sent to the respective customer via email. The customer email address is determined from a CSV file.

### **Prerequisites**

- Extraction rule for “CustomerNumber” created
- CSV file with customer data available

### **Configuration**

1. **Enabled:** Yes
2. **Recipient Type:** CSV File

3. **CSV File:** C:\Data\Customers.csv
4. **Key Column:** 1 (Customer Number)
5. **Email Column:** 3 (Email Address)
6. **Subject:** Your Invoice <RuleId:1(InvoiceNo)>
7. **Message:** (Standard text with placeholders)
8. **HTML Message:** No

## **Result**

Each processed invoice is automatically sent to the customer's email address determined from the CSV file.

---

## 16.9 Sending Methods

The email can be sent via different methods. Configuration is done in program options or in task-specific send settings.

### **SMTP**

Direct sending via an SMTP server: - Server address and port - Authentication (username/password) - Encryption (SSL/TLS)

### **Outlook**

Sending via locally installed Microsoft Outlook: - Uses the configured Outlook account - Supports drafts and delayed sending - Requires installed Outlook

### **Exchange Online**

Sending via Microsoft 365 / Exchange Online: - OAuth authentication - Tenant ID and Application ID required - No local Outlook installation needed

---

## 16.7 Tips and Notes

### **Validate Email Addresses**

When using placeholders or CSV files, ensure that determined email addresses are valid. Invalid addresses lead to sending errors.

### **Consider Attachment Size**

Many email servers limit attachment size (often 10-25 MB). For large PDF files: - Use the "Compress PDF" task beforehand (with a separate profile) - Split large files (with a separate profile) - Use alternative transfer methods (cloud links)

### **Limit Page Count (Recommended)**

When automatically sending confidential documents (payslips, invoices), it is recommended to set a filter for the maximum page count (e.g., maximum 3 pages). This prevents bulk documents from being sent as a whole if the upstream splitting is not configured correctly.

To configure this, set the filter criterion “Number of pages” with the operator “is less than or equal to” and the value 3 (or another value appropriate for your scenario) in the profile under “Filter”.

### **Avoid Spam Filters**

To prevent automatically sent emails from being classified as spam: - Use a reputable sender - Avoid typical spam phrases - Configure SPF, DKIM, and DMARC for your domain

### **Multiple Email Tasks**

You can add multiple “Send Email” tasks to the same profile: - Different recipients - Different message texts - Different email accounts

### **Error Handling**

If email sending fails: - Check the connection to the email server - Verify login credentials - Check the email address - See the log file for details

---

---

## 17 Merge PDFs

Task: Merge PDFs

*Task: Merge PDFs*

### 17.1 Description

The **Merge PDFs** task combines multiple PDF files or image files into a single PDF document. The order of merged files corresponds to the order in the configuration.

#### **Typical Use Cases**

- **Document Bundling:** Merge invoice with related delivery notes
- **Scan Correction:** Combine multiple scan pages into one document
- **Archiving:** Archive main document with attachments in one file
- **Image Collection:** Convert multiple image files (JPG, PNG, TIFF) into one PDF

**Important:** This task creates a new file in the configured target folder. The original files remain unchanged. Further tasks contained in the current profile all refer to the original file. The merged PDF created by this task must be further processed with a **separate profile** that monitors the corresponding output folder if needed.

---

---

## 17.2 General Settings

### Enabled

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

## 17.3 Files to Merge

### File and Subgroup Selection

In the file list, you specify which files should be merged. Via the toolbar you can:

- **Add File:** Select one or more PDF files that should always be added
- **Add Subgroup:** Add files from a configured file group (secondary file group)
- **Remove:** Deletes the selected entry from the list
- **Move Up/Down:** Changes the order of files in the merged PDF

### Order of Merging:

1. The triggering PDF file (or with file groups: all files from the primary file group)
2. Files from secondary file groups (in list order)
3. Manually added files (in list order)
4. All PDFs from the parent folder (if enabled)
5. All images from the parent folder (if enabled)

### Additionally Append All PDF Files from Parent Folder

Enable this option to automatically append all PDF files from the same folder as the triggering file. Files are added in alphabetical order.

**Use Case:** You receive multi-page scans as separate PDF files (Page1.pdf, Page2.pdf, Page3.pdf). With this option, all are automatically merged.

### Additionally Append All Image Files from Parent Folder

Enable this option to automatically convert and append all image files from the same folder. Images are sorted alphabetically and converted to PDF pages.

**Supported Image Formats:** - JPEG (.jpg, .jpeg, .jpe, .jfif, .jif) - PNG (.png) - TIFF (.tif, .tiff) - BMP (.bmp) - GIF (.gif)

**Use Case:** In addition to scanned PDFs, there are also photographed documents as JPEG files. All files should be combined in one PDF.

---

## 17.4 Storage Location

### Overwrite Original File

When enabled, the primary file (the file that triggered the profile) is replaced by the merged file. No separate target folder or filename is required.

### Directory

Specify the target directory for the merged file. You can: - Enter a fixed path (e.g., D:\Archive\Merged) - Select the folder via **Browse...** - Use placeholders for dynamic folder paths

### Examples with Placeholders:

Input	Result
D:\Archive\<<TodaysYear4>\<TodaysMonth>	D:\Archive\2024\12
D:\Customers\<<RuleId:1(Customer)>	D:\Customers\Sample Company Inc

**Note:** It's recommended to use a separate folder for each processing step to ensure clear separation.

### Filename

Set the name for the merged file. You can: - Leave the field empty (original name is used) - Enter a fixed name - Use placeholders for dynamic names

### Examples:

Input	Result
(empty)	Scan001.pdf (original name)
<RuleId:1(Date)>_Merged	2024-12-15_Merged.pdf
Complete_<FileName>	Complete_Scan001.pdf

### Name Collisions

Choose what should happen if a file with the target name already exists:

Option	Description
<b>Overwrite</b>	The existing file is replaced. <b>Caution:</b> Data may be lost!
<b>Append number</b>	Adds a number: File.pdf, File(1).pdf, File(2).pdf
<b>Append date</b>	Adds processing date: File(2024-12-15).pdf
<b>Append date and time</b>	Adds date and time: File(2024-12-15_14-30-25).pdf
<b>Cancel operation</b>	The merge is not performed

**Recommendation:** For most use cases, **Append number** is the safest option.

---

## 17.5 File Date

### Adjust Creation and Modification Date

Optionally, you can change the file date of the merged file:

Option	Description
<b>Do not change</b>	File automatically receives current date
<b>Creation date of original file</b>	Uses original creation date
<b>Modification date of original file</b>	Uses modification date
<b>PDF creation date</b>	Date from PDF metadata
<b>Extracted date</b>	A date obtained with an extraction rule
<b>Current date</b>	Sets today's date and current time

---

## 17.6 Afterwards

### Call External Program

After merging, an external program can be started automatically.

**Program:** Path to the executable file (e.g., C:\Program Files\...\program.exe)

**Parameters:** Command line parameters. Available placeholders: -

<PathIncludingFilename> - Full path of merged file - <ParentDirectory> - Path of parent folder - <Filename> - Filename of merged file

**Example:** Automatically open merged file: - Program: cmd.exe - Parameters: /c start "" "<PathIncludingFilename>"

---

## 17.7 Example: Merge Invoice with Delivery Note

### Initial Situation

You receive invoices and related delivery notes as separate PDF files. Both should be automatically merged into one document.

### Prerequisites

1. Two profiles for file recognition:
  - Profile "Invoices": Recognizes invoice PDFs
  - Profile "Delivery Notes": Configured as secondary subgroup
2. Extraction rule for customer number (for matching)

## Configuration

1. **Enabled:** Yes
2. **Files:** Subgroup “Delivery Notes” added
3. **Directory:** D:\Archive\Complete
4. **Filename:** <RuleId:1(InvoiceNo)>\_complete
5. **On name collision:** Append number

## Result

Original Files	Merged File
Invoice_12345.pdf + DeliveryNote_12345.pdf	D:\Archive\Complete\12345_complete.pdf

---

## 17.8 Example: Merge All Scans from a Folder

### Initial Situation

Your scanner saves each page as a separate PDF file (Scan\_001.pdf, Scan\_002.pdf, etc.). These should be automatically combined into one document.

### Configuration

1. **Enabled:** Yes
2. **Additionally append all PDF files from parent folder:** Yes
3. **Directory:** D:\Documents\Merged
4. **Filename:** <TodaysYear4>-<TodaysMonth>-<TodaysDay>\_Scan
5. **On name collision:** Append number

### Process

1. Profile recognizes Scan\_001.pdf as triggering file
2. All other PDFs in the same folder are automatically appended in alphabetical order
3. Merged file is saved in target folder

**Note:** To prevent each individual scan file from triggering a separate merge operation, configure file grouping to wait for all files.

---

## 17.6 Tips and Notes

### Consider Order

The order of files in the merged PDF corresponds to the configuration order. Use the Up/Down buttons to set the desired order.

### Combine with File Grouping

For merging related documents (e.g., invoice + delivery note), using file grouping is recommended: 1. Configure the main profile with the triggering file 2. Create secondary file groups for related documents 3. Merging then automatically waits for all needed files

### **Prepare Images Optimally**

When merging image files: - Name files so alphabetical sorting matches desired order (e.g., 01\_Cover.jpg, 02\_Content.jpg) - High image resolution leads to larger PDF files

### **Further Processing of Merged PDF**

The created merged PDF is in the configured target folder. To process it further (e.g., OCR, email), create a separate profile that monitors this target folder.

### **Consider Storage Space**

When merging many large files, the resulting file can become considerably large. Ensure sufficient storage space is available.

### **Password-Protected PDFs**

Password-protected PDFs can be merged if the password is stored in the password list (program options) or in the profile.

---

## **18 Split PDFs**

### Task: Split PDFs

*Task: Split PDFs*

### **18.1 Description**

The **Split PDFs** task divides a multi-page PDF document into multiple individual files. Splitting can be done according to various criteria: page count, file size, bookmarks, keywords, barcodes, blank pages, or separator pages.

#### **Typical Use Cases**

- **Batch Processing:** Split large scan files into individual documents
- **Invoice Separation:** Divide combined PDFs into individual invoices
- **Archiving:** Save documents by bookmark into separate files
- **Form Processing:** Split multi-page forms based on separator pages

**Important:** This task creates new files in the configured target folder. The original file remains unchanged. Further tasks contained in the current profile all refer to the original file. The partial documents created by this task must be further processed with a **separate profile** that monitors the corresponding output folder if needed.

---

## 18.2 General Settings

### Enabled

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

## 18.3 Splitting Method

Choose the criterion by which the PDF should be split:

Method	Description
<b>Page Count</b>	Split after a fixed number of pages
<b>File Size</b>	When a certain file size is exceeded
<b>Top-Level Bookmarks</b>	At each bookmark of the top level
<b>Keywords</b>	When certain texts are present or not present on pages
<b>Barcode</b>	Based on QR codes or barcodes
<b>Placeholder (Value Change)</b>	When an extracted value changes
<b>Separator Page with Text</b>	At pages with specific text content
<b>Blank Pages</b>	At blank pages

---

## 18.4 By Page Count

### Pages per Document

Specify after how many pages a new partial document should begin.

**Example:** With a 10-page PDF and setting “3 pages”, the result is: - Partial document 1: Pages 1-3 - Partial document 2: Pages 4-6 - Partial document 3: Pages 7-9 - Partial document 4: Page 10

---

## 18.5 By File Size

### Maximum File Size (MB)

Specify the maximum size per partial document in megabytes. When adding another page would exceed the limit, a new partial document begins.

**Note:** Actual size may vary depending on PDF content. The setting is a guideline.

---

## 18.6 By Top-Level Bookmarks

With this method, a new partial document is created at each top-level bookmark. This is particularly useful for structured documents like manuals or reports.

**Prerequisite:** The PDF must contain bookmarks.

### Special Placeholder

When splitting by bookmarks, an additional placeholder is available for the filename: - <SplittingBookmarkTitle> - The title of the bookmark

**Example:** - Filename template: <FileName>\_<SplittingBookmarkTitle> - Result: Manual\_Chapter1.pdf, Manual\_Chapter2.pdf, etc.

---

## 18.7 By Keywords

Splits the document when certain texts are found on a page.

### Start New Document at Pages With

Enter the text that should trigger a split. When this text is found on a page, a new partial document starts there.

**Example:** “Invoice No.” - Each page containing this text starts a new document.

### Start New Document at Pages Without

Enter text that is **not** present on separator pages. Useful when separator pages are blank pages or special pages.

### End New Document at Pages With

Optional: Text that marks the end of a partial document.

### End New Document at Pages Without

Optional: Ends partial document at pages that **don't** contain this text.

### Trigger Error if Keyword Not Found

Enable this option if processing should fail when the keyword is not found anywhere in the document.

---

## 18.8 By Barcode

Splits the document based on QR codes or barcodes.

### Barcode Format

Select the format of the barcode to recognize: - QR Code - Code 128 - Code 39 - EAN-13, EAN-8 - Data Matrix - PDF417 - Aztec - UPC-A, UPC-E

### Image Noise Suppression and Rotation Correction

Determines how thoroughly to search for barcodes: - **Low** - Fast recognition, lower accuracy - **Medium** - Balanced setting (default) - **High** - Thorough search - **Very High** - Maximum accuracy, slowest processing

### Trigger

Choose when a new split should occur:

Trigger	Description
<b>All Occurrences of Barcode</b>	At every found barcode
<b>Change in Barcode Data</b>	When barcode content changes
<b>Barcode Data with Text</b>	When barcode contains specific text

### Additional Options

- **Skip Pages Before First Valid Barcode** - Pages without barcode at the beginning are not included in partial documents
- **Skip Pages After Last Valid Barcode** - Pages without barcode at the end are not included
- **Exclude Separator Pages** - Pages with the triggering barcode are not included in partial documents

---

## 18.9 By Placeholder (Value Change)

Splits the document when an extracted value changes from page to page.

### Select Extraction Rule

Select a previously defined extraction rule. When the extracted value changes between two pages, a new partial document begins.

**Example:** You have a rule “CustomerNumber” that extracts the customer number from each page. With a combined invoice containing pages for different customers, it automatically separates at each new customer number.

### Skip Pages Until First Valid Extraction

Enable this option to skip pages at the beginning where no data can be extracted.

---

## 18.10 By Separator Page with Text

Recognizes special separator pages based on defined keywords.

### **Page Contains**

Enter the text present on separator pages (e.g., “— SEPARATOR —”).

### **Page Does Not Contain**

Optional: Text that **must not** be present on separator pages.

---

## 18.11 By Blank Pages

Splits the document at blank pages. A page is considered blank if it contains no text (0 characters).

**Use Case:** Scanned documents are often separated by blank pages between individual documents.

---

## 18.12 Page Exclusion

### **Timing of Exclusion**

- **After Splitting** - Exclusion rules are applied to partial documents
- **Before Splitting** - Exclusion rules are applied to the entire document before splitting

### **Exclude Pages with Less Than X Characters**

Removes pages with fewer than the specified character count. Useful for removing blank or separator pages.

### **Exclude Pages with Text**

Excludes pages containing the specified text.

### **Exclude Pages Without Text**

Excludes pages that **don't** contain the specified text.

---

## 18.13 Source for New Documents

### **Start Partial Documents with Empty Document**

Partial documents are created as new, empty PDFs and only relevant pages are inserted. This is the default setting and creates smaller files.

## Start Partial Documents with Original's Structure and Metadata

Partial documents retain the PDF structure and metadata (author, title, etc.) of the original. Choose this option when this information is important.

---

### 18.14 Storage Location

#### Directory

Specify the target directory for partial documents.

**Note:** It's recommended to use a separate folder for each processing step to ensure clear separation.

#### Filename

Set the naming scheme for partial documents. In addition to standard placeholders, special placeholders are available:

Placeholder	Description	Example
<SplittingNumber>	Sequential number of partial document	1, 2, 3, ...
<SplittingNumberWithLeadingZeros{N}>	Number with leading zeros (N digits)	001, 002, ...
<SplittingBookmarkTitle>	Bookmark title (only with bookmark splitting)	Chapter1

#### Examples:

Input	Result
<FileName>_Part<SplittingNumber>	Invoice_Part1.pdf, Invoice_Part2.pdf
<FileName>_<SplittingNumberWithLeadingZeros{3}>	Invoice_001.pdf, Invoice_002.pdf
<SplittingBookmarkTitle>	Introduction.pdf, MainPart.pdf

#### Name Collisions

Choose what should happen if a file with the target name already exists:

Option	Description
<b>Overwrite</b>	Existing file is replaced
<b>Append number</b>	Adds a number
<b>Append date</b>	Adds processing date
<b>Append date and time</b>	Adds date and time
<b>Cancel operation</b>	Split is not performed

---

## 18.15 File Date

### Adjust Creation and Modification Date

Optionally, you can change the file date of partial documents:

Option	Description
<b>Do not change</b>	Files automatically receive current date
<b>Creation date of original file</b>	Uses original creation date
<b>Modification date of original file</b>	Uses modification date
<b>PDF creation date</b>	Date from PDF metadata
<b>Extracted date</b>	A date obtained with an extraction rule
<b>Current date</b>	Sets today's date

---

## 18.16 Afterwards

### Call External Program

After splitting, an external program can be started automatically for each partial document.

**Program:** Path to executable file

**Parameters:** Command line parameters. Available placeholders: -

<PathIncludingFilename> - Full path of partial document - <ParentDirectory> - Path of parent folder - <Filename> - Filename of partial document

---

## 18.17 Example: Split Combined Invoice by Customer

### Initial Situation

You receive a monthly combined invoice as PDF containing invoices for multiple customers. Each customer invoice starts with the text "Invoice for:".

### Configuration

1. **Enabled:** Yes
2. **Splitting Method:** Keywords
3. **Start new document at pages with:** Invoice for:
4. **Directory:** D:\Invoices\Split
5. **Filename:** Invoice\_<SplittingNumber>\_<TodaysYear4><TodaysMonth>
6. **On name collision:** Append number

## Result

Original File	Partial Documents
CombinedInvoice_December.pdf (30 pages)	Invoice_1_202412.pdf, Invoice_2_202412.pdf, etc.

---

## 18.18 Example: Split Manual by Chapters

### Initial Situation

A manual with bookmarks for each chapter should be split into individual chapter PDFs.

### Configuration

1. **Enabled:** Yes
2. **Splitting Method:** Top-Level Bookmarks
3. **Directory:** D:\Documents\Chapters
4. **Filename:** <FileName>\_<SplittingBookmarkTitle>
5. **On name collision:** Append number

## Result

Bookmark	Partial Document
“Introduction”	Manual_Introduction.pdf
“Chapter 1 - Installation”	Manual_Chapter 1 - Installation.pdf
“Chapter 2 - Configuration”	Manual_Chapter 2 - Configuration.pdf

---

## 18.5 Tips and Notes

### Further Processing of Partial Documents

Created partial documents are in the configured target folder. To process them further (e.g., OCR, renaming, email), create a separate profile that monitors this target folder.

