



# **User Manual for PdfClerk**



# Table of contents

<b>Introduction</b> . . . . .	3
<b>Presets and example files</b> . . . . .	3
Preset: Simple Merge . . . . .	3
Preset: Merge with header and footer . . . . .	3
Preset: Simple linking . . . . .	4
Preset: Linked Pleading as Binder . . . . .	4
Preset: Linked Pleading Attachment Binder . . . . .	4
Preset: Set of Linked Pleadings . . . . .	4
Preset: Stamp existing binder . . . . .	5
<b>Input tab</b> . . . . .	5
Input files . . . . .	5
Item metadata . . . . .	5
Reorder documents . . . . .	6
Level . . . . .	6
Bookmarks . . . . .	6
Document properties . . . . .	6
Select the output folder . . . . .	7
Existing binders . . . . .	7
Include Exhibits in Subfolder(s) . . . . .	7
Document Filename Metadata Extraction . . . . .	7
Exhibit Filename Metadata Extraction . . . . .	8
Existing links . . . . .	8
<b>Output tab</b> . . . . .	8
General settings . . . . .	8
Table of contents . . . . .	9
Stamps . . . . .	9
Add Bookmarks . . . . .	10
Accessing Metadata . . . . .	10
Document Properties . . . . .	11
Digitally Signed Documents . . . . .	11
<b>Autolink tab</b> . . . . .	11
General . . . . .	11
Prerequisites . . . . .	11
How it Works . . . . .	12
Other Autolink Settings . . . . .	12
Multi binder collections . . . . .	13
Searches . . . . .	13
Norwegian Civil Litigation Configuration . . . . .	14
<b>Load/Save tab</b> . . . . .	15
Projects . . . . .	15
Excel format . . . . .	15
<b>Tools tab</b> . . . . .	16
Convert Clerk Links . . . . .	16
Create Index . . . . .	17
Reconnect links in binders . . . . .	19
Transform links . . . . .	19
<b>Global keyboard shortcuts</b> . . . . .	20
<b>Legal stuff</b> . . . . .	20
Data privacy . . . . .	20
License . . . . .	20



## Introduction

PdfClerk is a desktop software for efficiently creating feature-rich document collections and binders.

Main features:

- Automatically create hyperlinks to referenced documents, typically between pleadings and exhibits. Links are maintained if documents are moved or combined in binders.
- Create linked litigation bundles of pleadings and their exhibits
- Extract metadata to be used for header & footers (stamps) and bookmarks from the filename and add and manage additional metadata like exhibit numbers, titles, and dates
- Create stamps with graphic elements like borders, solid or transparent backgrounds and company logos in addition to text derived from metadata.
- Stamps can be added to binders or individual files in a document set
- Import and export metadata using Microsoft Excel format
- Update bookmarks and add stamps to existing binders based on data extracted from existing bookmarks by using the in-app data grid or Excel
- Create hyperlinks to specific pages in the same or other binders using Clerk Links.
- Split binders based on bookmark level
- Access to standard settings and workflows using configurable presets

PdfClerk's auto-link capabilities can be configured to work in different languages and on other materials and document types such as contracts, documentation and reports.

## Presets and example files

PdfClerk ships with predefined settings for various pdf building tasks. Examples showing the functionality of PdfClerk are available [here](#). The presets and example files may serve as an excellent introduction to how PdfClerk works. You can create your own presets by saving a project to the presets folder (by default located in *Documents\PdfClerk\Presets\*).

### Preset: Simple Merge

*Simple Merge* is the default setting, merging documents and adding the filenames (without the filename extension) as bookmarks.

1. Select and load the *Simple Merge* preset
2. Click the *Open input files* button select *Single folder* and then select the folder with the pdf files you want to merge
3. Use the "Select output folder" to select the folder to store the merged file
4. Click the Generate button

### Preset: Merge with header and footer

Same as *Simple Merge* but with the filename (without extension) as document name in the upper left corner and page number and total pages of binder in the bottom right corner.