### Numbering with Leading Zeros

For better sorting in file manager, use <SplittingNumberWithLeadingZeros{3}> instead of <SplittingNumber>. Files are then sorted correctly: 001, 002, ... 010, 011 instead of 1, 10, 11, 2, 3.

### Optimize Barcode Recognition

If barcodes are not reliably recognized: - Increase preprocessing to “High” or “Very High” - Ensure correct barcode format is selected - Enable machine learning for barcode recognition in program options

## Remove Separator Pages

To not have separator pages in partial documents: - With barcode splitting: Enable “Exclude separator pages” - With other methods: Use page exclusion with separator page text

## Combination of Criteria

If a single criterion is not sufficient, you can apply a second profile with a different splitting method to the partial documents after splitting.

## Consider Storage Space

When splitting large PDFs, many individual files are created. Ensure sufficient storage space is available.

---

# 19 OCR (Text Recognition)

Task: OCR

*Task: OCR*

## 19.1 Description

The **OCR** task (Optical Character Recognition) converts scanned documents or image PDFs into searchable PDFs. After processing, text in the PDF can be selected, copied, and searched.

### Typical Use Cases

- **Archiving:** Make scanned documents searchable
- **Data Extraction:** Extract text from scanned invoices for further processing
- **Compliance:** Prepare documents for full-text search in document management systems
- **Accessibility:** Make PDFs accessible for screen readers

**Important:** This task creates a new file in the configured target folder. The original file remains unchanged. Further tasks contained in the current profile all refer to the original file. The searchable PDF created by OCR must be further processed with a **separate profile** that monitors the corresponding output folder if needed.

---

## 19.2 General Settings

**Enabled**

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

## 19.3 Language

### Primary Language

Select the main language of the text to be recognized. Correct language selection significantly improves recognition accuracy.

**Available Languages:** Over 100 languages, including: - German (German Best) - English (English Best) - French, Spanish, Italian - And many more

**Note:** Languages not installed can be downloaded via **Tools → Install Languages**.

### Use Secondary Language

Enable this option when documents contain text in two languages (e.g., German-English contracts).

### Secondary Language

Select the second language appearing in the document.

---

## 19.4 DPI Setting (Resolution)

The DPI setting (Dots Per Inch) affects text recognition quality and output file size.

### Optimized

The program automatically selects an optimal DPI setting based on document content.

### Based on Images

DPI is determined based on the resolution of images contained in the PDF. This is the default setting.

### Custom

You can set a fixed DPI value:

DPI	Usage
96-150	Fast processing, lower quality
175-225	Good balance between quality and speed (recommended)
300	High quality, longer processing
600-1200	Maximum quality, only for special requirements

---

## 19.5 Image Optimization

### Correct Page Orientation Automatically

Detects and corrects page orientation (0°, 90°, 180°, 270°) before OCR processing. This option is especially useful for documents from scanners that feed pages in mixed portrait and landscape orientation. Correction is only applied when detected with high confidence.

**Note:** This option is disabled by default. Enable it if you regularly process incorrectly rotated scans.

### Correct Skew (Deskew)

Automatically corrects slightly skewed scanned documents. Improves recognition for non-exactly aligned scans.

### Sharpen

Increases image sharpness before text recognition. Helpful for slightly blurry scans.

### Increase Contrast

Improves contrast between text and background. Useful for faded or weak text.

### Convert PDF Content to Images First

Converts entire PDF content to images before OCR. Enable this option for PDFs containing both text and images with text.

### Upscale Low Resolutions

Automatically upscales low-resolution images to improve recognition accuracy.

---

## 19.6 Handling Files That Already Contain Text

Specify how to handle PDFs that already contain searchable text:

Option	Description
<b>Ignore</b>	PDF is not processed (default)
<b>Process Anyway</b>	OCR is performed even if text is present
<b>Only Copy to Target Directory</b>	PDF is copied to target folder without OCR

**Recommendation:** Leave setting on “Ignore” unless you have a specific reason for another choice.

---

## 19.7 Character Restriction

### Restrict Characters

Enable this option to limit recognition to specific characters. This can improve accuracy for specialized documents.

### Only Allow Following Characters (Whitelist)

Specify characters that should be recognized. All others are ignored.

#### Default Characters:

ABCDEFGHIJKLMNOPQRSTUVWXYZabcdefghijklmnopqrstuvwxyz0123456789.,-?!

### Exclude Following Characters (Blacklist)

Specify characters that should **not** be recognized.

**Use Case:** When recognizing article numbers consisting only of numbers and letters, special characters can be excluded.

---

---

## 19.8 Processing

### Multithreading

Enables parallel processing of multiple pages. Significantly speeds up OCR on systems with multiple processor cores.

**Recommendation:** Keep enabled unless stability issues occur.

---

---

## 19.9 Storage Location

### Directory

Specify the target directory for searchable PDFs.

**Note:** It's recommended to use a separate folder for each processing step to ensure clear separation.

### Filename

Set the name for the output file. You can: - Leave field empty (original name is used) - Enter a fixed name - Use placeholders for dynamic names

### Examples:

Input	Result
(empty)	Scan001.pdf (original name)

Input	Result
<FileName>_OCR	Scan001_OCR.pdf
<TodaysYear4>-<TodaysMonth>-<TodaysDay>_<FileName>	2024-12-15_Scan001.pdf

### Name Collisions

Choose what should happen if a file with the target name already exists:

Option	Description
<b>Overwrite</b>	Existing file is replaced
<b>Append number</b>	Adds a number
<b>Append date</b>	Adds processing date
<b>Append date and time</b>	Adds date and time
<b>Cancel operation</b>	OCR is not performed

## 19.10 File Date

### Adjust Creation and Modification Date

Optionally, you can change the file date of the output file:

Option	Description
<b>Do not change</b>	File automatically receives current date
<b>Creation date of original file</b>	Uses original creation date
<b>Modification date of original file</b>	Uses modification date
<b>PDF creation date</b>	Date from PDF metadata
<b>Extracted date</b>	A date obtained with an extraction rule
<b>Current date</b>	Sets today's date

## 19.11 Afterwards

### Call External Program

After OCR, an external program can be started automatically.

**Program:** Path to executable file

**Parameters:** Command line parameters. Available placeholders: -

<PathIncludingFilename> - Full path of OCR file - <ParentDirectory> - Path of parent folder - <Filename> - Filename of OCR file

## 19.12 Example: Make Scanned Invoices Searchable

### Initial Situation

Your scanner creates image PDFs without searchable text. These should be automatically processed with OCR so you can later search for invoice numbers or amounts.

### Configuration

1. **Enabled:** Yes
2. **Primary Language:** English Best
3. **DPI Setting:** Based on Images
4. **Correct Skew:** Yes
5. **Handling files with text:** Only copy to target directory
6. **Directory:** D:\Archive\OCR
7. **Filename:** <FileName>
8. **On name collision:** Append number

### Result

Original File	OCR File
C:\Scanner\Scan001.pdf (image)	D:\Archive\OCR\Scan001.pdf (searchable)

---

## 19.13 Example: Process Multilingual Documents

### Initial Situation

You process international contracts containing both German and English text.

### Configuration

1. **Primary Language:** German Best
2. **Use Secondary Language:** Yes
3. **Secondary Language:** English Best
4. **DPI Setting:** 225 (Custom)
5. **Sharpen:** Yes

### Note

Using two languages increases processing time but significantly improves recognition accuracy for mixed-language documents.

---

## 19.7 Tips and Notes

### Further Processing of OCR PDF

The created searchable PDF is in the configured target folder. To process it further (e.g., data extraction, renaming, email), create a separate profile that monitors this target folder.

### Install Languages

Not all OCR languages are installed by default. Use **Tools → Install Languages** to download additional languages.

### Improve Recognition Quality

If text recognition is unsatisfactory: 1. Increase DPI setting 2. Enable image optimizations (Sharpen, Contrast, Deskew) 3. Ensure correct language is selected 4. For problems with certain characters: Use character restriction

### Processing Time

OCR processing is computationally intensive. Factors affecting speed: - Document page count - Selected DPI setting - Enabled image optimizations - Use of two languages - Available processor power

### Quality vs. Speed

Priority	Recommended Settings
<b>Speed</b>	Low DPI (150), no image optimization
<b>Balance</b>	DPI 225, Correct skew (default)
<b>Quality</b>	DPI 300+, all image optimizations enabled

### Already Searchable PDFs

When you choose “Ignore” for files with text, already searchable PDFs are not processed again. This saves time and prevents quality loss from multiple processing.

### Password-Protected PDFs

Password-protected PDFs can be processed if the password is stored in the password list (program options) or in the profile and content extraction is allowed.

---

## 20 Save Attachments

Task: Save Attachments

*Task: Save Attachments*

## 20.1 Description

The **Save Attachments** task extracts embedded files from a PDF document and saves them as separate files. PDF attachments can be any file type, such as Excel spreadsheets, Word documents, images, or additional PDFs.

### Typical Use Cases

- **E-Invoice:** Extract XML data from ZUGFeRD/Factur-X invoices
  - **Document Archiving:** Archive attached source files separately
  - **Data Processing:** Extract embedded tables for further processing
  - **Backup:** Save all attachments of a PDF file
- 

## 20.2 General Settings

### Enabled

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

## 20.3 Attachment Filter

### Attachment Name Contains

Enter text that must be contained in the attachment name. Only attachments whose name contains this text are extracted.

**Examples:** - factur-x - Only ZUGFeRD XML files - .xlsx - Only Excel files - (empty) - All attachments

### Attachment Name Does Not Contain

Enter text that **must not** be contained in the attachment name. Attachments with this text in the name are excluded.

**Example:** - thumbnail - Exclude preview images - .tmp - Exclude temporary files

### Combined Filtering

You can combine both filters: - **Contains:** .xml - **Does Not Contain:** metadata

Result: All XML files except metadata files are extracted.

---

## 20.4 Storage Location

### Directory

Specify the target directory for extracted attachments. You can: - Enter a fixed path (e.g., D:\Attachments) - Select the folder via **Browse...** - Use placeholders for dynamic folder paths

### Examples with Placeholders:

Input	Result
D:\Attachments\<<TodaysYear4>\<Today sMonth>	D:\Attachments\2024\12
D:\Customers\<<RuleId:1(Customer)>\Attachments	D:\Customers\Sample Company Inc\Attachments

**Note:** It's recommended to use a separate folder for each processing step to ensure clear separation.

### Filename

The attachment filename is preserved by default. However, you can set a custom name:

- Leave field empty (original attachment name is used)
- Enter a fixed name
- Use placeholders for dynamic names

**Note:** When multiple attachments exist and you use a fixed name, files are handled according to the selected name collision option.

### Name Collisions

Choose what should happen if a file with the target name already exists:

Option	Description
<b>Overwrite</b>	Existing file is replaced
<b>Append number</b>	Adds a number: Attachment.pdf, Attachment(1).pdf
<b>Append date</b>	Adds processing date
<b>Append date and time</b>	Adds date and time
<b>Cancel operation</b>	Attachment is not saved

## 20.5 File Date

### Adjust Creation and Modification Date

Optionally, you can change the file date of extracted attachments:

Option	Description
<b>Do not change</b>	File automatically receives current date
<b>Creation date of original file</b>	Uses PDF's creation date
<b>Modification date of original file</b>	Uses PDF's modification date
<b>PDF creation date</b>	Date from PDF metadata
<b>Extracted date</b>	A date obtained with an extraction rule
<b>Current date</b>	Sets today's date

## 20.6 Afterwards

### Call External Program

After saving each attachment, an external program can be started automatically.

**Program:** Path to executable file

**Parameters:** Command line parameters. Available placeholders: -

<PathIncludingFilename> - Full path of attachment - <ParentDirectory> - Path of parent folder - <Filename> - Filename of attachment

**Example:** Automatically open extracted Excel file: - Program: cmd.exe - Parameters: /c start "" "<PathIncludingFilename>"

## 20.7 Example: Extract ZUGFeRD XML

### Initial Situation

You receive electronic invoices in ZUGFeRD format. These contain an embedded XML file with structured invoice data that you want to extract for your accounting software.

### Configuration

1. **Enabled:** Yes
2. **Attachment name contains:** factur-x or zugferd
3. **Attachment name does not contain:** (empty)
4. **Directory:** D:\ZUGFeRD\XML
5. **Filename:** <RuleId:1(InvoiceNo)>.xml
6. **On name collision:** Append number

### Result

PDF File	Extracted Attachment
Invoice_2024001.pdf (contains factur-x.xml)	D:\ZUGFeRD\XML\2024001.xml

---

## 20.8 Example: Extract All Attachments from Document Collection

### Initial Situation

You receive PDF documents with various embedded files (images, tables, additional PDFs) that should all be extracted.

### Configuration

1. **Enabled:** Yes
2. **Attachment name contains:** (empty - all attachments)
3. **Attachment name does not contain:** (empty)
4. **Directory:** D:\Extracted\<>FileName<
5. **Filename:** (empty - keep original names)
6. **On name collision:** Append number

### Result

For each PDF, a subfolder with the PDF name is created containing all extracted attachments:

```
D:\Extracted\  
├── Report2024\  
│   ├── Table.xlsx  
│   ├── Chart.png  
│   └── SourceData.csv  
└── Presentation\  
    ├── Logo.png  
    └── Notes.docx
```

---

## 20.5 Tips and Notes

### No Attachments Present

If a PDF contains no attachments, the task is skipped without error. Simply no files are extracted.

### Check Attachments

To check if a PDF contains attachments: 1. Open PDF in a PDF viewer 2. Look for a paperclip symbol or attachments section 3. Or use the “Attachment count” filter in profile settings

### Filtering with Regular Expressions

The “Attachment name contains” and “Attachment name does not contain” fields support regular expressions: - <BeginOfRegex>.\*\.xml\$<EndOfRegex> - All files with .xml extension

## Combine with Other Tasks

Typical combinations: 1. **Save Attachments + Copy File**: Archive invoice and extract XML 2. **Save Attachments + Send Email**: Send XML to accounting 3. **Save Attachments + Rename File**: Rename PDF based on extracted data

## ZUGFeRD/Factor-X Standard

For ZUGFeRD/Factor-X invoices, the embedded XML file is typically named: - factur-x.xml (Factor-X) - zugferd-invoice.xml (ZUGFeRD 1.0) - xrechnung.xml (XRechnung)

## File Types

PDF attachments can have any file type. The task extracts files unchanged. The file extension is preserved.

---

# 21 Save Extractions

Task: Save Extractions

*Task: Save Extractions*

## 21.1 Description

The **Save Extractions** task exports data obtained with extraction rules to an external file. Currently CSV format is supported, which can be imported by virtually all applications.

### Typical Use Cases

- **Accounting**: Export invoice data (number, date, amount) for import into accounting software
  - **Document Management**: Transfer metadata for indexing to a DMS
  - **Data Collection**: Collect extracted information in a central table
  - **Automation**: Provide structured data for subsequent processing steps
- 

## 21.2 General Settings

### Enabled

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

## 21.3 Rules to Export

### Rule Selection

Select the extraction rules whose values should be written to the file. Each selected rule is represented as a separate column in the CSV file.

**Note:** Only rules defined in the profile that extract data can be exported. The order of rules in the selection determines the column order in the CSV file.

---

## 21.4 CSV Settings

### Delimiter

The character that separates individual values (columns). By default, the system's list separator is used.

Delimiter	Description
, (Comma)	International standard
; (Semicolon)	German standard, recommended for German Excel versions
\t (Tab)	For TAB-delimited files

**Tip:** Use ; if you want to open the file with German Excel versions.

### Column Headers

Enable this option to output extraction rule names as column headers in the first row.

#### Example with column headers:

```
InvoiceNumber;Date;Amount  
INV-12345;12/15/2024;1250.00  
INV-12346;12/16/2024;890.50
```

#### Example without column headers:

```
INV-12345;12/15/2024;1250.00  
INV-12346;12/16/2024;890.50
```

### Collection File

Enable this option to collect all extracted data in **one** shared file. New records are appended at the end of the file.

- **Enabled:** All PDFs write to the same CSV file (one row per PDF)
- **Disabled:** Each PDF creates a separate CSV file

**Use Case:** You process multiple invoices daily and want to collect all data in a single overview file.

### Expand Multi-Line Values

When an extraction rule returns multi-line values (e.g., multiple line items of an invoice), you can specify how these are handled:

- **Don't expand (default):** Multi-line text stays in one cell
- **Select rule:** Multi-line value is split into separate CSV rows (with repetition of other column values)

**Example:** An invoice with 3 line items - Without expansion: Item 1^Item 2^Item 3 in one cell - With expansion: 3 separate rows in CSV

---

## 21.5 Character Encoding

Select the character encoding for the output file:

Encoding	Description	Recommended For
<b>ANSI</b>	Windows standard encoding	Older applications
<b>UTF-8</b>	Unicode without BOM	Web, modern applications
<b>UTF-8 with BOM</b>	Unicode with Byte Order Mark	Excel (recommended)
<b>UTF-16 LE/BE</b>	16-bit Unicode	Special applications
<b>ASCII</b>	Standard characters only	Legacy systems

**Recommendation:** Use “UTF-8 with BOM” for best compatibility with Excel and special characters.

---

## 21.6 Storage Location

### Directory

Specify the target directory for the CSV file.

**Note:** It's recommended to use a separate folder for each processing step to ensure clear separation.

### Filename

Set the name for the CSV file.

### Examples:

Input	Result
Export	Export.csv
<TodaysYear4>-<TodaysMonth>-	2024-12-15_Invoices.csv

Input	Result
<TodaysDay>_Invoices	
<FileName>_Data	Invoice123_Data.csv

**For collection file:** Use a fixed name or date placeholder for daily/monthly files.

### Name Collisions

Choose what should happen if a file with the target name already exists:

Option	Description
<b>Overwrite</b>	Existing file is replaced
<b>Append number</b>	Adds a number
<b>Append date</b>	Adds processing date
<b>Append date and time</b>	Adds date and time
<b>Cancel operation</b>	File is not written

**For collection file:** This setting only applies to new files. With collection file enabled, new rows are always appended.

## 21.7 File Date

### Adjust Creation and Modification Date

Optionally, you can change the file date of the CSV file:

Option	Description
<b>Do not change</b>	File automatically receives current date
<b>Creation date of original file</b>	Uses PDF's creation date
<b>Modification date of original file</b>	Uses PDF's modification date
<b>PDF creation date</b>	Date from PDF metadata
<b>Extracted date</b>	A date obtained with an extraction rule
<b>Current date</b>	Sets today's date

## 21.8 Afterwards

### Call External Program

After saving, an external program can be started automatically.

**Program:** Path to executable file

**Parameters:** Command line parameters. Available placeholders: -  
<PathIncludingFilename> - Full path of CSV file - <ParentDirectory> - Path of parent folder - <Filename> - Filename of CSV file

---

## 21.9 Example: Export Invoice Data for Accounting

### Initial Situation

Incoming invoices should be automatically processed. Invoice data (number, date, supplier, amount) should be exported to a CSV file that is imported monthly into accounting software.