1. Select and load the *Merge with header and footer* preset



2. Click the *Open input files* and *Single folder* or *Folder & subfolders* to select the folder with the pdf files you want to merge
3. Click the *Select output folder* to select the folder to store the merged file
4. Click the *Generate* button

## Preset: Simple linking

Converts references to pages in the other documents in the folder into active links. The reference text is the filename (less the extension) of the file to be referenced or any aliases provided. Links are underlined. Uncheck the *Process* checkbox in the document list for documents you do not want to convert references in.

1. Select *Single file to link* using the “Open input files” button
2. Select a Word document or a pdf document
3. The document list loads with all pdf documents in the same folder as the selected file selected. If the text reference does not match the filename, edit the text in the Aliases box to correspond with the text in the reference
4. Click the Generate button

## Preset: Linked Pleading as Binder

Merges a pleading with its exhibits and creates blue underlined links that are 2 pt wide (thick) between references to the exhibits and the exhibits. Adds a bookmark for each exhibit with *Bilag X* as the title. Exhibit numbering inside a rounded rectangle on the top right and page count on the bottom right.

1. Select and load the *Linked Pleading as Binder* preset
2. Click “Single folder” using the “Open input files” button
3. Select a folder that has the pleading stored as a pdf and its exhibits in a subfolder
4. Set the exhibit folder to the folder that contains the exhibits
5. Use the “Select output folder” to select the folder to store linked files
6. Click the Generate button

## Preset: Linked Pleading Attachment Binder

Binds a pleading together with its exhibits and creates blue underlined links highlighted in pink between references to the exhibits and the exhibits. The exhibits are added as attachments to the pleading and exhibit numbers are padded to allow natural ordering. Large exhibit numbering on the top right first page of each exhibit. Exhibit numbers read from the filename using the pattern *Bilag X*.

- Same as above, except select this preset.

## Preset: Set of Linked Pleadings

Converts references to exhibits and pleadings into links as described in section *Norwegian Civil Litigation Configuration*. In addition, this preset merges each pleading with its exhibits and creates page numbers in red. Note that in this configuration the target of the links is the individual files and not the merged ones. In order to merge them into one binder, merge the resulting files using the “Folder & subfolders” option.



- Same as above, except select this preset and in step 3 the folder should contain multiple sets of pleadings and their exhibits

## Preset: Stamp existing binder

Extracts exhibit numbers from bookmarks in an existing binder and stamps the first page of each exhibit with the relevant exhibit number.

1. Select and load the *Stamp existing Binder* preset
2. Select an existing binder using “Existing binder” on the “Open input files” button
3. Click the Generate button

## Input tab

### Input files

The input files are the files that you want to do something with - like merging, stamping and linking.

For operations that change multiple files, PdfClerk operates a folder basis. That means that an entire folder should be selected as input. For operations that only involve changing one file - like adding links to one file or adding stamps to a binder - the individual source file can be selected.

You select the input files using the dropdown button *Open input files*. The button has four options:

- *Single folder* - Use this to select a single folder as input. All pdf files in this folder will be loaded into the document list.
- *Folder & subfolders* - Same as above but also include any subfolders. Each folder is added as a bookmark by default.
- *Existing binder* - Opens an existing binder (document collection) and loads the bookmark hierarchy of that binder into the document list.
- *Single file to link* - Use this to select a single pdf or Word file that contains references to other documents in order to convert those references into links. It loads the selected file and all pdfs in the same folder into the document view. Only the selected file is marked for processing. In this mode, the setting “Process” in the document list controls whether or not a document is processed for references.
- *Single file to link w/subfolders* - same as above but will include subfolders and their files

### Item metadata

The metadata for each item is shown in the document list. By default the name of the item is the filename without extension for single documents and existing bookmarks for binders. Metadata can be used for stamps, bookmarks and table of contents.

Each item can be edited by clicking the text to be edited. You can navigate between entries using the up and down arrows.