### Prerequisites

Extraction rules defined for: - Rule 1: "InvoiceNumber" - Rule 2: "InvoiceDate" - Rule 3: "Supplier" - Rule 4: "GrossAmount"

### Configuration

1. **Enabled:** Yes
2. **Selected Rules:** All four rules
3. **Delimiter:** ;
4. **Column Headers:** Yes
5. **Collection File:** Yes
6. **Character Encoding:** UTF-8 with BOM
7. **Directory:** D:\Accounting\Import
8. **Filename:** Invoices\_<TodaysYear4>-<TodaysMonth>

### Result

All invoices processed in December 2024 are collected in one file:

**File:** D:\Accounting\Import\Invoices\_2024-12.csv

```
InvoiceNumber;InvoiceDate;Supplier;GrossAmount
INV-12345;12/15/2024;Sample Company Inc;1250.00
INV-12346;12/16/2024;Smith Corp;890.50
INV-12347;12/17/2024;Example Ltd;2100.00
```

---

## 21.10 Example: Individual CSV per PDF

### Initial Situation

Each processed invoice should have its own CSV file with extracted data to add as companion file to a document management system.

## Configuration

1. **Enabled:** Yes
2. **Selected Rules:** All relevant rules
3. **Collection File:** No
4. **Directory:** D:\Archive\<>Today'sYear4>\<Today'sMonth>
5. **Filename:** <FileName>

## Result

PDF File	CSV File
Invoice_12345.pdf	D:\Archive\2024\12\Invoice_12345.csv
Invoice_12346.pdf	D:\Archive\2024\12\Invoice_12346.csv

---

## 21.6 Tips and Notes

### Special Characters in Values

If extracted values contain the delimiter (e.g., , in an amount), values are automatically enclosed in quotation marks:

```
"Sample, Company Inc";12/15/2024;1250.00
```

### Empty Values

If an extraction rule returns no value for a specific PDF, an empty field is written:

```
INV-12345;;Sample Company Inc;1250.00
```

(Here the date is missing)

### Column Order

Column order in the CSV file corresponds to the order of selected rules. Change the selection order to adjust column order.

### Excel Import

For trouble-free import into Excel: 1. Use ; as delimiter (for German Excel version) or , (for English version) 2. Choose UTF-8 with BOM as encoding 3. Enable column headers

### Combination with Other Tasks

The “Save Extractions” task combines well with other tasks: 1. **Rename File:** Rename PDF based on extracted data 2. **Copy File:** Copy PDF to archive 3. **Save Extractions:** Export data for import 4. **Send Email:** Send notification with extracted data

---

# 22 Use Template

## Task: Use Template

*Task: Use Template*

### 22.1 Description

The **Use Template** task fills a text template with extracted data and saves the result as a separate file. The placeholders in the template are replaced with actual values from the PDF.

#### Typical Use Cases

- **Cover Letters:** Automatically generated letters with invoice data
  - **Data Export:** Create XML or JSON files with extracted values
  - **Reports:** Generate structured text files for other systems
  - **Index Files:** Metadata files for document management systems
- 

### 22.2 General Settings

#### Enabled

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

### 22.3 Template

#### Template Path

*Template Path*

#### Template File

Specify the path to the template file. The template is a text file (e.g., .txt, .xml, .json) that contains placeholders.

**Supported File Types:** All text-based files

#### Character Encoding

Select the encoding of the template file:

Encoding	Description
<b>ANSI</b>	Windows standard encoding
<b>UTF-8</b>	Unicode (recommended)
<b>UTF-8 with BOM</b>	Unicode with Byte Order Mark

Encoding	Description
UTF-16	16-bit Unicode

---

## 22.4 Replacements

### Replacements

#### *Replacements*

#### **Define Replacement Pairs**

For each placeholder in the template, define a replacement pair:

1. **Search Term:** The text in the template to be replaced (e.g., [INVOICENO])
2. **Extraction Rule:** The rule whose value will be inserted

#### **Example Template:**

```
<?xml version="1.0"?>
<Invoice>
  <Number>[INVOICENO]</Number>
  <Date>[DATE]</Date>
  <Amount>[AMOUNT]</Amount>
  <Customer>[CUSTOMER]</Customer>
</Invoice>
```

**Replacement Pairs:** | Search Term | Extraction Rule | | [INVOICENO]  
| Rule "InvoiceNumber" | | [DATE] | Rule "InvoiceDate" | | [AMOUNT] | Rule "GrossAmount" | |  
[CUSTOMER] | Rule "CustomerName" |

---

## 22.5 Embedding (optional)

### Embed

#### *Embed*

#### **Embed File in PDF**

Enable this option to embed the generated file as an attachment in the PDF.

#### **Embedding Filename**

Specify the name under which the file should be displayed in the PDF.

**Use Case:** You create an XML file with structured invoice data and want to make it directly available in the PDF, similar to ZUGFeRD invoices.

---

## 22.6 Storage Location

### Directory

Specify the target directory for the generated file.

**Note:** It's recommended to use a separate folder for each processing step to ensure clear separation.

### Filename

Set the name for the output file. The file extension should match the content (e.g., .xml, .json, .txt).

### Name Collisions

Choose what should happen if a file with the target name already exists:

Option	Description
<b>Overwrite</b>	Existing file is replaced
<b>Append number</b>	Adds a number
<b>Append date</b>	Adds processing date
<b>Cancel operation</b>	File is not created

---

---

## 22.7 File Date

### Adjust Creation and Modification Date

Optionally, you can change the file date of the generated file.

---

---

## 22.8 Afterwards

### Call External Program

After creating the file, an external program can be started automatically.

**Parameters:** Available placeholders: - <PathIncludingFilename> - Full path of the generated file - <ParentDirectory> - Path of parent folder - <Filename> - Filename of the generated file

---

---

## 22.9 Example: XML File for Accounting System

### Initial Situation

Your accounting system expects an XML import in a specific format.

### Template File (invoice\_template.xml):

```
<?xml version="1.0" encoding="UTF-8"?>
<AccountingImport>
  <Document>
    <DocNo>%%DOCNO%%</DocNo>
    <Date>%%DATE%%</Date>
    <Supplier>%%SUPPLIER%%</Supplier>
    <Net>%%NET%%</Net>
    <Tax>%%TAX%%</Tax>
    <Gross>%%GROSS%%</Gross>
  </Document>
</AccountingImport>
```

### Configuration

1. **Template File:** C:\Templates\invoice\_template.xml
2. **Replacements:** All 6 placeholders linked with corresponding rules
3. **Directory:** D:\Import\Accounting
4. **Filename:** <RuleId:1(DocNo)>.xml

### Result

For an invoice with document number "INV-2024001":

```
<?xml version="1.0" encoding="UTF-8"?>
<AccountingImport>
  <Document>
    <DocNo>INV-2024001</DocNo>
    <Date>12/15/2024</Date>
    <Supplier>Sample Company Inc</Supplier>
    <Net>1050.42</Net>
    <Tax>199.58</Tax>
    <Gross>1250.00</Gross>
  </Document>
</AccountingImport>
```

---

## 22.6 Tips and Notes

### Placeholder Format

Use unique placeholders that won't accidentally appear in regular text: - **Good:** [INVOICENO], %%DATE%%, {{CUSTOMER}} - **Bad:** Date, No (could appear in text)

### Special Characters in XML

If extracted values contain XML special characters (<, >, &), they must be cleaned or escaped in the extraction rules.

### Multiple Replacements

You can define any number of replacement pairs. Each pair is processed independently.

### Empty Values

If an extraction rule returns no value, the placeholder is replaced with an empty string.

---

## 23 Change File Protection

Task: Change File Protection

*Task: Change File Protection*

### 23.1 Description

The **Change File Protection** task allows adding, changing, or removing password protection and permissions for PDF files. You can manage both user passwords (for opening) and permission passwords (for restrictions).

#### Typical Use Cases

- **Add Protection:** Protect confidential documents with a password
  - **Remove Protection:** Remove existing password protection for further processing
  - **Set Permissions:** Restrict printing or editing
  - **Dynamic Passwords:** Generate passwords from extracted data (e.g., customer number)
- 

### 23.2 General Settings

#### Enabled

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

### 23.3 Remove Password

File Protection

*File Protection*

#### Remove Password Protection

Enable this option to completely remove existing password protection. The file can then be opened and edited without a password.

**Prerequisite:** The current password must be stored in the password list (program options) or in the profile.

---

## 23.4 User Password

The user password (also called “open password”) is required to open the PDF.

### Enable Encryption

Enable this option to set a user password.

### Password Type

Type	Description
<b>Static</b>	A fixed password that you enter
<b>Dynamic</b>	The password is obtained from an extraction rule

### Static Password

Enter the desired password. This will be used for all processed PDFs.

### Dynamic Password

Select an extraction rule whose value will be used as the password.

**Example:** The rule “CustomerNumber” extracts “CU-12345”. This is used as the password. The customer can only open the PDF if they know their customer number.

### Exclude Metadata from Encryption

Enable this option so PDF metadata (title, author, etc.) remains readable without a password. This allows indexing by search systems.

---

## 23.5 Permission Password

The permission password (also called “owner password”) protects the permission settings. Without this password, the restrictions cannot be removed.

### Enable Encryption

Enable this option to set a permission password.

### Password Type

Same as user password: Static or Dynamic.

---

## 23.6 Permissions

For each permission, you can specify: - **Leave unchanged** - Keep current setting - **Allow** - Grant permission - **Disallow** - Deny permission

### Available Permissions

Permission	Description
<b>Print (low quality)</b>	Allow printing in low resolution
<b>Print (high quality)</b>	Allow printing in full quality
<b>Modify document</b>	Edit document content
<b>Extract content</b>	Copy text and graphics
<b>Accessibility extraction</b>	Extraction for assistive technologies
<b>Annotations/Forms</b>	Edit annotations and forms
<b>Fill forms</b>	Only fill form fields (not edit)
<b>Assemble document</b>	Insert, rotate, delete pages

---

---

## 23.7 File Date

### Adjust Creation and Modification Date

Optionally, you can adjust the file date after changing the protection.

---

---

## 23.8 Afterwards

### Call External Program

After changing the file protection, an external program can be started automatically.

**Parameters:** Available placeholders: - <PathIncludingFilename> - Full path of the file - <ParentDirectory> - Path of parent folder - <Filename> - Filename

---

---

## 23.9 Example: Protect Documents with Customer Number

### Initial Situation

You send personalized documents to customers. Each document should be protected with the respective customer number as password.

### Prerequisites

- Extraction rule “CustomerNumber” defined

## Configuration

1. **Enabled:** Yes
2. **User Password:** Enabled
3. **Password Type:** Dynamic
4. **Extraction Rule:** CustomerNumber
5. **Permissions:** All set to “Leave unchanged”

## Result

PDF	Extracted Customer Number	Password
Invoice_001.pdf	CU-12345	CU-12345
Invoice_002.pdf	CU-67890	CU-67890

---

## 23.10 Example: Allow Printing, Forbid Editing

### Initial Situation

Documents should be printable but not editable.

### Configuration

1. **Enabled:** Yes
2. **Permission Password:** Enabled, static password
3. **Permissions:**
  - Print (high quality): **Allow**
  - Modify document: **Disallow**
  - Extract content: **Allow**
  - Annotations/Forms: **Disallow**
  - Fill forms: **Allow**
  - Assemble document: **Disallow**

### Result

The document can be opened, printed, and forms can be filled out. Editing and inserting/deleting pages is not possible.

---

## 23.5 Tips and Notes

### Difference Between User and Permission Password

Aspect	User Password	Permission Password
<b>Required for</b>	Opening the PDF	Changing permissions
<b>Without</b>	PDF not readable	PDF readable but restricted

Aspect	User Password	Permission Password
<b>password</b>		
<b>Typical use</b>	Confidentiality	Protection against manipulation

### Combining Both Passwords

You can set both passwords simultaneously: - User password for opening - Permission password for restrictions

### Dynamic Passwords

With dynamic passwords, ensure the extraction rule reliably returns a value. An empty password would provide no protection.

### Communicating Passwords

Don't forget to communicate the password to the recipient, e.g.: - In the email message (separate email recommended) - Via SMS or phone call - In the cover letter

### Encryption Strength

Encryption is performed by default with AES-256, the current industry standard for PDF encryption.

## 24 Compress PDF

Task: Compress PDF

*Task: Compress PDF*

### 24.1 Description

The **Compress PDF** task reduces the file size of PDF documents by optimizing contained images and removing unnecessary elements. This is particularly useful for archiving or emailing large files.

#### Typical Use Cases

- **Email Sending:** Reduce large PDFs for attachments
- **Archiving:** Save storage space for long-term archiving
- **Web Upload:** Optimize files for online portals
- **Scan Optimization:** Reduce oversized scan files

**Important:** This task creates a new file in the configured target folder. The original file remains unchanged. Further tasks contained in the current profile all refer to the original file. The compressed PDF must be further processed with a **separate profile** that monitors the corresponding output folder if needed.

---

## 24.2 General Settings

### Enabled

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

## 24.3 Image Compression

### Image Compression

#### *Image Compression*

### Image Quality

Set the JPEG quality for compressed images (1-100).

Value	Description
90-100	High quality, low compression
70-89	Good balance (recommended)
50-69	Medium quality, significant compression
1-49	Low quality, maximum compression

### Reduce Image Resolution

Enable this option to reduce the resolution of images.

### Maximum DPI

The target resolution for images. Images with higher resolution will be downscaled.

DPI	Recommended for
300	High-quality prints
150-200	Screen display, standard archiving
96-120	Screen viewing only, maximum compression

### Convert to Grayscale

Converts color images to grayscale. This significantly reduces file size but removes color information.

**Use Case:** Black and white documents like invoices or contracts where color is not relevant.

---

## 24.4 Remove Elements

### Remove Content

*Remove Content*

#### **Remove Metadata**

Removes PDF metadata such as author, title, subject, creation date.

**Note:** Only enable this option if the metadata is not needed.

#### **Remove Attachments**

Removes embedded files from the PDF.

#### **Remove Bookmarks**

Removes the table of contents (bookmarks).

#### **Remove Thumbnails**

Removes embedded page thumbnails.

#### **Remove JavaScript Actions**

Removes JavaScript code from the PDF (also increases security).

#### **Remove Logical Structure**

Removes structure information for accessibility.

**Note:** Do not enable if the PDF needs to be accessible.

---

## 24.5 Additional Options

### Behavior

*Behavior*

#### **Skip if Result is Larger**

Enable this option to discard the compressed file if it's larger than the original. In this case, no output file is created.

**Recommendation:** This option should be enabled to avoid unintended file enlargement.

#### **Use Object Streams**

Optimizes internal PDF storage. Can further reduce file size.

---

## 24.6 Storage Location

### Directory

Specify the target directory for the compressed file.

**Note:** It's recommended to use a separate folder for each processing step to ensure clear separation.

### Filename

Set the name for the compressed file.

### Examples:

Input	Result
(empty)	Document.pdf (original name)
<FileName>_compressed	Document_compressed.pdf

### Name Collisions

Choose what should happen if a file with the target name already exists.

---

## 24.7 File Date

### Adjust Creation and Modification Date

Optionally, you can change the file date of the compressed file.

---

## 24.8 Afterwards

### Call External Program

After compression, an external program can be started automatically.

**Parameters:** Available placeholders: - <PathIncludingFilename> - Full path of the compressed file - <ParentDirectory> - Path of parent folder - <Filename> - Filename of the compressed file

---

## 24.9 Example: Optimize Scan Files for Archiving

### Initial Situation

Your scanner creates high-resolution PDFs with 5-10 MB per page. A lower quality is sufficient for archiving.

## Configuration

1. **Enabled:** Yes
2. **Image Quality:** 75
3. **Reduce Image Resolution:** Yes
4. **Maximum DPI:** 150
5. **Remove Thumbnails:** Yes
6. **Skip if Larger:** Yes
7. **Directory:** D:\Archive\Compressed
8. **Filename:** <FileName>

## Result

Original	Compressed	Savings
8.5 MB	1.2 MB	85%

---

## 24.5 Tips and Notes

### Quality vs. File Size

Find the right balance for your use case: - **Archiving:** Higher quality (80-90) for long-term readability - **Email:** Medium quality (60-75) for smaller attachments - **Quick View:** Lower quality (40-60) for temporary files

### Already Compressed PDFs

With already heavily compressed PDFs, recompression can even increase the file size. Enable “Skip if Result is Larger”.

### Text PDFs vs. Image PDFs

- **Text PDFs:** Less compression potential, as text is already efficiently stored
- **Image PDFs (Scans):** Large savings possible through image compression

### Further Processing

The compressed PDF is in the configured target folder. Create a separate profile for further processing if needed.

---

## 25 Embed Attachments

Task: Embed Attachments

*Task: Embed Attachments*

## 25.1 Description

The **Embed Attachments** task inserts external files as attachments into a PDF document. The embedded files can be any type and are visible as attachments in the PDF viewer.

### Typical Use Cases

- **Attach Source Files:** Add Excel spreadsheets or Word documents to an invoice
  - **Archive Bundles:** Combine multiple related files in one PDF
  - **Data Attachments:** Add XML files or raw data to the PDF
  - **Supporting Documents:** Integrate supplementary documents into the main file
- 

## 25.2 General Settings

### Enabled

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

## 25.3 Files to Embed

### Files to Embed

*Files to Embed*

### File List

Add the files that should be embedded in the PDF. For each entry, you can specify:

1. **File Path:** The path to the file (can contain placeholders)
2. **Description:** An optional description displayed in the PDF viewer

### Placeholders in File Path

You can use placeholders to create dynamic filenames:

### Examples:

Input	Result
D:\Data\ <ruleid:1(customer)&gt;.xlsx< td=""><td>D:\Data\Sample Company Inc.xlsx</td></ruleid:1(customer)&gt;.xlsx<>	D:\Data\Sample Company Inc.xlsx
<ParentDirectory>\<FileName>.xml	C:\Incoming\Invoice_001.xml

---

## 25.4 Duplicate Handling

Choose what should happen if an attachment with the same name already exists in the PDF:

Option	Description
<b>Add with numbering</b>	Adds the file with a number (e.g., file (1).pdf)
<b>Overwrite</b>	Replaces the existing attachment
<b>Skip</b>	The file is not embedded

---

---

## 25.5 Storage Location

### Directory

Specify the target directory for the PDF with embedded attachments.

**Note:** It's recommended to use a separate folder for each processing step to ensure clear separation.

### Filename

Set the name for the output PDF.

### Name Collisions

Choose what should happen if a file with the target name already exists.

---

---

## 25.6 File Date

### Adjust Creation and Modification Date

Optionally, you can change the file date of the output file.

---

---

## 25.7 Afterwards

### Call External Program

After embedding, an external program can be started automatically.

**Parameters:** Available placeholders: - <PathIncludingFilename> - Full path of the file - <ParentDirectory> - Path of parent folder - <Filename> - Filename

---

---

## 25.8 Example: Add Excel Raw Data to Invoice

### Initial Situation

You create invoices as PDFs and want to embed the associated Excel data directly in the PDF.