Additional metadata columns can be added by clicking the plus sign in the floating menu to the right in the document list.



Metadata is read from filenames or bookmarks using the regular expression specified under [Metadata extraction](#). This is useful to extract information that is already present in filenames and bookmarks. Metadata can be imported and exported in [Excel format](#).

## Reorder documents

By default, the documents in the document list are sorted alphabetically.

The documents can be reordered by selecting one or more documents by clicking the circle to the left and then move the selected documents using the arrow links in the floating menu to the right.

To more efficiently reorder the documents, the following keyboard shortcuts are available. Select a document by clicking the circle to the left in the document list.

- ctrl+up/down: Move selected document(s) up or down
- alt+up/down: Move document selection (cursor) up or down
- shift+up/down: Expand or reduce selection

In addition, there are some [global keyboard shortcuts](#).

## Level

In a pdf, document bookmarks can be nested in a hierarchy. In PdfClerk this is achieved by changing the level of an item. Since this is a tree structure, the first item needs to have 1 as level and the increase in level from one item to the next cannot be more than one.

The level can be changed for multiple items at once by selecting several items as described in *Reorder documents* and then using the level arrow for one of the selected items.

## Bookmarks

Additional bookmarks can be added by selecting the wrench icon on the right of the document in the document list. The bookmark will link to the first page of the relevant document. By default, folders are added as bookmarks.

## Document properties

By clicking the wrench icon on the right of the document in the document list you can change the following properties of each individual document/bookmark:

- Change the name of the document. The name (and metadata) of each document and bookmark can also be edited from the documents list by clicking on the text that should be edited.
- Each document and bookmark can be excluded from stamping, from being included as a bookmark and from appearing in the table of contents
- The appearance of each bookmark can be adjusted by setting font colour and font style
- By default, all bookmarks are expanded. This can be changed on a per-level basis in the bookmarks section in the *Output* tab. In addition, this can be set per bookmark by selecting either *default*, *collapsed* or *expanded* in the document properties..



## Select the output folder

Use this button to select the output folder where files created by PdfClerk should be saved.

Click the “New folder” item on the ribbon in the *Select folder* dialog to create a new folder.

## Existing binders

If an existing binder is selected as input using the *Open input files* button, PdfClerk will work on that binder only and the document structure based on the bookmarks will be loaded into the document list. In this mode, you can add stamps to the binder, rebuild bookmarks, update the document properties or split the binder to individual files based on the bookmarks. The information to be used in stamps and bookmarks can be retrieved from existing bookmarks, edited in PdfClerk or loaded from Excel via the Load/save menu. You can also save bookmarks and their page numbers from an existing binder to Excel. This is useful for creating index files using Clerk Links.

## Include Exhibits in Subfolder(s)

PdfClerk can operate either on individual files and folders or one or more sets consisting of a main document and exhibits to the main document. To work with sets of main documents and exhibits turn on the setting *Include Exhibits in Subfolder(s)* under the *Input* tab.

- The typical use is for pleadings and their exhibits. Based on defined searches, PdfClerk can identify references from the pleadings to their exhibits: from one pleading to another or from one pleading to exhibits of other documents.
- Identified references are converted into pdf links to the relevant document or exhibit, either to the document or to a specific page within a document.
- Exhibits can be stamped with their exhibit numbers and other metadata.
- The pleadings and their exhibits can be merged together into one file.
- This function requires that all the exhibits are numbered using numbers.

## Document Filename Metadata Extraction

PdfClerk can extract information from the filename of each file, like exhibit number or date. Extraction is made using regular expressions and named capturing groups. The named capturing groups will be available as metadata for the document. Setting up regular expressions require some technical knowledge. Here are some examples that can be edited to your use case

- “(.\*)(?<exhibit>[0-9]{1,5}[A-z]{0,1})” - extracts 24 from *Lore ipsum bilag 24*
- “.{6}\_(?<dok>.+)” - extracts “Dokument 1” from *2001-02-03\_Document 1.pdf*
- “(?<date>.+?)\s{1,3}(?<name>.\*)(\s){0,3}\[(?<dok>[0-9]{1,5})?-(?<exhibitNo>[0-9]{1,5})\]?” - extracts date as “2019-03-08”, name as “Minutes from meeting” dok as “4” and exhibitno as “7” from *2019-03-08 Minutes from meeting [4-7]*

To run a new regular expression, press the refresh icon.

In order to extract different data, you can run different regular expressions with different group names. If you run a regular expression with the same group names(s) they will overwrite data in the existing columns.



## Exhibit Filename Metadata Extraction

Same as above but for the filename of the exhibits to documents. This option is only available when the option *Include Exhibits in Subfolder(s)* above is selected. Note that the regular expression needs to have a group with the name “exhibitNo” in order for the linking described under “Linking Documents” to work. Additional groups can be added to the regular expression and they will be accessible for stamps and bookmarks.

## Existing links

Any existing cross-links between pdf files merged into a bundle will be maintained. This is useful for links added to a document using Microsoft Word or a pdf editor. Matching of links with a relative path is made both with the source folder and the target folder. For links added via Microsoft Word you can also specify which page number to link to by adding a ‘#’ and the page number at the end of the filename (*myFile.pdf#3* will link to page 3)

## Output tab

### General settings

- *Create binder* – Whether or not PdfClerk should merge the documents in the document list. If the option *Include Exhibits in Subfolder(s)* is selected, one binder is created for each main document.
- *Start documents on odd page* – Ensures that all documents start on an odd page to allow double-sided printing
- *Add blank page before* – Use this option to add a blank page in front of other documents. Use in combination with stamps to create named separation sheets and similar. By default, a blank page is added in front of all documents, except the first document in the binder. Select a field from the dropdown list to only add blank pages in front of documents that have a value for this field including the first document. Optionally select a background color.
- *Create binders as attachment binders* – Creates a binder with the first document in the document list as the main document and the other documents in the document list as attachments to that document. If the option *Include Exhibits in Subfolder(s)* is selected the exhibits will be attachments to the main document.
- *Pad exhibit numbers in attachment binders* – This is a function to pad exhibit names with additional zeros to ensure that they are shown in the correct order in the pdf viewer software’s attachment pane. If there are 100 exhibits, this means that “Exhibit 1.pdf” is renamed to “Exhibit 001.pdf” etc.
- *Split binder by bookmarks* – (Binder Source mode only) Select this option to split a binder by its bookmarks. You can select which level of the bookmark should be used for splitting on. The default name of the file will be the name of the bookmark but this can be changed by specifying another value in the *Filename string* field. This works like stamps and bookmarks. Different levels of metadata can be accessed using the <|1-name> syntax (where 1 represents the level) and documents can be organised in subfolders by adding a slash (1) between different properties names. If there are file attachments to the binder, these will be extracted to the subfolder *attachments*.
- *Open target folder* – Opens the target folder after the document(s) have been created





- *Debug* – If this option is selected information is collected during execution to enable debugging. The information is stored in a folder named “debug” in the target folder. The information is only stored locally on the machine and not shared.

## Table of contents

PdfClerk can add a clickable table of contents to the binder.

- Select which template to use in the dropdown.
- For templates in Word format, you can edit the project template by clicking the wrench icon in the document.
- Once you have selected to add a table of contents it appears in the documents list. By default the name as it appears in documents list will be the title of the table of contents. You can change it by clicking on the name as it appears in the document list (default name is *Table of contents*)
- Once inserted the table of contents can be moved to another location in the documents list.
- Templates are Word or HTML documents located in the PdfClerk’s *Templates* folder (normally found in the Windows Documents folder).
- You can edit the templates or add your own by adding them to the folder. In addition to changing the layout and adding company stationery, you can configure them to output additional metadata fields for the documents like exhibit numbers and dates.
- The metadata in the documents list for the item *table of contents* is accessible as <m-name> in Word templates and as *meta.name* in HTML templates. This is useful in different scenarios like setting a title in the table of contents. If you update the name of the item *table of contents* to “List of documents in project Green” this text becomes the title of the table of the contents in the actual table of contents. You can access other metadata items by replacing *name* with the name of the metadata column.
- PdfClerk also has an alternative, even more flexible, method for creating tables of contents. See [Create Index](#).