## Configuration

1. **Enabled:** Yes
2. **Files:**
  - Path: D:\Data\  - Description: “Invoice Data”
3. **Duplicate Handling:** Add with numbering
4. **Directory:** D:\Archive\Invoices
5. **Filename:** <FileName>

## Result

The PDF Invoice\_2024001.pdf contains the file 2024001.xlsx as an attachment with the description “Invoice Data”.

---

## 25.9 Example: Embed Multiple Supporting Files

### Initial Situation

Each main document has multiple associated files (attachments, images, supplements) that should all be bundled in the PDF.

### Configuration

1. **Enabled:** Yes
2. **Files:**
  - <ParentDirectory>\<FileName>\_Attachment1.pdf - “Attachment 1”
  - <ParentDirectory>\<FileName>\_Attachment2.pdf - “Attachment 2”
  - <ParentDirectory>\<FileName>\_Data.xlsx - “Source Data”
3. **Duplicate Handling:** Skip (if file doesn’t exist)
4. **Directory:** D:\Archive\Complete

## Result

All existing supporting files are embedded in the PDF. Non-existing files are skipped.

---

## 25.5 Tips and Notes

### Consider File Size

Embedded attachments increase the PDF size accordingly. Keep this in mind with large attachments.

### Supported File Types

Any file type can be embedded: - Office documents (Word, Excel, PowerPoint) - Images (PNG, JPG, TIFF) - PDF files - XML, JSON, CSV - Archives (ZIP) - And all other file types

### **Non-Existing Files**

If a file doesn't exist, it is skipped. Processing does not fail.

### **Dynamic Filenames**

Use placeholders to embed files based on extracted values. This way you can automatically assign customer-related files.

### **Viewing Attachments**

Embedded attachments are accessible in PDF viewers via the paperclip icon or an attachments panel.

---

## 26 Create ZUGFeRD/Factur-X

Task: Create ZUGFeRD

*Task: Create ZUGFeRD*

### 26.1 Description

The **Create ZUGFeRD/Factur-X** task converts a PDF into an electronic invoice according to the ZUGFeRD or Factur-X standard. A pre-existing structured XML file with invoice data is embedded in the PDF and the PDF is converted to PDF/A-3b format.

### **Typical Use Cases**

- **E-Invoice:** Prepare PDF invoices for B2B or B2G requirements
  - **XRechnung:** Create German government invoices according to EU standard
  - **Factur-X:** Create French/European e-invoices
  - **Automation:** Combine invoice PDFs with existing XML data
- 

### 26.2 General Settings

#### **Enabled**

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

## 26.3 ZUGFeRD Version

### ZUGFeRD Configuration

#### ZUGFeRD Configuration

Select the ZUGFeRD version:

Version	Description
<b>ZUGFeRD 2.1</b>	Widely used version
<b>ZUGFeRD 2.3.2</b>	Current version (recommended)

---

## 26.4 Conformance Level

Choose the conformance level based on your requirements:

Level	Description	Usage
<b>Minimum</b>	Minimum requirements	Simple invoices
<b>Basic</b>	Basic invoice data	Standard invoices
<b>Basic WL</b>	Basic without line item details	Invoices without individual items
<b>Comfort</b>	Extended invoice data	Detailed invoices
<b>Extended</b>	Complete invoice data	Complex invoices
<b>EN16931</b>	EU standard for e-invoices	EU-wide exchange
<b>XRechnung</b>	German government standard	Invoices to authorities

---

## 26.5 XML File

### XML Source

#### XML Source

#### Path to XML File

Specify the path to the XML file with invoice data. You can use placeholders.

#### Examples:

Input	Result
D:\XML\ <ruleid:1(invoiceno)&gt;.xml< td=""><td>D:\XML\2024001.xml</td></ruleid:1(invoiceno)&gt;.xml<>	D:\XML\2024001.xml
<ParentDirectory>\<FileName>.xml	C:\Invoices\Invoice_001.xml

#### Search in Parent Folder

Enable this option to automatically search for the XML file in the same folder as the PDF.

## Filename in Parent Folder

When “Search in Parent Folder” is enabled, specify the filename of the XML file. Default: factur-x.xml

---

## 26.6 Storage Location

### Directory

Specify the target directory for the ZUGFeRD file.

**Note:** It’s recommended to use a separate folder for each processing step to ensure clear separation.

### Filename

Set the name for the resulting ZUGFeRD PDF file.

### Name Collisions

Choose what should happen if a file with the target name already exists.

---

## 26.7 File Date

### Adjust Creation and Modification Date

Optionally, you can change the file date of the ZUGFeRD file.

---

## 26.8 Afterwards

### Call External Program

After creating the ZUGFeRD file, an external program can be started automatically.

**Parameters:** Available placeholders: - <PathIncludingFilename> - Full path of the ZUGFeRD file - <ParentDirectory> - Path of parent folder - <Filename> - Filename

---

## 26.9 Example: Invoices for Government (XRechnung)

### Initial Situation

You need to submit invoices to German authorities in XRechnung format. The XML files are created by your accounting software in the same folder as the PDFs.

### Prerequisites

- PDF invoice present
- XML file with XRechnung-compliant data in the same folder

### Configuration

1. **Enabled:** Yes
2. **ZUGFeRD Version:** ZUGFeRD 2.3.2
3. **Conformance Level:** XRechnung
4. **Search in Parent Folder:** Yes
5. **Filename in Parent Folder:** xrechnung.xml
6. **Directory:** D:\Outgoing\Authorities
7. **Filename:** <FileName>

### Result

Input Files	Output
Invoice_2024001.pdf + xrechnung.xml	D:\Outgoing\Authorities\Invoice_2024001.pdf (ZUGFeRD/XRechnung)

## 26.10 Example: Factur-X for French Customers

### Initial Situation

You send invoices to French customers and must comply with the Factur-X format.

### Configuration

1. **Enabled:** Yes
2. **ZUGFeRD Version:** ZUGFeRD 2.3.2
3. **Conformance Level:** EN16931
4. **Path to XML File:** D:\Export\<<RuleId:1(InvoiceNo)>\_facturx.xml
5. **Directory:** D:\Sending\FR
6. **Filename:** Facture\_<RuleId:1(InvoiceNo)>

## 26.5 Tips and Notes

### PDF/A-3b Conversion

The PDF is automatically converted to PDF/A-3b if it's not already in this format. This is a requirement for ZUGFeRD/Factur-X.

### XML Structure

The XML file must conform to the selected ZUGFeRD/Factor-X schema. Invalid XML files will cause processing errors.

### Validation

Validate your ZUGFeRD files after creation with an appropriate validator: - Free online validators - ZUGFeRD SDK - Government portals (for XRechnung)

### Choosing Conformance Level

- **Minimum/Basic:** For simple invoices without special requirements
- **EN16931:** For EU-wide B2B exchange
- **XRechnung:** Required for German government invoices

### XML Sources

The XML file can come from various sources: - ERP system - Accounting software - Manually created - Other automation processes

### Combination with “Save Attachments” Task

If you’re processing an existing ZUGFeRD invoice, you can use “Save Attachments” to extract the XML file.

---

---

## 27 Convert to PDF/A

Task: Convert to PDF/A

*Task: Convert to PDF/A*

### 27.1 Description

The **Convert to PDF/A** task converts a PDF to PDF/A format, which is designed for long-term archiving of documents. PDF/A ensures that the document can be displayed faithfully in the future.

### Typical Use Cases

- **Long-term Archiving:** Make documents archivable for 10+ years
- **Compliance:** Meet legal archiving requirements
- **DMS Import:** Prepare documents for document management systems
- **Government Communication:** Meet requirements for electronic files

---

---

### 27.2 General Settings

**Enabled**

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

---

## 27.3 PDF/A Conformance Level

### PDF/A Configuration

#### *PDF/A Configuration*

Select the desired PDF/A version:

Version	ISO Standard	Description	Usage
<b>PDF/A-1b</b>	ISO 19005-1	Basic conformance, visual representation guaranteed	Maximum compatibility, older systems
<b>PDF/A-2b</b>	ISO 19005-2	Extended, with JPEG2000 and transparencies	Recommended standard for archiving
<b>PDF/A-3b</b>	ISO 19005-3	With arbitrary attachments	ZUGFeRD, Factur-X, e-invoices

**Note:** The program supports the “b” variants (Basic), which ensure reliable visual reproduction. The “a” and “u” variants (with additional structure information or Unicode mapping) are not supported.

#### **Recommendations**

Use Case	Recommended Version
Standard archiving	PDF/A-2b
ZUGFeRD/E-Invoice	PDF/A-3b
Maximum compatibility	PDF/A-1b

---

---

## 27.4 Storage Location

### **Directory**

Specify the target directory for the PDF/A file.

**Note:** It’s recommended to use a separate folder for each processing step to ensure clear separation.

### **Filename**

Set the name for the PDF/A file.

## Examples:

Input	Result
(empty)	Document.pdf (original name)
<FileName>_PDF/A	Document_PDF/A.pdf

## Name Collisions

Choose what should happen if a file with the target name already exists.

---

## 27.5 File Date

### Adjust Creation and Modification Date

Optionally, you can change the file date of the PDF/A file.

---

## 27.6 Afterwards

### Call External Program

After conversion, an external program can be started automatically.

**Parameters:** Available placeholders: - <PathIncludingFilename> - Full path of the PDF/A file - <ParentDirectory> - Path of parent folder - <Filename> - Filename

---

## 27.7 Example: Invoice Archiving

### Initial Situation

Incoming invoices should be stored in a long-term stable format for the legally required archiving period.

### Configuration

1. **Enabled:** Yes
2. **PDF/A Version:** PDF/A-2b
3. **Directory:** D:\Archive\PDF-A\<TodaysYear4>\<TodaysMonth>
4. **Filename:** <FileName>

### Result

Original	Converted
C:\Incoming\Invoice_001.pdf	D:\Archive\PDF-A\2024\12\Invoice_001.pdf (PDF/A-2b)

---

## 27.8 Example: Preparation for DMS

### Initial Situation

Your document management system requires PDF/A-1b for all incoming documents.

### Configuration

1. **Enabled:** Yes
  2. **PDF/A Version:** PDF/A-1b
  3. **Directory:** \\Server\DMS\Import
  4. **Filename:** <RuleId:1(DocumentNo)>\_<TodaysYear4><TodaysMonth><TodaysDay>
- 

## 27.4 Tips and Notes

### What Happens During Conversion?

PDF/A conversion may include: - Embedding all used fonts - Removing JavaScript and multimedia - Converting transparencies - Standardizing color spaces - Adding XMP metadata

### Non-Convertible Elements

The following elements may cause problems during conversion: - Non-embeddable fonts (will be substituted) - Encryption (will be removed) - JavaScript (will be removed) - Embedded multimedia content (will be removed)

### File Size

PDF/A files may be larger than originals because all fonts must be embedded.

### PDF/A-1 vs. PDF/A-2/3

Aspect	PDF/A-1	PDF/A-2/3
Base	PDF 1.4	PDF 1.7
JPEG2000	No	Yes
Transparencies	No	Yes
Attachments	No	Only PDF/A-3
Compression	Less efficient	More efficient options

### Validation

Validate your PDF/A files with an appropriate tool: - Adobe Acrobat (Preflight) - veraPDF (Open Source) - Online validators

### Already Converted PDFs

If a PDF is already in PDF/A format, it will be converted to the selected version or copied unchanged during conversion.

---

## 28 Remove Pages

Task: Remove Pages

*Task: Remove Pages*

### 28.1 Description

The **Remove Pages** task deletes specific pages from a PDF document. The pages to remove can be determined by various criteria: text content, character count, page numbers, or page ranges.

#### Typical Use Cases

- **Remove Cover Pages:** Remove automatically generated cover pages from scans
  - **Remove Blank Pages:** Delete pages with little or no content
  - **Remove Separator Pages:** Remove pages with specific marker text
  - **Remove Advertising:** Delete pages with promotional content from documents
  - **Standardization:** Keep only specific pages (e.g., odd pages)
- 

### 28.2 General Settings

#### Enabled

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

### 28.3 Removal Criteria

#### Removal Criteria

##### *Removal Criteria*

You can combine multiple criteria. Pages that match at least one of the enabled criteria will be removed.

#### **Remove Pages with Specific Text**

Removes all pages that contain the specified text.

**Examples:** - SEPARATOR - Removes all pages with the word “SEPARATOR” - Advertising - Removes all pages containing “Advertising”

## Remove Pages Without Specific Text

Removes all pages that do **not** contain the specified text. Useful when you only want to keep pages with specific content.

**Example:** Invoice - Keeps only pages containing the word “Invoice”.

## Remove Pages with Fewer Than X Characters

Enable this option and specify the minimum character count. Pages with fewer characters are removed.

**Use Case:** Automatically remove blank or nearly blank pages (e.g., only with page number).

Setting	Effect
10 characters	Removes almost empty pages
50 characters	Removes pages with very little text
100 characters	Removes pages with little content

## Remove Odd/Even Pages

Option	Description
<b>None</b>	No page number-based removal
<b>Remove odd pages</b>	Pages 1, 3, 5, 7... are removed
<b>Remove even pages</b>	Pages 2, 4, 6, 8... are removed

**Use Case:** With duplex-scanned documents, blank back pages (e.g., all even pages) can be removed.

## Remove Page Ranges

Specify the pages to remove as a range.

**Syntax:** | Input | Effect | |———|———| | 1 | Removes page 1 | | 1-3 | Removes pages 1, 2, and 3  
| | 1, 5, 7-9 | Removes pages 1, 5, 7, 8, and 9 | | last | Removes the last page | | last-2 |  
Removes the last 2 pages |

### Examples:

Input	For a 10-page PDF
1	Page 1 is removed
1-3, last	Pages 1, 2, 3, and 10 are removed
last-3	Pages 8, 9, and 10 are removed

---

## 28.4 Storage Location

### Directory

Specify the target directory for the processed PDF file.

**Note:** It's recommended to use a separate folder for each processing step to ensure clear separation.

### Filename

Set the name for the processed file.

### Examples:

Input	Result
(empty)	Document.pdf (original name)
<FileName>_cleaned	Document_cleaned.pdf

### Name Collisions

Choose what should happen if a file with the target name already exists.

---

## 28.5 File Date

### Adjust Creation and Modification Date

Optionally, you can change the file date of the processed file.

---

## 28.6 Afterwards

### Call External Program

After removing pages, an external program can be started automatically.

**Parameters:** Available placeholders: - <PathIncludingFilename> - Full path of the processed file - <ParentDirectory> - Path of parent folder - <Filename> - Filename

---

## 28.7 Example: Remove Cover Pages from Scans

### Initial Situation

A scanner automatically adds a cover page with device information that should be removed before archiving.

### Configuration

1. **Enabled:** Yes
2. **Remove Page Ranges:** 1

3. **Directory:** D:\Archive\Cleaned
4. **Filename:** <FileName>

### Result

Original	Processed
Scan_001.pdf (5 pages)	Scan_001.pdf (4 pages, without cover page)

---

---

## 28.8 Example: Remove Blank Pages from Scans

### Initial Situation

Duplex-scanned documents contain blank back pages that should be removed.

### Configuration

1. **Enabled:** Yes
2. **Remove Pages with Fewer Than X Characters:** Yes, 20 characters
3. **Directory:** D:\Documents\Cleaned
4. **Filename:** <FileName>

### Result

All pages with fewer than 20 characters are automatically removed.

---

---

## 28.9 Example: Remove Separator Pages from Batch Scan

### Initial Situation

A batch scan contains separator pages with the text “— SEPARATOR —” that should be removed after splitting.

### Configuration

1. **Enabled:** Yes
2. **Remove Pages with Text:** --- SEPARATOR ---
3. **Directory:** D:\Documents\Cleaned
4. **Filename:** <FileName>

### Result

All pages containing the separator text are removed.

---

---

## 28.5 Tips and Notes

### Combining Criteria

You can enable multiple criteria simultaneously. A page is removed if it matches **at least one** of the criteria.

**Example:** “Remove pages with text ‘Advertising’” + “Remove pages with fewer than 50 characters” → Removes both advertising pages and nearly empty pages.

### Caution with “Remove Pages Without Text”

The option “Remove pages without specific text” can unintentionally delete many pages. Test this setting carefully with sample files.

### At Least One Page

The program ensures that at least one page remains in the document. If all pages would match the removal criteria, the first page is kept.

### Order in Task List

Typically place “Remove Pages” before other tasks like Copy or Send Email, so subsequent tasks work with the cleaned document.

### Backup Copy

Create a backup copy before removing pages by placing a “Copy File” task before the “Remove Pages” task.

---

## 29 Add Stamp/Watermark

Task: Stamp/Watermark

*Task: Stamp/Watermark*

### 29.1 Description

The **Add Stamp/Watermark** task adds stamps or watermarks to PDF documents. You can use predefined standard stamps or define custom text.

#### Typical Use Cases

- **Document Status:** Mark documents as “Draft”, “Approved”, or “Confidential”
- **Paid Stamp:** Mark incoming invoices with “PAID” after payment
- **Copyright Protection:** Watermarks with company name on documents
- **Copy Marking:** Mark printed copies as such
- **Workflow Status:** Add processing status to documents

---

## 29.2 General Settings

### Enabled

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

## 29.3 Annotation Type

### Annotation Type

#### *Annotation Type*

Choose between two annotation types:

Type	Description	Application
<b>Stamp</b>	Rectangular text with optional border	Paid stamps, status markers
<b>Watermark</b>	Semi-transparent, diagonal text	Copyright protection, draft marking

---

## 29.4 Stamp Settings

### Stamp Settings

#### *Stamp Settings*

### Standard Text or Custom Text

You can choose between predefined standard texts and custom text.

### Standard Texts

Icon	Description
<b>Draft</b>	Draft
<b>Approved</b>	Approved
<b>Experimental</b>	Experimental
<b>NotApproved</b>	Not Approved
<b>AsIs</b>	As Is
<b>Expired</b>	Expired
<b>NotForPublicRelease</b>	Not For Public Release
<b>Confidential</b>	Confidential
<b>Final</b>	Final
<b>Sold</b>	Sold

Icon	Description
<b>Departmental</b>	Departmental
<b>ForComment</b>	For Comment
<b>TopSecret</b>	Top Secret
<b>ForPublicRelease</b>	For Public Release

### Custom Text

Enter your own text for the stamp. Placeholders are supported.

**Examples:** - PAID - PAID on <TodaysDate> - Reviewed by <RuleId:1(Reviewer)> - Copy - <TodaysYear4>-<TodaysMonth>-<TodaysDay>

### Position

Select the position of the stamp on the page:

Position	Description
<b>Top left</b>	Left upper corner
<b>Top center</b>	Top edge, centered
<b>Top right</b>	Right upper corner (Default)
<b>Middle left</b>	Left edge, vertically centered
<b>Middle center</b>	Page center
<b>Middle right</b>	Right edge, vertically centered
<b>Bottom left</b>	Left lower corner
<b>Bottom center</b>	Bottom edge, centered
<b>Bottom right</b>	Right lower corner
<b>Custom</b>	Free positioning in percent

### Custom Position

When selecting “Custom”, you can specify X and Y coordinates in percent: - **X Position:** 0% (left) to 100% (right) - **Y Position:** 0% (top) to 100% (bottom)

### Text Color

Set the color of the stamp text: - **RGB Values:** Red, Green, Blue (0-255 each) - **Default:** Dark red (139, 0, 0)

## 29.5 Watermark Settings

### Watermark Settings

*Watermark Settings*

## Text

Enter the text for the watermark. Placeholders are supported.