## Stamps

PdfClerk can apply header and footer text to documents. This function is typically used to mark documents with exhibit numbers, page numbers and document titles.

There are two types of stamps:

- *Text stamps* – Inserts the text defined in the Text field.
- *Graphic stamps* – Inserts a graphic element in addition to the text. The graphic elements can be borders, solid or transparent backgrounds, logos, fixed text or other. If no text is provided in the *Text* field, only the graphic will be added. If an expression is added in the *Text* field that evaluates to an empty string, then the graphic stamp is not added.

PdfClerk ships with some predefined stamps located in the stamps folder. Additional stamps can be added in the stamps folder by adding a configuration file and an image (svg, png, jpeg, gif, bmp or tiff)

Each stamp has the following setting:



- *Position* – Defines where on the page the stamp should be placed
- *X margin* – The vertical margin. For top and middle positions, it is calculated from the top edge of the page. The bottom position is calculated from the bottom of the page. Negative values are allowed.
- *Y margin* – The horizontal margin. For right and centre positions, it is calculated from the right edge side. For left positions, it is calculated from the left. Negative values are allowed.
- *Font* – The font used for the text in the stamp
- *Font size* – The size of the font
- *Color* – Colors are specified using the color picker or alternatively using their hex values.
- *Text* – The text to be inserted. Any text enclosed in < > will be handled as metadata and looked up in the metadata table. Available values are listed below the stamps. Click the value to insert it in the active stamp.
- *First binder page* - Uncheck this box to not print this stamp on the first page of a binder.
- *First document page only* – Checking this box means that the stamp is applied on the first page of the relevant document only.
- *Use this* – Uncheck this box in order to not apply the stamp.

## Add Bookmarks

Select this option to add bookmarks to a new binder or to replace the bookmarks in an existing binder. Bookmarks in pdfs are sometimes referred to as outlines. Specify the metadata fields to use for the bookmark text in the *Build string* field. Note that adding bookmarks to an existing binder will remove any existing bookmarks. In order to remove all bookmarks in a binder without adding new click the *Remove existing* option.

By default all bookmarks are expanded, meaning that the whole bookmark tree is shown when the file is opened. Use the *Collapse Bookmarks* function to collapse one or more levels.

## Accessing Metadata

### Default and additional properties

The following default metadata items are available for use in stamps and bookmarks:

- <name>: the name of the bookmark or file without extension (“filename”)
- <filename>: the filename of the document (“filename.pdf”)
- <page>: the page number of the document
- <totpages>: the total number of pages in the document or binder
- <docpage>: the page number of the merged document (stamps only)

You can add additional metadata items by clicking the plus sign in the floating menu to the right of the document list in the *Input tab* view or by loading from Excel. Additional metadata items may typically be exhibit numbers, dates etc.

If you use the function *Include Exhibits in Subfolder(s)*, the following items are available for use in stamps and bookmarks for the exhibits:

- <namex>: the filename of the exhibit file without extension
- <filename>: the filename of the exhibit including extension



The <name> and <filename> and other items defined for the main document is still available.

## Formatting

For metadata which is numbers you can apply the following formatting:

- *add* - Use this function if you would like to start the pagination on another page number than 1. If you want to start the pagination of a binder on page 101 use the following format: <page|add:100>
- *pad* - Use this function to ensure that the number is represented by a fixed number of digits by adding leading zeros. Example “003” instead of “3”. Use the following format <page|pad:3>

## Empty values

If metadata for an item is unset or empty PdfClerk ignores the metadata item but inserts the rest of the string provided in the input box. This can be changed by enclosing the relevant part of the string in curly brackets. Examples:

- “Exhibit <exhibit>” inserts “Exhibit “ if there is no value for exhibit
- “{Exhibit <exhibit>}” inserts “Exhibit 1” if value “1” is given for exhibit and nothing if there is no value for the exhibit.

## Accessing metadata at specific levels

When accessing metadata items (like <name>), PdfClerk by default uses the highest metadata level for the relevant item. For stamps and when splitting binders, lower-level metadata items can be accessed by adding the letter “l”, the level of the metadata item to be accessed and a hyphen in front of the property name like <l1-name>. This is useful to include the full context on the header or footer of a document.

## Document Properties

For binders the document properties *title*, *author* and *subject* can be set under *Document Properties*. If the checkbox next to the value is not checked, the original values will be kept.

## Digitally Signed Documents

Digitally signed documents or documents locked for editing cannot be merged or stamped. To overcome this limitation PdfClerk creates a copy of the document with the digital signature removed. For the same purposes, PdfClerk ignores the protected status for some pdfs solely for the purpose of merging and stamping such documents.

## Autolink tab

### General

PdfClerk’s autolink feature works by identifying textual patterns within the text that refer to other documents. PdfClerk then adds hyperlinks from the identified points of reference to the document or specific page being referenced.

### Prerequisites



The most practical use is to create links from pleadings to their exhibits. In order for this to work, both the options *Convert References to Links* under the Linking tab and the *Include Exhibits in Subfolder(s)* under the Linking tab need to be enabled.

The pleadings should be placed in a folder on your computer with the exhibits to each pleading in subfolders. If you use the same number series for more than one pleading, all the exhibits can be placed in the same folder.

PdfClerk copies all pleadings and exhibits to the *Output folder* specified in the *Input tab*.

## How it Works

For each link, the software looks for three items of information

- *Document* – the main document that may or may not have exhibits
- *Exhibit* – an exhibit to a document.
- *Page number* – the page number in the referenced document or exhibit

In order for documents to be identified, the software needs to understand which names the document is referenced under. Each of these names is called an *alias*, and each document might have one or more aliases.

If the software identifies a reference to an exhibit without any document reference, it assumes that the reference references an exhibit of the document being processed.

If there are references in the document that should not be linked up, linking may be avoided by adding the names used to reference the documents that you want to exclude to the *Documents to Ignore* section under the File tab. A practical example would be a pleading that contains references to an agreement that has exhibits. By adding the word “agreement” to the *Documents to Ignore*, PdfClerk will avoid linking under those circumstances.

PdfClerk processes all PDF documents in the document list by default. If a document in the document list is linked to from other documents, but you do not want to process the document itself, uncheck the “Process” checkbox in front of the document. The document and its exhibits can still be linked to from other documents. To completely ignore a document, press “remove”.

If the software adds links that are not correct, the links can be removed in a PDF editing tool such as Adobe Acrobat (not Adobe Reader) or Foxit Reader (free, go to “Home”, “Links” in order to edit and remove links). Please note that if the option *Highlight links* are checked, the highlighting will remain after a link has been removed in the PDF editing tool.

## Other Autolink Settings

- *Underline Links* – If checked, the links will be underlined in the selected color. Width is the thickness of the line. Default is 1 pt.
- *Highlight Links* – If checked, the links will be highlighted in the selected color. Note that the highlight color becomes part of the pdf once applied and cannot be removed. Link highlighting may not work on some pdfs.
- *Open Links in New Window* – If checked, the links are opened in a window (or tab, based on which pdf viewer that is used) instead of reusing the same window.



- *Tag Links and Documents to Enable Binder Recognition* – This option enables relinking of documents once inserted into a binder by other software than PdfClerk. The function does not embed information about the text in the reference, but it is possible to see that a page has been referred to.
- *Adjust page link target* – Sometimes there is an offset between the page number stated in the reference and the page you would like to link to. The target page to link to can be adjusted by selecting the wrench icon on the right of the document in the Input tab and changing the value for *Adjust page link target*. Both positive and negative numbers can be used.