**Examples:** - DRAFT (Default) - CONFIDENTIAL - © Sample Company Inc - COPY

## Font

Setting	Description	Default
<b>Font Family</b>	Name of the font	Arial
<b>Font Size</b>	Size in points	72
<b>Bold</b>	Use bold font	Yes

**Available Font Sizes:** 24, 36, 48, 60, 72, 84, 96, 120, 144

## Color and Transparency

Setting	Description	Default
<b>Text Color (RGB)</b>	Red, Green, Blue (0-255 each)	Light gray (200, 200, 200)
<b>Opacity</b>	0.0 (transparent) to 1.0 (opaque)	0.3 (30%)

**Recommendation:** A low opacity (0.2-0.4) ensures the watermark doesn't cover the document content.

## Rotation

Select the rotation angle of the watermark:

Angle	Description
<b>0°</b>	Horizontal
<b>-45°</b>	Diagonal from bottom left to top right (Default)
<b>-90°</b>	Vertical (readable from bottom to top)
<b>45°</b>	Diagonal from top left to bottom right
<b>90°</b>	Vertical (readable from top to bottom)

**Available Angles:** 0°, ±15°, ±30°, ±45°, ±60°, ±90°

## Position

Position	Description
<b>Centered</b>	One watermark in the page center
<b>Tiled</b>	Multiple watermarks distributed across the entire page

## Tile Spacing

With tiled arrangement, you can set the spacing between watermarks: - **Horizontal Spacing (X):** Default 200 points - **Vertical Spacing (Y):** Default 200 points

---

## 29.6 Page Selection

### Page Selection

#### *Page Selection*

Determine on which pages the stamp or watermark should appear:

Selection	Description
<b>All pages</b>	On every page of the document
<b>First page</b>	Only on the first page
<b>Last page</b>	Only on the last page
<b>Page range</b>	Specific pages or ranges
<b>Pages with text</b>	Pages containing specific text
<b>Pages without text</b>	Pages not containing specific text

#### **Page Range**

Specify pages as a range:

Input	Effect
1	Only page 1
1-3	Pages 1, 2, and 3
1, 3, 5	Pages 1, 3, and 5
1-3, 5, 7-10	Pages 1, 2, 3, 5, 7, 8, 9, and 10
last	Only the last page
last-2	The last 2 pages
8-last	From page 8 to the end

#### **Text-Based Selection**

For “Pages with text” or “Pages without text”, enter the search text. Extended filter syntax is supported: - **AND/OR Logic:** Invoice<AND>2024 - **Regular Expressions:** <BeginOfRegex>INV-\d{5}<EndOfRegex> - **Number Ranges:** <NumberRange{1,999}>

---

## 29.7 Embedding Mode

Select how the annotation should be embedded in the PDF:

Mode	Description	Application
<b>As Annotation</b>	Editable, can be removed later	Temporary markings
<b>Permanent</b>	Permanently integrated into page	Final markings (Default)

Mode	Description	Application
	content	

**Recommendation:** For legally relevant markings (e.g., “PAID”), use “Permanent” mode so the marking cannot be removed.

## 29.8 Storage Location

### Directory

Specify the target directory for the processed PDF file.

**Note:** It’s recommended to use a separate folder for each processing step to ensure clear separation.

### Filename

Set the name for the processed file.

### Examples:

Input	Result
(empty)	Document.pdf (original name)
<FileName>_stamped	Document_stamped.pdf

### Name Collisions

Choose what should happen if a file with the target name already exists.

## 29.9 File Date

### Adjust Creation and Modification Date

Optionally, you can change the file date of the processed file.

## 29.10 Afterwards

### Call External Program

After adding the stamp/watermark, an external program can be started automatically.

**Parameters:** Available placeholders: - <PathIncludingFilename> - Full path of the processed file - <ParentDirectory> - Path of parent folder - <Filename> - Filename

## 29.11 Example: Paid Stamp on Invoices

### Initial Situation

Incoming invoices should be marked with a “PAID” stamp after payment and archived.

### Configuration

1. **Enabled:** Yes
2. **Annotation Type:** Stamp
3. **Standard Icon:** No (custom text)
4. **Text:** PAID on <TodaysDate>
5. **Position:** Top right
6. **Text Color:** Dark red (139, 0, 0)
7. **Page Selection:** First page
8. **Embedding Mode:** Permanent
9. **Directory:** D:\Archive\Invoices\Paid
10. **Filename:** <FileName>

### Result

Original	Processed
Invoice_001.pdf	Invoice_001.pdf (with “PAID on 12/15/2024” on page 1)

---

## 29.12 Example: Watermark for Confidential Documents

### Initial Situation

Confidential documents should be marked with a “CONFIDENTIAL” watermark appearing on all pages.

### Configuration

1. **Enabled:** Yes
2. **Annotation Type:** Watermark
3. **Text:** CONFIDENTIAL
4. **Font Size:** 72
5. **Bold:** Yes
6. **Text Color:** Red (255, 0, 0)
7. **Opacity:** 0.2 (20%)
8. **Rotation:** -45°
9. **Position:** Centered
10. **Page Selection:** All pages
11. **Embedding Mode:** Permanent

12. **Directory:** D:\Documents\Confidential

### **Result**

All pages of the document show a diagonal, semi-transparent “CONFIDENTIAL” watermark in the center.

---

## 29.13 Example: Tiled Copyright Watermark

### **Initial Situation**

Documents should be protected with a tiled copyright watermark.

### **Configuration**

1. **Enabled:** Yes
2. **Annotation Type:** Watermark
3. **Text:** © Sample Company Inc
4. **Font Size:** 36
5. **Opacity:** 0.15 (15%)
6. **Rotation:** -30°
7. **Position:** Tiled
8. **Horizontal Spacing:** 250
9. **Vertical Spacing:** 200
10. **Page Selection:** All pages

### **Result**

The copyright watermark appears multiple times distributed across each page, so it cannot be easily cut out.

---

## 29.6 Tips and Notes

### **Placeholders in Text**

Both stamps and watermarks support placeholders: - <TodaysDate> - Current date - <TodaysYear4>, <TodaysMonth>, <TodaysDay> - Date parts - <FileName> - Filename - <RuleId:N(RuleName)> - Extracted data

### **Text Color for Stamps**

For good visibility and professional appearance: - **Dark red (139, 0, 0):** Classic for “PAID” or status stamps - **Dark blue (0, 0, 139):** Professional for approvals - **Dark green (0, 100, 0):** Positive for “Approved”

### **Watermark Opacity**

Opacity	Effect
0.1 - 0.2	Very subtle, text remains easily readable
0.2 - 0.4	Clearly visible, slight readability impact
0.4 - 0.6	Dominant, may impair readability

### Permanent vs. As Annotation

- **Permanent:** More secure, as it cannot be removed. Ideal for legal or final markings.
- **As Annotation:** More flexible, can be removed if needed. Useful for temporary markings or review processes.

### Combination with Other Tasks

Typically place “Add Stamp/Watermark” before tasks like “Send Email” or “Copy File”, so subsequent tasks use the stamped document.

---

## 30 Insert Barcode

Task: Insert Barcode

*Task: Insert Barcode*

### 30.1 Description

The **Insert Barcode** task adds QR codes or barcodes to PDF documents. The barcode content can be dynamically generated from extracted data or placeholders.

### Typical Use Cases

- **Document Identification:** Unique ID as QR code for quick retrieval
  - **Archive Reference:** Barcode with archive number for DMS integration
  - **Tracking:** Shipment numbers or transaction numbers as scannable code
  - **Automation:** Barcode for subsequent automatic processing
  - **Inventory:** Article or document numbers as scannable code
- 

### 30.2 General Settings

#### Enabled

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

## 30.3 Barcode Format

### Barcode Settings

#### *Barcode Settings*

Select the desired barcode format:

#### **2D Formats (two-dimensional)**

Format	Description	Capacity	Application
<b>QR Code</b>	Square 2D code	Up to 4,296 characters	URLs, text, contact data
<b>DataMatrix</b>	Compact 2D code	Up to 2,335 characters	Industrial applications
<b>PDF417</b>	Stacked 2D code	Up to 1,850 characters	ID cards, transportation
<b>Aztec</b>	Compact 2D code	Up to 3,832 characters	Transport tickets

#### **1D Formats (one-dimensional)**

Format	Description	Character Set	Application
<b>Code 128</b>	Alphanumeric	A-Z, a-z, 0-9, special characters	General, logistics
<b>Code 39</b>	Alphanumeric	A-Z, 0-9, some special characters	Industry, military
<b>EAN-13</b>	13 digits	Digits only	Retail (Europe)
<b>EAN-8</b>	8 digits	Digits only	Small products
<b>UPC-A</b>	12 digits	Digits only	Retail (USA)
<b>UPC-E</b>	8 digits	Digits only	Small products (USA)

#### **Format Recommendations**

Use Case	Recommended Format
URLs or longer texts	QR Code
Short alphanumeric codes	Code 128
Numbers only	EAN-13 or Code 128
Maximum density	DataMatrix
High error tolerance	QR Code, Aztec

---

## 30.4 Barcode Content

### **Content**

Enter the content to be encoded in the barcode. Placeholders are supported.

### Examples:

Input	Result
https://company.com/doc/<FileName> <RuleId:1(InvoiceNo)> DOC- <TodaysYear4><TodaysMonth><TodaysDa y>-<FileName> INV:<RuleId:1(CustomerNo)>:<RuleId: 2(Amount)>	https://company.com/doc/Invoice_001.pdf INV-2024-12345 DOC-20241215-Invoice_001 INV:CU-123:1,234.56

### Placeholders

All standard placeholders can be used: - <FileName> - Filename without extension - <TodaysDate> - Current date - <TodaysYear4>, <TodaysMonth>, <TodaysDay> - Date parts - <RuleId:N(RuleName)> - Extracted data from rules - <FileCreatedDate>, <FileModifiedDate> - File dates

### Character Restrictions

Note the character restrictions of different formats:

Format	Allowed Characters
<b>QR Code</b>	All characters (Unicode)
<b>Code 128</b>	ASCII (0-127)
<b>Code 39</b>	A-Z, 0-9, -, ., \$, /, +, %, space
<b>EAN/UPC</b>	Digits only (fixed length)

---

## 30.5 Position

Select the position of the barcode on the page:

Position	Description
<b>Top left</b>	Left upper corner
<b>Top center</b>	Top edge, centered
<b>Top right</b>	Right upper corner
<b>Middle left</b>	Left edge, vertically centered
<b>Middle center</b>	Page center
<b>Middle right</b>	Right edge, vertically centered
<b>Bottom left</b>	Left lower corner
<b>Bottom center</b>	Bottom edge, centered

Position	Description
<b>Bottom right</b>	Right lower corner (Default)
<b>Custom</b>	Free positioning in percent

### Custom Position

When selecting “Custom”, you can specify X and Y coordinates in percent: - **X Position:** 0% (left) to 100% (right) - **Y Position:** 0% (top) to 100% (bottom)

## 30.6 Size

### Size

*Size*

### Width and Height

Set the size of the barcode in points:

Setting	Default	Description
<b>Width</b>	100 pt	Width of the barcode
<b>Height</b>	100 pt	Height of the barcode

**Note:** For square 2D formats (QR Code, DataMatrix, Aztec), the width is automatically used as height to achieve a square result.

### Size Recommendations

Format	Recommended Size
<b>QR Code</b>	80-150 pt (square)
<b>DataMatrix</b>	60-120 pt (square)
<b>Code 128</b>	150-250 pt wide, 50-80 pt high
<b>EAN-13</b>	100-150 pt wide, 50-70 pt high

### Margin

The distance from the page edge in points (Default: 20 pt). This value determines the minimum distance between barcode and page edge.

## 30.7 Page Selection

### Page Selection

*Page Selection*

Determine on which pages the barcode should be inserted:

Selection	Description
<b>All pages</b>	On every page of the document
<b>First page</b>	Only on the first page
<b>Last page</b>	Only on the last page
<b>Page range</b>	Specific pages or ranges
<b>Pages with text</b>	Pages containing specific text
<b>Pages without text</b>	Pages not containing specific text

### Page Range

Specify pages as a range:

Input	Effect
1	Only page 1
1-3	Pages 1, 2, and 3
1, 3, 5	Pages 1, 3, and 5
last	Only the last page
last-2	The last 2 pages

## 30.8 Storage Location

### Directory

Specify the target directory for the processed PDF file.

**Note:** It's recommended to use a separate folder for each processing step to ensure clear separation.

### Filename

Set the name for the processed file.

### Examples:

Input	Result
(empty)	Document.pdf (original name)
<FileName>_barcode	Document_barcode.pdf

### Name Collisions

Choose what should happen if a file with the target name already exists.

## 30.9 File Date

### Adjust Creation and Modification Date

Optionally, you can change the file date of the processed file.

---

## 30.10 Afterwards

### Call External Program

After inserting the barcode, an external program can be started automatically.

**Parameters:** Available placeholders: - <PathIncludingFilename> - Full path of the processed file - <ParentDirectory> - Path of parent folder - <Filename> - Filename

---

## 30.11 Example: QR Code with Archive Link

### Initial Situation

Archived documents should receive a QR code that links directly to the document in the company intranet.

### Configuration

1. **Enabled:** Yes
2. **Barcode Format:** QR Code
3. **Content:** `https://archive.company.com/doc/<RuleId:1(ArchiveNo)>`
4. **Position:** Bottom right
5. **Width:** 80
6. **Margin:** 15
7. **Page Selection:** First page
8. **Directory:** `D:\Archive\With-QR`
9. **Filename:** `<FileName>`

### Result

Original	Processed
Invoice_001.pdf	Invoice_001.pdf (with QR code to <code>https://archive.company.com/doc/A-2024-001234</code> on page 1)

---

## 30.12 Example: Barcode with Invoice Number

### Initial Situation

Incoming invoices should receive a Code 128 barcode with the invoice number to enable scanning in the accounting system.

### Configuration

1. **Enabled:** Yes
2. **Barcode Format:** Code 128
3. **Content:** <RuleId:1(InvoiceNo)>
4. **Position:** Top right
5. **Width:** 180
6. **Height:** 50
7. **Margin:** 20
8. **Page Selection:** First page
9. **Directory:** D:\Invoices\Scanned
10. **Filename:** <FileName>

### Result

The first page contains a scannable barcode with the invoice number (e.g., “INV-2024-12345”).

---

## 30.13 Example: Tracking Code on All Pages

### Initial Situation

Multi-page contracts should have a unique tracking code on each page.

### Configuration

1. **Enabled:** Yes
2. **Barcode Format:** QR Code
3. **Content:** CONTRACT:<RuleId:1(ContractNo)>:P<PageNumber>
4. **Position:** Bottom left
5. **Width:** 60
6. **Margin:** 10
7. **Page Selection:** All pages
8. **Directory:** D:\Contracts\Tracking

### Result

Each page contains a QR code with the contract number and page number (e.g., “CONTRACT:C-2024-001:P1”, “CONTRACT:C-2024-001:P2”, etc.).

---

## 30.6 Tips and Notes

### Ensuring Scannability

For optimal scannability: - **Size:** QR codes should be at least 60-80 pt - **Contrast:** Black barcode on white background - **Margin:** Sufficient white space around the barcode (quiet zone) - **Position:** Don't place over existing text or in the fold zone

### Choose Format by Data Volume

Data Volume	Recommendation
Short number sequence (< 20 characters)	Code 128 or QR Code
Medium text length (20-100 characters)	QR Code
Longer texts or URLs	QR Code, PDF417
Binary data	DataMatrix

### EAN/UPC Check Digit

For EAN and UPC codes, the check digit is calculated automatically. Enter only the payload data: - **EAN-13:** 12 digits (check digit is calculated) - **EAN-8:** 7 digits (check digit is calculated) - **UPC-A:** 11 digits (check digit is calculated)

### Combination with Other Tasks

Typical workflow order: 1. Extract data (extraction rules) 2. Insert barcode (with extracted data) 3. Convert to PDF/A (for archiving) 4. Send email or copy

### Error Correction in QR Codes

QR codes have built-in error correction. Even with slight damage, they can still be scanned. For critical applications, use larger QR codes.

### Test Barcode Content

Test the generated barcode before production use: 1. Create a test PDF with the configured barcode 2. Scan the barcode with a smartphone or scanner 3. Verify the expected content is correctly recognized

---

## 31 Replace Content

Task: Replace Content

*Task: Replace Content*

## 31.1 Description

The **Replace Content** task enables privacy-compliant replacement of text content in PDF documents. The original text is actually removed from the document (true redaction) and replaced with a new value.

### Typical Use Cases

- **Anonymization:** Replace personal data (names, addresses, customer numbers) with placeholders
- **Data Protection:** Make sensitive information unrecognizable for GDPR compliance
- **Standardization:** Replace dynamic content with uniform values
- **Numbering:** Replace document numbers with sequential numbers
- **Archiving:** Replace original reference numbers with archive IDs

**Important:** This task creates a new file in the configured target folder. The original file remains unchanged.

---

## 31.2 General Settings

### Enabled

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

## 31.3 Configure Replacements

### Replacements

#### *Replacements*

Replacements are managed through a list. Each replacement consists of:

1. **Placeholder (Extraction Rule):** Defines which area in the PDF should be replaced
2. **Replacement Source:** Determines where the new value comes from
3. **Scope:** Specifies how many occurrences are replaced

### Add Replacements

Click **Add** to configure a new replacement.

### Edit Replacements

Select a replacement from the list and click **Edit** to adjust the configuration.

### Remove Replacements

Select a replacement and click **Remove** to delete it.

### **Change Order**

Use the arrow buttons to adjust the order of replacements. Replacements are processed from top to bottom.

---

## 31.4 Placeholder (Extraction Rule)

The placeholder determines the area in the PDF whose content should be replaced. The profile's configured data extraction rules are used.

### **Prerequisites**

- The extraction rule must be configured in the profile under "Data Extraction"
- The rule must have a defined area (position or keyword-based)

### **Fallback Rules**

If multiple rules with the same name exist (fallback configuration), the first successful rule is used.

---

## 31.5 Replacement Sources

Select where the new value for the replacement comes from:

### **Fixed Text**

An unchangeable text that is always inserted the same way.

Examples	Application
ANONYMIZED	Data protection marking
***REMOVED***	Mark sensitive data
(empty)	Completely remove text
[COMPANY XY]	Uniform company designation

### **Current Date/Time**

The current date and/or time, formatted according to your specifications.