## Multi binder collections

Sometimes a logical document collection with sequential page numbers is spread over multiple pdf binders (split binders). PdfClerk can still map references to the correct page in the correct binder if you add the page ranges and page adjustments in the options for the entry in the documents list (click the wrench icon). The binders in the logical document collection should have the same alias.

Example: All exhibits are collected in two binders named *ExhibitsPart1.pdf* with pages 1-1000 and *ExhibitsPart2.pdf* with pages 1001-2000. In order to convert references like *Exhibits page 12* and *Exhibits page 1222* to links to the relevant page in the corresponding binder do the following:

- Add “Exhibit” as an alias for both files
- For *ExhibitsPart1.pdf* set “Start link page” to 1 and “End link page” to 1000
- For *ExhibitsPart2.pdf* set “Start link page” to 1001 and “End link page” to 2000
- For *ExhibitsPart2.pdf* set the “Adjust page link target” to -1000.

## Searches

This list contains the searches that PdfClerk performs to identify references to other documents.

There are two text boxes for each search item. The first textbox is for naming the search for later reference. The second textbox is for the actual search.

The main building blocks of the searches are the base patterns {d}, {e} and {p} which represents a document, an exhibit, and a page respectively. If an exclamation mark is added after the letter, it means that the search will only match against text that contains this element. One example of when this is useful is to enable a search that finds references both with and without page numbers. There needs to be at least one mandatory base pattern or additional text in each search.

The system uses regular expressions to find references in the text. Each of the base patterns is defined using regular expression syntax and are defined in the *Base patterns* setting.

In addition to the base patterns, the searches can be supplemented with additional words and symbols. The base patterns are replaced with their definition when the searches are executed.



In addition to the mentioned base patterns, there is also a building block for white space. The exhibit number and other metadata are extracted from the file name of the exhibit using the regular expression specified under “Document filename metadata extraction” in the Input tab. The exhibit number should be found in the regex group “ex”.

A phrase of text in a document may match against more than one of the searches. In most cases, a reference to an exhibit of the document itself may be simply referred to as “Exhibit B” while a reference to an exhibit of another document may be referred to as “Exhibit B to Pleading of 3. January”. A search for exhibits without qualifying document references will also find the qualified references. For that reason, the searches should be executed in order, starting with the most specific first. The searches are done in the same order as they are listed in the interface.

### *Advanced*

- In most cases, PdfClerk is able to recognize patterns also when they are spread on two lines, and also if the page has multiple columns. PdfClerk will also handle line breaks within single words as long as words are separated with a hyphen or underscore
- Regex lookaheads and lookbehinds in a search pattern need to be wrapped in a comment with the text “prtlgnore”. Example lookahead for “:” is `(?#prtlgnore)(?=:)(?#prtlgnore)`
- By default PdfClerk only allows one link on each section of text. That means that if there are two searches that match the same text, only the search that runs first is used. Add `#i#` in front of the search to override this behaviour. This could be useful in order to link to both page 100 and 155 in the reference “Report A page 100 and 155”.
- The `#doclist#` in the definition of the document base pattern is replaced runtime with a list of the document aliases (separated by |)

## Norwegian Civil Litigation Configuration

The Norwegian Civil Litigation Pack Configuration identifies the following patterns:

- Bilag 3
- Bilag 3 side 5
- Stevningen side 4
- Stevningens bilag 3 side 4
- Stevningen bilag nr 3
- Bilag 4 til stevningen, side 4

The software understands both “side”, “s” and “s.” to indicate page numbers.

When looking for documents the software accepts alternative forms of the alias, meaning that “stevning” will also match “stevningen” and “stevningens”

The software is case-insensitive, meaning that “Stevning” and “stevning” is treated in the same way.



## Load/Save tab

### Projects

PdfClerk allows presets and documents to be saved and loaded. When using the Load & Save dialog user should select what category/ies of project data should be saved or loaded to enable easy reuse between projects.