Format	Example
MM/dd/yyyy	01/12/2026
yyyy-MM-dd	2026-01-12
HH:mm:ss	14:30:45

Format	Example
MM/dd/yyyy HH:mm	01/12/2026 14:30
dddd, MMMM d, yyyy	Monday, January 12, 2026

**UTC Time:** Optionally, UTC time can be used instead of local time.

### Sequential Number

An automatically incremented number saved per profile and rule.

Setting	Description
<b>Start Value</b>	The first number (e.g., 1 or 1000)
<b>Increment</b>	Increase per document (e.g., 1, 10, 100)
<b>Format</b>	Output format (e.g., “00000000” for 8 digits with leading zeros)
<b>Delete After Days</b>	Automatic reset after X days (optional)

**Example:** Start value 1000, increment 1, format “00000000” → “00001000”, “00001001”, “00001002”, ...

### Random Number

A randomly generated number within a defined range.

Setting	Description
<b>Number of Digits</b>	Total length of the number (with leading zeros)
<b>Minimum</b>	Smallest possible value (optional)
<b>Maximum</b>	Largest possible value (optional)

**Example:** 8 digits, Min 0, Max 99999999 → “00123456”, “87654321”, ...

### Random Value from List

A randomly selected value from a text file.

Setting	Description
<b>List Path</b>	Path to the text file (one value per line)

#### Example List File:

```
John Smith
Jane Doe
Robert Johnson
Lisa Williams
```

Values are randomly selected. Previously used values are stored internally to minimize repetitions.

## CSV Lookup

A value is looked up from a configured DynamicQueryList (CSV file).

Setting	Description
---------	-------------

<b>List Name</b>	Name of the DynamicQueryList (from program options)
------------------	---

**Process:** 1. The extracted value (from the placeholder rule) is used as search term 2. The corresponding replacement value is found in the CSV list 3. The replacement value is inserted

**Example:** Extracted value “CU-123” → CSV lookup → “Anonymous Customer A”

---

## 31.6 Scope

The scope determines how many occurrences in the document are replaced:

Scope	Description	Use Case
<b>Single Occurrence</b>	Only the occurrence found by the rule	Standard for individual data fields
<b>All Pages (Same Position)</b>	The text at the same position on all pages	Headers/footers with recurring content
<b>All Occurrences in Document</b>	Every occurrence of the text throughout the document	Consistent replacement (e.g., company name)

### Single Occurrence (Default)

Replaces only the one location found by the extraction rule. Ideal for: - Individual data fields (customer number, invoice number) - One-time text locations

### All Pages (Same Position)

Finds the position on the first page and replaces the text at exactly this position on all pages. Ideal for: - Headers with document number - Footers with page number or company name - Recurring elements at fixed position

### All Occurrences in Document

Searches the entire document for all occurrences and replaces every match. Ideal for: - Company names that appear multiple times - Person names in contracts - Terms that should be consistently replaced

---

## 31.7 Replace Images with Placeholder

In addition to text replacement, all images in the PDF document can be replaced with single-color placeholders. This feature is particularly useful for fully anonymizing documents where photos, logos, or signatures need to be removed alongside text content.

### Activation

Check the **Replace all images with placeholder** checkbox to enable this feature.

### Placeholder Color

Use the color picker to customize the color of the placeholder rectangle. The default is a light gray. The placeholder image automatically receives a subtle, slightly darker border.

### Size Filter

Use the **Min. width**, **Max. width**, **Min. height**, and **Max. height** fields to specify which images should be replaced based on their size (in PDF points). This allows you to exclude small logos or symbols from replacement.

Setting	Default	Description
Min. width	0	Only replace images at least this wide
Max. width	10000	Only replace images up to this wide
Min. height	0	Only replace images at least this tall
Max. height	10000	Only replace images up to this tall

### Notes

- Each image is replaced with a single-color rectangle at the original size and position
- Image replacement can be used without text replacements (just check the box, no replacement entries required)
- Image replacement is performed after text replacement
- Use the size filter to exclude small logos or symbols from replacement

---

## 31.8 Preview

Use the **Preview** button to test replacements on a sample file.

### Preview Window

The preview window shows a comparison: - **Left:** Original PDF - **Right:** PDF with applied replacements

### Select Sample File

Choose one of the configured sample files from the dropdown menu.

## Preview Notes

- Preview does not save any changes
  - Sequential numbers are simulated (not saved)
  - Random values are regenerated
- 

## 31.9 Storage Location

### Directory

Specify the target directory for the processed file.

**Note:** It's recommended to use a separate folder for each processing step.

### Filename

Set the name for the processed file.

### Examples:

Input	Result
(empty)	Document.pdf (original name)
<FileName>_anon	Document_anon.pdf
<RuleId:1(ArchiveNo)>	A-2024-001234.pdf

### Name Collisions

Choose what should happen if a file with the target name already exists.

---

## 31.10 File Date

### Adjust Creation and Modification Date

Optionally, you can change the file date of the processed file.

---

## 31.11 Afterwards

### Call External Program

After replacement, an external program can be started automatically.

**Parameters:** Available placeholders: - <PathIncludingFilename> - Full path of the processed file - <ParentDirectory> - Path of parent folder - <Filename> - Filename

---

## 31.12 Example: Anonymize Customer Data

### Initial Situation

Invoices contain customer names and customer numbers that should be anonymized for archiving.

### Configuration

- 1. Create Extraction Rules:**
  - Rule “CustomerName” (position or keyword-based)
  - Rule “CustomerNumber” (position or keyword-based)
- 2. Replacement 1: Customer Name**
  - Placeholder: CustomerName
  - Replacement Source: Fixed Text
  - Fixed Text: ANONYMIZED
  - Scope: All Occurrences in Document
- 3. Replacement 2: Customer Number**
  - Placeholder: CustomerNumber
  - Replacement Source: Sequential Number
  - Start Value: 1, Increment: 1, Format: D6
  - Scope: Single Occurrence
- 4. Storage Location:**
  - Directory: D:\Archive\Anonymized
  - Filename: <FileName>

### Result

Original	Anonymized
“John Smith”	“ANONYMIZED”
“CU-123456”	“000001”

---

## 31.13 Example: Replace Header on All Pages

### Initial Situation

Multi-page documents have a header with the company name on each page that should be replaced with a new designation.

### Configuration

- 1. Create Extraction Rule:**
  - Rule “CompanyName” with header position
- 2. Replacement:**

- Placeholder: CompanyName
- Replacement Source: Fixed Text
- Fixed Text: New Company Inc
- Scope: **All Pages (Same Position)**

### **Result**

The company name is replaced on all pages at the header position with “New Company Inc”.

---

## 31.14 Example: Update Date

### **Initial Situation**

Template documents should be updated with the current date.

### **Configuration**

1. **Create Extraction Rule:**
  - Rule “Date” with date field position
2. **Replacement:**
  - Placeholder: Date
  - Replacement Source: Current Date/Time
  - Format: MM/dd/yyyy
  - Scope: Single Occurrence

### **Result**

Original	Replaced
“01/01/2024”	“01/12/2026”

---

## 31.15 Tips and Notes

### **Font Handling**

The new text is inserted in a similar font to the original text. With embedded subset fonts, the system automatically falls back to available system fonts.

### **True Redaction**

Unlike simple text overlays, the original text is actually removed from the PDF: - The original text is no longer extractable via copy & paste - Forensic recovery is not possible - GDPR-compliant anonymization

### **Consider Order**

If multiple replacements have overlapping areas, the list order is followed. Ensure replacements don't interfere with each other.

### Use Fallback Rules

For robust configurations, you can use fallback rules: - Multiple rules with the same name - The first successful rule is used - Enables different position variants

### Combination with Other Tasks

Typical workflow order: 1. Extract data (extraction rules) 2. **Replace content** (anonymization) 3. Convert to PDF/A (archiving) 4. Send email or copy

### Performance with Large Documents

With documents having many pages and the "All Pages (Same Position)" setting, processing may take longer. Test the configuration with sample files.

### Manage Sequential Numbers

Current counter values are stored in the AppData configuration. If needed, you can: - Reset the start value - Enable automatic deletion after X days

---

## 32 Flatten Pages

Task: Flatten Pages

*Task: Flatten Pages*

### 32.1 Description

The **Flatten Pages** task permanently burns annotations and form fields into the page content of a PDF document. After flattening, these elements become a fixed part of the page and can no longer be edited or removed.

#### What Does "Flattening" Mean?

PDF documents can contain multiple layers: - **Form Fields:** Fillable text fields, checkboxes, dropdown menus - **Annotations:** Comments, highlights, stamps, freehand drawings - **Page Content:** The actual text and graphics of the page

Flattening converts form fields and annotations into page content. The separate layers are removed and merged into a single, non-editable layer.

#### Typical Use Cases

- **Seal Forms:** Fix filled-in forms so that entered values can no longer be changed
- **Make Annotations Permanent:** Permanently burn comments and markings into the document

- **Increase PDF Compatibility:** Ensure that the document displays identically in all PDF viewers
  - **Archiving:** Convert documents to an immutable state
  - **Distribution:** Prepare PDFs for distribution without recipients being able to edit form fields
- 

## 32.2 General Settings

### Enabled

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

## 32.3 Settings

### Settings

#### *Settings*

The “Flatten Pages” task requires no task-specific settings. All form fields and annotations in the document are automatically flattened.

**Note:** Flattening is an irreversible operation. After flattening, form fields can no longer be filled in and annotations can no longer be edited.

---

## 32.4 Storage Location

### Directory

Specify the target directory for the flattened PDF file.

**Note:** It’s recommended to use a separate folder for each processing step to ensure clear separation.

### Filename

Set the name for the processed file.

### Examples:

Input	Result
(empty)	Form.pdf (original name)
<FileName>_flattened	Form_flattened.pdf

### Name Collisions

Choose what should happen if a file with the target name already exists.

### **Overwrite original file**

When this option is enabled, the source file is overwritten directly. The settings for directory, filename, and name collisions are ignored — the processed file replaces the original file at the same location.

**Warning:** This option is not recommended as the original file is irreversibly lost. Only use this option if you are certain that you no longer need the original.

**Note:** All subsequent tasks in the same profile will operate on the newly created (overwritten) file.

---

## 32.5 File Date

### **Adjust Creation and Modification Date**

Optionally, you can change the file date of the processed file.

---

## 32.6 Afterwards

### **Call External Program**

After flattening, an external program can be started automatically.

**Parameters:** Available placeholders: - <PathIncludingFilename> - Full path of the processed file - <ParentDirectory> - Path of parent folder - <Filename> - Filename

---

## 32.7 Example: Seal Filled-In Forms

### **Initial Situation**

Customers fill in PDF forms and return them. Before archiving, the forms should be sealed so that the entered values cannot be changed afterwards.

### **Configuration**

1. **Enabled:** Yes
2. **Directory:** D:\Archive\Forms\Sealed
3. **Filename:** <FileName>

### **Result**

Original

Processed

---

Original

Processed

Application\_Smith.pdf (with editable fields)

Application\_Smith.pdf (all fields fixed)

---

## 32.8 Example: Fix Annotations Before Distribution

### Initial Situation

Documents are internally annotated with comments and markings. Before distribution to external partners, these annotations should be permanently burned in.

### Configuration

1. **Enabled:** Yes
2. **Directory:** D:\Documents\External
3. **Filename:** <FileName>\_final

### Result

All comments, highlights, and stamps are converted to page content. Recipients see the annotations as a fixed part of the page but can no longer edit or delete them.

---

## 32.5 Tips and Notes

### Order in Task List

Typically place “Flatten Pages” before tasks like Copy or Send Email, so subsequent tasks work with the flattened document.

### Create a Backup Copy

Since flattening is irreversible, it’s recommended to create a backup copy first. Place a “Copy File” task before the “Flatten Pages” task.

### Combination with File Protection

For maximum document security, you can combine flattening with the “Change File Protection” task: First flatten, then set a password.

### Documents Without Form Fields/Annotations

If a PDF document contains no form fields or annotations, the task is still executed successfully. The document remains unchanged.

---

# 33 Auto-rotate Pages

Task: Auto-rotate Pages

*Task: Auto-rotate Pages*

## 33.1 Description

The **Auto-rotate Pages** task detects the text orientation of each page and automatically corrects misrotated pages. Optionally, OCR text recognition can be used for non-searchable (image-based) pages to determine the correct orientation there as well.

### Typical Use Cases

- **Scanner Correction:** Automatic correction of incorrectly fed or rotated scans
  - **Mixed Orientation:** Standardize documents with portrait and landscape pages
  - **Batch Processing:** Correct large volumes of scanned documents with varying orientation
  - **Archive Preparation:** Bring documents to uniform orientation before archiving
- 

## 33.2 General Settings

### Enabled

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

## 33.3 Settings

Settings

*Settings*

### Use OCR for Non-Searchable Pages

Enable this option if your PDFs also contain image-based pages (e.g., scans without text recognition). In this case, OCR analysis is performed to determine text orientation.

**Without OCR:** Only pages with existing text are analyzed. Image-based pages remain unchanged.

**With OCR:** Image-based pages are also analyzed using optical character recognition and rotated if necessary.

### OCR Language

Select the document language for OCR analysis. The correct language significantly improves recognition accuracy.

Setting	Description
<b>Language</b>	The main language of the documents to be analyzed

**Note:** The selected language must be installed. You can download additional languages via the **Install Language...** button.

## 33.4 Storage Location

### Directory

Specify the target directory for the corrected PDF file.

**Note:** It's recommended to use a separate folder for each processing step to ensure clear separation.

### Filename

Set the name for the processed file.

### Examples:

Input	Result
(empty)	Scan_001.pdf (original name)
<FileName>_corrected	Scan_001_corrected.pdf

### Name Collisions

Choose what should happen if a file with the target name already exists.

### Overwrite original file

When this option is enabled, the source file is overwritten directly. The settings for directory, filename, and name collisions are ignored — the processed file replaces the original file at the same location.

**Warning:** This option is not recommended as the original file is irreversibly lost. Only use this option if you are certain that you no longer need the original.

**Note:** All subsequent tasks in the same profile will operate on the newly created (overwritten) file.

## 33.5 File Date

### Adjust Creation and Modification Date

Optionally, you can change the file date of the processed file.

---

## 33.6 Afterwards

### Call External Program

After auto-rotation, an external program can be started automatically.

**Parameters:** Available placeholders: - <PathIncludingFilename> - Full path of the processed file - <ParentDirectory> - Path of parent folder - <Filename> - Filename

---

## 33.7 Example: Correct Scanned Documents

### Initial Situation

A document scanner regularly produces PDFs where individual pages are misaligned — some are rotated 90° or 180°.

### Configuration

1. **Enabled:** Yes
2. **Use OCR for Non-Searchable Pages:** Yes
3. **OCR Language:** English
4. **Directory:** D:\Scans\Corrected
5. **Filename:** <FileName>

### Result

Original	Processed
Scan_001.pdf (3 of 5 pages rotated)	Scan_001.pdf (all pages correctly aligned)

---

## 33.8 Example: Landscape Pages in Mixed Documents

### Initial Situation

Business documents contain both portrait text and landscape tables. For archiving, all pages should be uniformly aligned.

### Configuration

1. **Enabled:** Yes
2. **Use OCR for Non-Searchable Pages:** No (documents are already searchable)
3. **Directory:** D:\Archive\Cleaned
4. **Filename:** <FileName>

### Result

Landscape pages are detected and automatically rotated to portrait. The text orientation is correctly maintained.

---

## 33.5 Tips and Notes

### **When to Enable OCR?**

Enable the OCR option only when your documents contain image-based pages (e.g., pure scans without prior text recognition). For PDFs that already contain searchable text, OCR is not required and would only slow down processing.

### **Choose the Right Language**

The choice of the correct OCR language affects recognition accuracy: - Use the main language of your documents - For multilingual documents, choose the predominant language - For mixed batches (e.g., German and English), English works as a good compromise

### **Order in Task List**

Typically place “Auto-rotate Pages” before other tasks like Copy or Send Email. It’s particularly useful before the OCR task, as correct page orientation improves OCR results.

### **Already Correct Pages**

Pages that are already correctly aligned are not changed. The task only rotates pages where incorrect orientation is detected.

### **Backup Copy**

Create a backup copy if needed by placing a “Copy File” task before the “Auto-rotate Pages” task.

---

## 34 Redact Text

Task: Redact Text

*Task: Redact Text*

### 34.1 Description

The **Redact Text** task permanently and irreversibly removes sensitive text areas from a PDF document. The areas to be redacted are determined via data extraction rules. Unlike a purely visual overlay, the underlying text is actually removed from the document.

### **Important: True Redaction**

With true redaction, text is not just visually covered but physically removed from the PDF. The redacted areas cannot be restored — not even through copying, searching, or technical analysis.

### Typical Use Cases

- **GDPR Compliance:** Remove personal data such as names, addresses, or account numbers from documents
  - **Anonymization:** Anonymize documents for distribution to third parties
  - **Confidential Information:** Remove salary data, social security numbers, or internal reference numbers
  - **Authority Requests:** Prepare documents for information requests in a privacy-compliant manner
  - **Archiving:** Remove sensitive data before long-term archiving
- 

## 34.2 General Settings

### Enabled

Enable this option so the task is executed for matching PDF files. Disabled tasks are skipped.

---

## 34.3 Redactions

### Redactions

#### *Redactions*

In the redactions list, you define which text areas should be redacted. Each entry references a data extraction rule and specifies the appearance of the redaction.

### Redaction Entry

Each entry consists of:

Property	Description
<b>Rule</b>	The data extraction rule that determines the text area to be redacted
<b>Color</b>	The color of the redaction rectangle (default: Black)
<b>Overlay Text</b>	Optional text displayed over the redaction (e.g., “REDACTED”)

### Managing Redaction Entries

Button	Description
<b>Add</b>	Create a new redaction entry
<b>Edit</b>	Edit the selected entry
<b>Remove</b>	Delete the selected entry
<b>Move Up/Down</b>	Adjust the order of entries
<b>Preview</b>	Preview redactions using the example files

### Rule Reference

Redaction uses your data extraction rules to locate the text areas to be redacted. First create an extraction rule that finds the desired text, then reference it in the redaction configuration.

**Example:** A rule “Name” extracts the recipient’s name. In the redaction configuration, this rule is selected — the found text area is then redacted.

### Preview

With the **Preview** button, you can check the redactions in advance using the example files stored in the profile. This ensures that the correct areas are redacted before processing starts.

---

## 34.4 Storage Location

### Directory

Specify the target directory for the redacted PDF file.

**Note:** It’s recommended to use a separate folder for each processing step to ensure clear separation.

### Filename

Set the name for the processed file.

### Examples:

Input	Result
(empty)	Contract.pdf (original name)
<FileName>_anonymized	Contract_anonymized.pdf

### Name Collisions

Choose what should happen if a file with the target name already exists.

### Overwrite original file

When this option is enabled, the source file is overwritten directly. The settings for directory, filename, and name collisions are ignored — the processed file replaces the original file at the same location.

**Warning:** This option is not recommended as the original file is irreversibly lost. Only use this option if you are certain that you no longer need the original.

**Note:** All subsequent tasks in the same profile will operate on the newly created (overwritten) file.

---

## 34.5 File Date

### Adjust Creation and Modification Date

Optionally, you can change the file date of the processed file.

---

## 34.6 Afterwards

### Call External Program

After redaction, an external program can be started automatically.

**Parameters:** Available placeholders: - <PathIncludingFilename> - Full path of the processed file - <ParentDirectory> - Path of parent folder - <Filename> - Filename

---

## 34.7 Example: Anonymize Personnel Data

### Initial Situation

Pay slips contain personal data (name, address, account number) that must be redacted before forwarding to the accounting department.

### Configuration

#### 1. Create Data Extraction Rules:

- Rule “Name” — Extracts the employee name
- Rule “Address” — Extracts the mailing address
- Rule “IBAN” — Extracts the account number

#### 2. Redaction Entries:

- Rule “Name” → Color: Black, Overlay Text: “REDACTED”
- Rule “Address” → Color: Black, Overlay Text: “REDACTED”
- Rule “IBAN” → Color: Black, Overlay Text: “REDACTED”

#### 3. Directory: D:\Accounting\Anonymized

4. **Filename:** <FileName>\_anonymized

### Result

Original	Processed
Name: John Smith	Name: ██████████ REDACTED
IBAN: GB29 NWBK 6016 1331 9268 19	IBAN: ██████████ REDACTED

---

---

## 34.8 Example: Remove Confidential Reference Numbers

### Initial Situation

Contracts contain internal reference numbers that should be removed before distribution to external partners.

### Configuration

1. **Create Data Extraction Rule:**
  - Rule “Reference” — Extracts the internal reference number (e.g., by keyword “Ref:”)
2. **Redaction Entry:**
  - Rule “Reference” → Color: White, Overlay Text: (empty)
3. **Directory:** D:\Contracts\External
4. **Filename:** <FileName>

### Result

The reference numbers are covered with white rectangles and the text is removed. Due to the white color and missing overlay text, the redactions are visually unobtrusive.

---

---

## 34.5 Tips and Notes

### Prepare Data Extraction Rules

The quality of redaction depends directly on the quality of your extraction rules. Test the rules thoroughly with example files before enabling redaction.

### Use the Preview

Use the preview function to check redactions before productive processing. This helps avoid unintended redactions or overlooked areas.

### Color Selection

Color	Usage
<b>Black</b>	Classic redaction, clearly visible

Color	Usage
<b>White</b>	Unobtrusive removal, text “disappears”
<b>Other Colors</b>	Color coding by category (e.g., red for names, blue for addresses)

### Overlay Text

The optional overlay text is displayed over the redaction. Typical texts: - “REDACTED” - “\*\*\*”  
- “[REMOVED]”

### Irreversibility

Redaction is irreversible. Create a backup copy first by placing a “Copy File” task before the “Redact Text” task.

### Mind the Order

Place “Redact Text” before tasks like Copy, Move, or Send Email. This way, subsequent tasks work with the redacted document.

### GDPR Compliance

The redaction function can contribute to fulfilling requirements of the General Data Protection Regulation (GDPR). However, note that proper implementation of data protection requires additional organizational and technical measures.

---

## 35 Program Options

### 35.1 Overview

The **Program Options** contain central settings that apply to all profiles. Here you configure the email account for sending, processing times, backup settings, and other program-wide functions.

**Access:** Menu **Tools** → **Options...** or keyboard shortcut **Ctrl+O**

---

### 35.2 General

#### Language

Select the program language: - **Deutsch** (German) - **English** (default)

The change takes effect after restarting the program.

#### Autostart

Option	Description
<b>Start with Windows</b>	The program runs automatically at Windows startup
<b>Start to system tray</b>	The program starts minimized in the system tray

### Minimize/Close Behavior

Option	Description
<b>Minimize to system tray</b>	When minimizing, the program moves to the system tray
<b>Close to system tray</b>	The close button minimizes the program instead of closing it

### Working Directory

Optionally, you can specify a custom working directory where temporary files required for processing are created. By default, a subfolder in the Windows temporary directory is used.

**Use Case:** When experiencing problems with the default temp directory or when temporary files should be created on a specific drive.

### Centrally Managed Profile File

For enterprise environments, you can specify a centrally managed profile file that is read by all workstations.

Option	Description
<b>Profile file</b>	Path to the central Profiles.json
<b>Allow editing</b>	When disabled, profiles can only be read, not modified
<b>Suppress processing in admin mode</b>	Processing is suspended when the central file is being edited

### Password Storage

Determines where sensitive passwords are stored:

Mode	Description
<b>Encrypted in configuration</b>	Passwords are stored AES-encrypted in the JSON files (default)
<b>Windows Credential Manager</b>	Passwords are stored in the Windows Credential Manager

**Recommendation:** For service operation or shared configurations, Windows Credential Manager is more secure.

---

## 35.3 Processing

### Check Interval

Specify how often the program checks the monitored folders for new files.

Setting	Default
<b>Seconds between checks</b>	120 (2 minutes)

**Note:** A shorter interval leads to faster processing but increases system load.

### Pause Between Processing

Optionally, you can set a pause between processing individual files. This is useful for external systems that need time for synchronization.

### Restrict Processing Period

Limit processing to specific times:

Setting	Description
<b>Weekdays</b>	Process only on selected days
<b>From/To</b>	Time range for processing (e.g., 09:00-17:00)

**Use Case:** Processing only during business hours, for example, to print only when staff is present.

### Error Handling

#### Notify on Errors

Enable this option to automatically receive an email when processing errors occur.

Setting	Description
<b>Recipients</b>	Email addresses for error notifications (comma-separated)

#### Move Files on Errors

Setting	Description
<b>Move files with errors</b>	Move files that could not be processed to a separate folder
<b>Target folder</b>	Path to the error folder

### Files Without Text

Options for PDFs that contain no extractable text:

Option	Description
--------	-------------

Option	Description
<b>Move</b>	Move to a separate folder
<b>Send via email</b>	Send to the specified recipients
<b>Ignore</b>	Exclude these files from processing

### Files Without Filter Match

Option	Description
<b>Move</b>	Move files that don't match any profile filter to a folder
<b>Target folder</b>	Path to the folder for unassigned files

### Advanced Settings

Option	Default	Description
<b>Treat moved files as known</b>	Yes	Files moved by the program are not processed again when reappearing at the original location
<b>Enable ML for barcode detection</b>	Yes	Machine learning (local) for improved barcode detection
<b>Remember already processed files</b>	Yes	Stores checksums of processed files to avoid duplicate processing
<b>Checksum retention (days)</b>	90	How long file checksums are stored (0 = unlimited)
<b>File processing timeout (seconds)</b>	0	Maximum processing time per file. If exceeded, the file is skipped and marked as an error. 0 = disabled. Recommended: 120 seconds for large batch processing.
<b>Metadata reading timeout (seconds)</b>	10	Maximum wait time when reading PDF metadata. If exceeded, the file is treated as unreadable. 0 = unlimited.
<b>Max. wait time for locked files (seconds)</b>	10	Wait time before a locked file is skipped. The file will be retried at the next check interval.
<b>File size stability for new files (seconds)</b>	0	Checks whether the file size has not changed for the specified number of seconds before processing begins. Useful for network drives where files are still being copied. 0 = disabled.
<b>Background processor is managed externally</b>	No	Enable this option when the background processor (PDFProcessor.exe) is started

Option	Default	Description
		externally, e.g. as a Windows service or via Task Scheduler. The main application will no longer start, monitor, or terminate the background processor. IPC communication (e.g. immediate check, report generation) continues to work.

**Note on external processor mode:** When the background processor is managed externally, it cannot be paused during automatic backups. Individual files that the processor is currently accessing will be skipped during backup. Make sure to manage the service via Windows Service Manager (services.msc) or `sc stop`.

---

## 35.4 Email Account for Sending

Configure the email account used for all email tasks and notifications.

### Account Type

Type	Description
<b>SMTP</b>	Direct sending via an SMTP server
<b>Outlook</b>	Sending via installed Microsoft Outlook (classic)
<b>Exchange Online</b>	Sending via Microsoft 365 / Exchange Online

### SMTP Settings

Setting	Description
<b>Server</b>	Hostname of the SMTP server (e.g., smtp.provider.com)
<b>Port</b>	SMTP port (default: 25, SSL: 465, STARTTLS: 587)
<b>Username</b>	Login name for the SMTP server
<b>Password</b>	Password for authentication
<b>Sender name</b>	Name displayed as sender
<b>Sender email</b>	Email address of the sender

### Encryption

Option	Description
<b>Auto-negotiate</b>	Use best available encryption (recommended)
<b>None</b>	Unencrypted connection
<b>SSL/TLS immediately</b>	Encryption from the start (port 465)
<b>STARTTLS required</b>	STARTTLS must be supported

Option	Description
<b>STARTTLS if available</b>	Use STARTTLS if supported

### Certificate Validation

Option	Default	Description
<b>Check whether the TLS/SSL certificate has been revoked</b>	Yes	Checks the Certificate Revocation List (CRL) of the server certificate
<b>Ignore SSL/TLS certificate errors</b>	No	Ignores certificate errors such as self-signed, expired, or mismatched certificates. <b>Security note:</b> Only use this option if you trust the server and understand the cause of the certificate error.

### S/MIME Signing

For digital signatures, you can configure an S/MIME certificate:

Source	Description
<b>Windows Certificate Store</b>	Select certificate from Windows store
<b>Certificate file</b>	PFX/P12 file with private key

### Outlook Settings

Setting	Description
<b>Use default account</b>	Use the default account configured in Outlook
<b>Custom account</b>	Select a specific Outlook account via SMTP address

### Exchange Online Settings

Setting	Description
<b>Sender name</b>	Name of the sender
<b>Sender email</b>	Email address (must be linked to Microsoft 365)
<b>Timeout</b>	Connection timeout (default: 30 seconds)

**Note:** Exchange Online requires one-time authentication with Microsoft.

### Pause After Sending

Optionally, you can set a pause between email transmissions to respect email server limits.

---

## 35.5 Excel Report via Email

Automatic sending of an Excel report with processing statistics.

### Send Report

Setting	Description
<b>Send report via email</b>	Enables automatic report sending
<b>Recipients</b>	Email addresses for the report (comma-separated)

### Save Report

Setting	Description
<b>Save report locally</b>	Saves the report as an Excel file
<b>Directory</b>	Storage location for report files

### Interval

When should the report be created?

Interval	Description
<b>Daily</b>	Every day at the specified time
<b>Weekly</b>	On a specific day of the week
<b>Monthly</b>	On a specific day of the month

### Report Period

What time period should the report cover?

Period	Description
<b>Current day</b>	Only today
<b>Previous day</b>	Yesterday
<b>Previous week</b>	The complete previous week (Mon-Sun)
<b>Previous month</b>	The complete previous month
<b>Last X days</b>	The last N days
<b>From first of month to today</b>	From the beginning of the month until today

## 35.6 Log

### Retention Period

Setting	Default
<b>Days</b>	365

Log entries older than this period are automatically deleted.

### Display Limit

Setting	Default
<b>Maximum number of displayed entries</b>	25,000

Limits the number of log entries displayed in the main window to optimize performance.

---

## 35.7 Processing Log

Creates a daily CSV log of all processed files. The log is automatically appended after each file is processed.

### Enable Log

Setting	Default
<b>Create processing log as CSV file</b>	No

### Format

Format	Description
<b>Compact</b>	One row per file with a single details column (like the log area)
<b>Expanded</b>	One row per file with one column per task type (like the Excel report)
<b>Task level</b>	One row per executed task — ideal for automated analysis and database import

### Storage Directory

By default, log files are stored in the AppData directory under Logs/. You can specify an alternative directory.

**File name:** ProcessingLog\_YYYY-MM-DD.csv (one log file per day)

### Delimiter

Option	Character
<b>Semicolon</b>	; (default)
<b>Comma</b>	,
<b>Tab</b>	Tab

### Encoding

Choose the character encoding for the CSV file (default: UTF-8 with BOM).

## Retention

Setting	Default
<b>Retention period in days</b>	30

Older log files are automatically deleted when the background processor starts.

**Note:** Status values in the log (e.g., “Success”, “Error”) follow the configured program language.

---

---

## 35.8 Backups

Automatic backups of the program configuration.

### Enable Backup

Setting	Default
<b>Create backups</b>	Yes

### Backup Interval

Setting	Default	Description
<b>Every X hours</b>	4	Time-based backup
<b>After profile changes</b>	No	Backup after each profile change

### Backup Directory

By default, backups are stored in the AppData folder. You can specify an alternative directory (e.g., a network drive).

### Maximum Count

Setting	Default
<b>Maximum number of backups</b>	50

Older backups are automatically deleted when this count is exceeded.

---

---

## 35.9 Print Settings

Default settings for the print task.

### Profile Presets

Setting	Default	Description
<b>Stop processing after</b>	Yes	Additional profiles are not applied to

Setting	Default	Description
<b>applying</b>		the same file
<b>Treat recurring files as new</b>	No	Files that reappear in the folder are processed again

### Print Options

Setting	Default	Description
<b>Auto-rotate pages</b>	Yes	Pages are optimally aligned to paper size
<b>Auto-center content</b>	No	Print content is centered on the page
<b>Size mode</b>	Fit	How content is adapted to paper
<b>Scaling</b>	100%	Percentage scaling of print content

### Size Mode

Mode	Description
<b>Fit</b>	Content is scaled to page size
<b>Original size</b>	Content is printed at original size
<b>Custom</b>	Manual scaling in percent

## 35.10 Dynamic Lists

Dynamic lists enable centralized management of values that can be used in extraction rules or for filtering.

### List Management

Action	Description
<b>New</b>	Create new dynamic list
<b>Edit</b>	Edit selected list
<b>Delete</b>	Remove selected list

### List Types

Type	Description
<b>Static</b>	Fixed defined values
<b>CSV file</b>	Values are loaded from a CSV file
<b>Excel file</b>	Values are loaded from an Excel file

### Usage

In extraction rules, you can reference dynamic lists with: - <QueryWithList{ListName}> - Query value from the list - <EntryFromList{ListName}> - In filter: Check if value is in list

**Use Case:** A list of customer numbers that is regularly updated can be maintained centrally and is available to all rules.

**Note:** When lists are used as filters, the column separator is optional for single-column lists and the column specification for the result is irrelevant. However, when using the data type “Query (with List)”, the separator must be specified.

---

## 35.11 Password List

Central management of passwords for encrypted PDF files.

### Adding Passwords

Add passwords that the program automatically tries when opening encrypted PDFs.

Action	Description
<b>Add</b>	Add new password to the list
<b>Remove</b>	Delete selected password

### How It Works

When the program detects an encrypted PDF file: 1. Passwords from the current profile are tried 2. Passwords from this global list are tried 3. On success, the file is opened with the found password

**Security Note:** Passwords are stored encrypted, depending on the setting under “General” → “Password Storage”.

---

## 35.12 Other

### Check Folder Overlap

Setting	Default
<b>Check monitored folders for overlap</b>	No

When enabled, the program warns when saving if multiple profiles with insufficient filter criteria monitor the same folder or overlapping folders.

### Trim Function for New Rules

Setting	Default
<b>Automatically add trim function</b>	Yes

New extraction rules automatically receive the cleanup function “Remove spaces at beginning and end”.

### Printer Availability Check

Setting	Default
<b>Simplified printer availability check</b>	Yes

When enabled, only checks whether the printer is installed, not whether it is available. This can avoid problems with network printers.

### Folders for Encrypted Files Without Password

Specify folders where encrypted PDF files should be processed even without a password. The files are processed in the state in which they are accessible.

**Use Case:** PDFs with permission password that are readable but cannot be edited.

---

## 35.13 Tips and Notes

### Backup Configuration

Use the automatic backup function to protect your configuration. In case of problems, you can restore a previous state.

**Backup Path:** %APPDATA%\Gillmeister Software\Automatic PDF Processor 2\Backups\

### Email Test

Before using email tasks in production, test the email settings with the test function in the Options dialog.

### Processing Times

The “Restrict processing period” setting is useful for: - Print tasks that should only be executed when staff is present - Environments where files are only expected at certain times - Systems reserved for maintenance or backups at night

### Centralized Profile Management

For enterprises with multiple workstations: 1. Save profiles on a network drive 2. Enter the path under “Centrally managed profile file” 3. Optional: Disable “Allow editing” for end users 4. Optional: Enable “Suppress processing in admin mode” for the administrator

---

# 36 Advanced Features

## 36.1 Command Line Parameters

The PDF Processor (PDFProcessor.exe) can be controlled via command line parameters. This enables integration into scripts, scheduled tasks, or operation as a Windows service.

### Available Parameters

Parameter	Short Form	Description
--headless	-h	Starts in headless mode (without GUI dependency)
--service	-s	Starts as native Windows service (Service Control Manager)
--run_once	-r	Performs a single processing run, then exits
--quit	-q	Terminates a running instance
--profile "Name"	-p "Name"	Execute only the specified profile
--profiles [A,B,C]	-ps [A,B,C]	Execute only the specified profiles (comma-separated)

### Examples

```
# Single processing run
```

```
PDFProcessor.exe --run_once
```

```
# Short form
```

```
PDFProcessor.exe -r
```

```
# Start as Windows service
```

```
PDFProcessor.exe --service
```

```
# Start as headless console (endless loop)
```

```
PDFProcessor.exe --headless
```

```
# Terminate running instance
```

```
PDFProcessor.exe --quit
```

```
# Headless with single processing
```

```
PDFProcessor.exe --headless --run_once
```

```
# Process only a specific profile
```

```
PDFProcessor.exe --headless --run_once --profile "Invoice Import"
```

```
# Process multiple profiles
```

```
PDFProcessor.exe -h -r --profiles [Invoices,Archive,Backup]
```

## Profile Selection

With the `--profile` and `--profiles` parameters, you can limit processing to specific profiles. This is particularly useful for Windows Task Scheduler when different scheduled tasks should execute different profiles.

### Single Profile:

```
PDFProcessor.exe -h -r -p "Invoice Import"
```

```
PDFProcessor.exe --headless --run_once --profile "Orders"
```

### Multiple Profiles:

```
PDFProcessor.exe -h -r -ps [Invoices,Orders,Archive]
```

```
PDFProcessor.exe --headless --run_once --profiles [NightProcessing,Backup]
```

### Profile Names with Commas:

If a profile name itself contains a comma, it must be enclosed in quotes. Both double and single quotes are supported:

```
# Profile name "Invoices, Import" contains a comma
```

```
PDFProcessor.exe -h -r --profiles ["Invoices, Import",Archive]
```

```
# With single quotes
```

```
PDFProcessor.exe -h -r --profiles ['Invoices, Import',Archive]
```

```
# Multiple profiles with commas
```

```
PDFProcessor.exe -h -r --profiles ["Invoices, Incoming","Orders, Outgoing",Backup]
```

```
# Mixed: profiles with and without commas
```

```
PDFProcessor.exe -h -r --profiles [Normal,"With, Comma",AlsoNormal]
```

**Behavior:** - Profile names are not case-sensitive (“Invoices” = “invoices” = “INVOICES”) - Non-existent profiles are skipped (with warning in console) - Inactive profiles are also skipped (Active flag takes precedence) - Without profile specification, all active profiles are processed (default behavior)

### Console Output:

```
[08:00:00] Profiles filter: Invoices, Archive
```

or for all profiles:

```
[08:00:00] Processing all active profiles
```

### Combinations

Combination	Behavior
--headless	Autonomous continuous operation without GUI (console mode)
--service	Native Windows service (continuous operation)
--run_once	Process once, then exit
--headless --run_once	Process once without GUI, then exit
--quit	Terminate running instance

### Exit Codes

Code	Meaning
0	Successful termination
1	Unhandled exception
2	License expired or invalid
3	Configuration error
4	Another instance already active

## 36.2 Headless Mode / Service Operation

In headless mode, PDFProcessor runs without dependency on the GUI application. This is ideal for server operation or running as a Windows service.

### Use Cases

Scenario	Recommended Parameters
Windows Service (SC)	--service or no parameter
Windows Service (NSSM)	--headless
Docker Container	--headless
Scheduled Task (one-time)	--run_once
Script Integration	--headless --run_once

### Installation as Windows Service

#### With SC (Windows Service Control) — recommended

PDFProcessor automatically detects when it is started by the Service Control Manager. The -s parameter is optional:

```
# Create service (automatic detection)
sc create PDFProcessor binPath= "C:\Program Files\Automatic PDF Processor\PDF Processor.exe"
```

```

# Alternatively: Create service with explicit service parameter
sc create PDFProcessor binPath= "C:\Program Files\Automatic PDF Processor\PDF
Processor.exe -s"

# Set startup type to automatic
sc config PDFProcessor start= auto

# Start service
sc start PDFProcessor

# Stop service
sc stop PDFProcessor

# Remove service
sc delete PDFProcessor

```

### With NSSM (Non-Sucking Service Manager)

NSSM manages the process externally and therefore requires headless mode (`--headless`):

```

# Install service
nssm install PDFProcessor "C:\...\PDFProcessor.exe" --headless

# Start service
nssm start PDFProcessor

# Stop service
nssm stop PDFProcessor

# Remove service
nssm remove PDFProcessor

```

### Difference: `--service` vs. `--headless`

Mode	Usage	Description
<code>--service / -s</code>	SC, Windows Services	Registers with the Service Control Manager. Required for native Windows services.
<code>--headless / -h</code>	NSSM, Docker, Scripts	Console headless mode without GUI dependency. For all other scenarios.

**Note:** When PDFProcessor is started by the Service Control Manager, it detects this automatically — even without the `-s` parameter. The parameter serves as an explicit safeguard.

### Diagnostics

If problems occur with the Windows service, a diagnostic file is automatically created:

- **File:** ServiceDiagnostics.log (in the same directory as PDFProcessor.exe)
- **Content:** Start messages, error messages, mode detection

## Setting Up Scheduled Tasks

For regular processing without continuous operation:

1. **Open Task Scheduler** (taskschd.msc)
2. **Create New Task**
3. **Action:** Start program → PDFProcessor.exe -r
4. **Trigger:** As needed (e.g., hourly, daily)

*# Example: Daily processing at 8:00 AM (all active profiles)*

```
schtasks /create /sc daily /tn "PDFProcessor" /tr "\"C:\...\PDFProcessor.exe\"  
" -r" /st 08:00
```

*# Example: Daily invoice processing at 9:00 AM (specific profile only)*

```
schtasks /create /sc daily /tn "PDF Invoices" /tr "\"C:\...\PDFProcessor.exe\"  
" -h -r -p \"Invoice Import\"" /st 09:00
```

*# Example: Nightly archiving at 2:00 AM (multiple profiles)*

```
schtasks /create /sc daily /tn "PDF NightArchive" /tr "\"C:\...\PDFProcessor.  
exe\" -h -r --profiles [Archive,Backup,Cleanup]" /st 02:00
```

**Tip:** With different scheduled tasks for different profiles, you can set up flexible processing scenarios - e.g., invoices hourly during the day, archiving once nightly.

## Console Output

In headless mode, the processor outputs status information:

```
[08:00:00] Automatic PDF Processor 2.0.0 started  
[08:00:00] Version: 2.0.0  
[08:00:00] License: VALID  
[08:00:00] Mode: Headless + RunOnce + Profiles: [Invoices, Archive]  
[08:00:00] Profiles filter: Invoices, Archive  
[08:00:00] Culture: en-US  
[08:00:00] AppData: C:\Users\...\AppData\...
```

## Configuration

Headless mode uses the same configuration as the GUI application: - **Configuration**

**Path:** %APPDATA%\Gillmeister Software\Automatic PDF Processor 2\ - **Profiles:** Profiles.json - **Options:** Options.json

**Note:** With centralized profile management, the path configured in options is used.

---

## 36.3 Processing Time Diagnostics

The diagnostics window provides an overview of processing times and program activity.

**Access:** Menu **Tools** → **Processing Time Diagnostics...**

### Tab: Runtime

Shows program activity for the last 30 days.

Column	Description
<b>Date</b>	Day of activity
<b>Start</b>	Program start time
<b>End</b>	End time (or last known activity)
<b>Duration</b>	Total runtime

### Tab: Global Processing Times

Shows the global time window settings configured in program options: - Active weekdays - Processing period (From/To) - Next scheduled time

### Tab: Profile-Specific Processing Times

Lists all profiles that use their own processing times:

Column	Description
<b>Profile</b>	Profile name
<b>Status</b>	Active/Inactive in current time window
<b>Next Window</b>	When the next time window begins
<b>Processed Today</b>	Number of files processed today

### Tab: Waiting Files

Shows files waiting for their time window:

Column	Description
<b>File</b>	Path to the waiting file
<b>Profile</b>	Assigned profile
<b>Waiting Until</b>	Start of the next time window

---

## 36.4 Catch-Up Processing

The “Catch Up” function allows applying one or more profiles to all PDF files from a specific time period.

### How It Works

1. **Select Profiles:** Mark the desired profiles in the profile list (multiple selection with Ctrl+Click or Shift+Click possible)
2. **Start Catch-Up:** Click **Catch Up** in the toolbar above the profile list
3. **Set Time Period:** If needed, you can restrict the time period (From/To)

4. **Processing:** The selected profiles are applied to all PDF files in their monitored folders

### Dialog Options

Option	Description
<b>From/To</b>	Restrict time period (based on file modification date)
<b>Without Restriction</b>	Process all files in monitored folders

### Use Cases

- **Historical Data:** Migrate existing files to a new archive
- **Error Correction:** Reprocess files after profile adjustment
- **Manual Trigger:** Process files from another source folder
- **Test:** Test new profiles with existing files

## 36.5 Retry on Errors

The retry system enables automatic repetition of failed processing.

### Configuration (per Profile)

The retry settings can be found in **Profile Settings** under **General** → **Retries**.

Setting	Default	Description
<b>Active</b>	No	Enable retries
<b>Maximum Attempts</b>	3	Number of retry attempts
<b>Period</b>	1 hour	Attempts are distributed within this period

### Interval Calculation

Attempts are distributed evenly over the period:

Attempts	Period	Interval Between Attempts
3	1 hour	20 minutes
3	3 hours	1 hour
5	1 hour	12 minutes

### Period Options

Unit	Usage
<b>Seconds</b>	For quick retries (e.g., network issues)
<b>Minutes</b>	For medium-term retries
<b>Hours</b>	For longer-term retries (default)

Unit	Usage
<b>Days</b>	For very long retry periods

### Process Flow

1. **Error during processing** → File is added to retry list
2. **Wait** → System waits for calculated interval
3. **Retry** → Processing is attempted again
4. **On success** → File is removed from retry list
5. **On another error** → Next attempt after another interval

### Automatic Cleanup

- **7-Day Rule:** Entries whose last error was more than 7 days ago are automatically removed
- **Maximum Attempts:** No further retry is performed after reaching maximum attempts

### Retry Status in Log

Errors with active retry function show additional information:

Information	Description
<b>Next Attempt</b>	Time of next retry attempt
<b>Remaining</b>	Number of remaining attempts

### Use Cases

Scenario	Recommended Configuration
Network Printer	3 attempts, 10 minutes
Email Server	3 attempts, 30 minutes
External Storage	5 attempts, 1 hour
Batch Processing	3 attempts, 2 hours

### Example

**Configuration:** - Maximum attempts: 3 - Period: 1 hour

**Process on error at 10:00:** 1. First error: 10:00 2. Second attempt: 10:20 3. Third attempt: 10:40 4. Final attempt: 11:00

After the final attempt, the file is marked as “failed” and appears in the log under “Failed Retries”.

### Notification on Exhausted Attempts

In the profile notification settings, you can enable email notification when all retry attempts have failed.

---

## 36.6 Tips and Notes

### Service Operation

- **Permissions:** The service runs by default as SYSTEM account. For access to network drives, a user account with appropriate rights must be used.
- **User account recommended:** The SYSTEM account uses its own AppData directory (C:\Windows\System32\config\systemprofile\AppData\Roaming\). To ensure the service uses the same configuration as the GUI application, it should run under the same user account.
- **Logging:** In headless and service mode, errors are written to the console and error log. Additionally, ServiceDiagnostics.log is created next to the EXE.
- **Configuration Changes:** Changes to profiles are automatically applied.

### Retry Strategy

- **Temporary Errors:** Short intervals (minutes) for network issues, printer availability
- **External Dependencies:** Longer intervals (hours) for external systems that may be under maintenance
- **Not Too Many Attempts:** Too many retries can burden the system

### Catch-Up Processing

- **Multiple Profiles:** You can select multiple profiles at once to apply them together to a time period
  - **Consider Filters:** Only files matching profile filters are processed
  - **Large Time Periods:** Processing may take longer for very large time periods
- 

## 37 Language Installation

### 37.1 Overview

The **Language Installation** window allows downloading and installing additional languages for OCR text recognition. By default, only German is installed. For recognition of documents in other languages, the corresponding language packages must be installed.

**Access:** Menu **Tools** → **Install OCR Languages...**

---

## 37.2 Available Languages

The system supports over 100 languages, including:

**Western European Languages** - German (pre-installed) - English - French - Spanish - Italian - Portuguese - Dutch

**Eastern European Languages** - Polish - Czech - Hungarian - Russian - Ukrainian

**Asian Languages** - Chinese (simplified and traditional) - Japanese - Korean - Arabic - Hebrew

**Other Languages** - The complete list is available in the installation window

---

## 37.3 Installation

### Selecting a Language

1. Open the **Install OCR Languages** window
2. Find the desired language in the list
3. Check the checkbox next to the language
4. Click **Install**

### Download

Language packages (Tessdata) are downloaded from official Tesseract repositories: - File size: 1-50 MB depending on language - Storage location: %APPDATA%\Gillmeister Software\Automatic PDF Processor 2\tessdata\

### Multiple Languages

You can select and install multiple languages at once. Downloads proceed sequentially.

---

## 37.4 Uninstallation

To remove an installed language: 1. Uncheck the checkbox for the language 2. Click **Apply Changes** 3. The language file is deleted

**Note:** German cannot be uninstalled as it is one of the default languages.

---

## 37.5 Usage in Profiles

After installation, languages are available in the OCR settings of profiles:

1. Open a profile → **Tasks** → **OCR**

2. Select the **Primary Language**
3. Optional: Select a **Secondary Language** for multilingual documents

### Primary and Secondary Language

Setting	Usage
<b>Primary Language</b>	Main language of the document (required)
<b>Secondary Language</b>	For documents with mixed text (optional)

**Example:** A German document with English technical terms: - Primary Language: German - Secondary Language: English

---

## 37.6 Tips and Notes

### Language Quality

Recognition quality depends on the language: - **Very good:** Western European languages with Latin script - **Good:** Eastern European languages, Greek - **Variable:** Asian languages, depending on print quality

### Storage Requirements

Language Type	Typical Size
European Languages	1-5 MB
Asian Languages	10-50 MB
All Languages	~500 MB

### Offline Installation

For environments without internet access, Tessdata files can be manually downloaded and copied to the tessdata folder.

---

## 38 Catch-Up Processing

### 38.1 Overview

The **Catch Up** function allows applying one or more profiles to all PDF files from a specific time period. This is useful for reprocessing files or applying profiles to historical data.

**Access:** Select profiles in the profile list → **Catch Up** button in the profile toolbar

---

## 38.2 Procedure

### Step 1: Select Profiles

Select the profiles in the profile list that should be applied to the files: - **Single Profile:** Simply click - **Multiple Profiles:** Add individual profiles to selection with **Ctrl+Click** - **Range:** Select a contiguous range with **Shift+Click**

### Step 2: Start Catch-Up

Click the **Catch Up** button in the toolbar above the profile list. The catch-up processing dialog opens.

### Step 3: Set Time Period (optional)

Option	Description
<b>From</b>	Start date for processing (based on file modification date)
<b>To</b>	End date for processing
<b>Without Restriction</b>	Process all files in monitored folders

### Step 4: Start Processing

Click **Start** to begin processing. The selected profiles are applied to all PDF files in their respective monitored folders that match the time period filter.

---

## 38.3 Processing

### Process Flow

1. For each selected profile, the configured monitored folders are searched
2. PDF files matching the time period filter are identified
3. Profile filter rules are applied
4. Matching files are processed

### Progress

Processing progress is displayed in the main window: - Current file - Number of processed files - Processing status

### Cancel

The running catch-up processing can be stopped via the **Cancel** button in the main window. Files already processed remain processed.

---

## 38.4 Log

Catch-up processed items appear in the regular processing log: - Processing can be filtered via the log filter - Successes and errors are logged normally - The profile used is displayed

---

## 38.5 Use Cases

### **Migrate Historical Data**

Transfer existing PDF files to a new archiving system: 1. Create archive profile with desired tasks (e.g., Rename, Move, PDF/A) 2. Ensure monitored folders are correctly configured 3. Select profile in the list 4. Click **Catch Up** and optionally restrict time period 5. Start processing

### **Error Correction**

Reprocess files after adjusting a profile: 1. Correct profile and save 2. Select profile in the list 3. Click **Catch Up** 4. Restrict time period to the error period 5. Start processing

### **Test Before Production**

Test new profile with existing files: 1. Create new profile with desired settings 2. Configure monitored folder with test files 3. Select profile and click **Catch Up** 4. Start processing and check results

### **Apply Multiple Profiles Simultaneously**

Apply different profiles to the same time period: 1. Select all relevant profiles with **Ctrl+Click** 2. Click **Catch Up** 3. Set common time period 4. Start processing

---

## 38.6 Tips and Notes

### **Large Data Volumes**

With many files: - Restrict time period to form batches - Monitor system load during processing - Process in multiple runs if needed

### **Consider Filter Rules**

Filter rules of selected profiles are applied: - Only files matching the respective profile filter are processed - Non-matching files are skipped (not counted as errors)

### **Monitored Folders**

Catch-up processing uses the monitored folders configured in the profiles: - Ensure folders contain the desired files - Ensure stable connection for network folders

---

## 39 Processing Time Diagnostics

### 39.1 Overview

The **Processing Time Diagnostics** window provides a comprehensive overview of program activity and configured processing times. It helps with analysis and troubleshooting of time-controlled processing.

**Access:** Menu **Tools** → **Processing Time Diagnostics...**

---

### 39.2 Tab: Runtime

Shows program activity for the last 30 days based on the heartbeat mechanism.

#### Display

Column	Description
<b>Date</b>	Day of the activity period
<b>Start</b>	Program start time
<b>End</b>	Time of last known activity
<b>Duration</b>	Total runtime of the activity period

#### Heartbeat

The program regularly updates a timestamp (every 60 seconds). Activity periods are reconstructed from these timestamps.

#### Usage

- **Check availability:** Was the program active at certain times?
  - **Detect restarts:** Multiple periods in one day indicate restarts
  - **Analyze runtime:** How long did the program run in total?
- 

### 39.3 Tab: Global Processing Times

Shows the global time window settings configured in program options.

#### Display

Information	Description
<b>Status</b>	Whether global time windows are enabled
<b>Weekdays</b>	Active processing days

Information	Description
<b>Period</b>	Configured From-To time range
<b>Currently Active</b>	Whether a time window is currently active
<b>Next Window</b>	When the next time window begins

### Changes

Global processing times are configured in **Program Options** under **Processing**.

## 39.4 Tab: Profile-Specific Processing Times

Lists all profiles that have defined their own processing times (independent of global settings).

### Display

Column	Description
<b>Profile</b>	Profile name
<b>Schedule</b>	Summary of configured times
<b>Status</b>	Whether the time window is currently active
<b>Next Window</b>	Start of the next time window
<b>Window Ends</b>	End of current time window (when active)
<b>Processed Today</b>	Number of files processed today

### Detail View

Double-clicking a profile opens the detail view with: - Complete schedule configuration - Processing history - Next scheduled times

### Configuring Profile-Specific Times

1. Open profile → **General**
2. Enable **Use profile-specific processing times**
3. Configure schedule (weekdays, time ranges, repetitions)

## 39.5 Tab: Waiting Files

Shows files waiting for their time window.

### Display

Column	Description
<b>File</b>	Full path to the waiting file

Column	Description
<b>Profile</b>	The profile that will process the file
<b>Detected At</b>	Time when the file was found
<b>Waiting Until</b>	Start of the next time window
<b>Remaining</b>	Remaining wait time

#### Waiting Files Occur When:

1. A file is found outside the global time window
2. A file belongs to a profile whose time window is not active
3. A profile with “Once per time window” has already processed in this window

#### Actions

Action	Description
<b>Process Now</b>	Process selected file immediately (ignores time window)
<b>Remove from Queue</b>	File will not be processed

## 39.6 Schedule Types

### Time Range (FromTo)

Processing during a time range (e.g., 09:00-17:00): - Files are processed immediately when they appear in the time window - Unlimited processing per time window

### Time Point (At)

Processing at a specific time (e.g., 09:00): - Processing only once per time window - Files that appear later wait for the next window

### Catch-Up

When enabled, missed processing is caught up: - Example: Time point 09:00, Catch-up 30 minutes - If program starts at 09:15, processing still occurs - After 09:30, the time window is missed

## 39.7 Troubleshooting Tips

### Files Are Not Being Processed

1. **Check global time windows:** Is the current time window active?
2. **Check profile-specific times:** Does the profile have its own restrictions?
3. **Check waiting files:** Is the file waiting for a time window?

4. **Check runtime:** Was the program active at the expected time?

#### **Processing Does Not Occur at Expected Time**

1. **Time point vs. time range:** With “time point”, only one processing per window
2. **Catch-up setting:** Was the window possibly missed?
3. **Check system time:** Is the computer’s system time correct?

#### **Unexpected Delays**

1. **Other profiles:** Are FileGroups blocking processing?
  2. **Check interval:** How often does the program check? (Default: 2 minutes)
  3. **System load:** Is the system overloaded?
-