Files are saved as .clerk files which are associated with the application. The internal format is JSON, meaning that more technical users with particular needs could create and alter PdfClerk project files using a text editor.

### Excel format

PdfClerk can read and write document metadata from and to Microsoft Excel format (.xlsx). This enables the use of a familiar and powerful tool as part of the pdf creation workflow.

The columns and format used are different depending on whether the input is separate files or an existing binder.

#### Separate files

Document names and other metadata is stored in a sheet named "Documents" with the column names:

- Level - the bookmark level
- Filename - the full name of the file - including extension, but not the path
- Name - the name of the file - not including the extension

Additional metadata is saved in additional columns with the name of the column as the header. The path to the documents is stored in the "Config" sheet as "SourceFolder".

An example file for a document assembly can be found [here](#).

#### Existing binder

Bookmarks and other metadata is stored in a sheet named "Binder" with the column names:

- Level - the bookmark level
- Name - the name of the bookmark
- Additional metadata is saved in additional metadata columns with the name of the column as the header

The full path and filename to the binder are stored in the "Config" sheet as "BinderToBeUpdated".

An example file for a binder can be found [here](#).

#### General

The order of the columns or formatting does not matter. But there should not be any empty columns or rows before non-empty ones.



Reading and writing to Excel is not supported if the option Include *Exhibits in Subfolder(s)* is selected.

## Tools tab

The Tools section provides various tools to work with existing files and binders.

### Convert Clerk Links

This function translates Clerk Links to pdf links. Clerk Links can be used to create custom tables of contents or other documents with direct links to specific pages in pdf binders. The custom documents can be created as Microsoft Word or HTML documents. The Clerk Links is inserted in the custom document as a hyperlink using the format described below.

The default way to create a table of contents in PdfClerk is to automatically create and insert one when creating a binder using the insert [Table of contents](#) function on the *Output* tab. With Clerk Links you can create linked documents with other data sources and full control over how the resulting document will look.

Clerk links can be inserted manually into the source document or you can use the function [Create Index](#) (described below) to combine a template and data from a Microsoft Excel datasheet into a document.

The created pdf can either be a separate pdf that links to other pdfs or be included in a binder referencing other pages in that binder.

The Clerk Links support the following link formats:

- `clerk://MyBinder.pdf` - Creates a link to page 1 in *MyBinder.pdf*
- `clerk://MyExcelSheet.xlsx` - Creates a link to open the Excel sheet (or any other file type)
- `clerk://MyBinder.pdf?page=27` - Creates a link to page 27 in *MyBinder.pdf*
- `clerk://subdir/MyOtherBinder.pdf?page=3` - Creates a link to page 3 in *MyOtherBinder.pdf* located in the subfolder *subdir*.
- `clerk://n/?page=3` - Creates a link to page 3 in the same file
- `clerk://n?anchor=someName` - Creates an anchor (named destination) that can be linked to
- `clerk://n/?dest=someName` - Links to an anchor (named destination)
- `clerk://n/?embed=MyExcelSheet.xlsx` - Embeds *MyExcelSheet.xlsx* into the binder as an attachment. The link will open the file. If the path is not absolute, PdfClerk looks for the file in the source folder

The pdf file created should be placed in the same folder as the files that are linked to or if paths are provided (like the example link to *MyOtherBinder* above) in the parent directory.

When you want to include a file in a binder that links to other parts of such binder you should just convert the file from Word to pdf and then merge the pdf using PdfClerk. You can convert from Word format to pdf directly using Word or you can use this tool in PdfClerk and check *Do not convert links*.





A possible workflow is to create links in Excel using the Excel Hyperlink function for each relevant document, then copy the links from Excel to Word as a table, and then format the table in Word before activating the links using PdfClerk.

In order to be able to convert a file from Word to pdf format, MS Word version 2010 or newer needs to be installed on the computer. If the file has been converted from Word format to pdf format using another tool, PdfClerk can update the links as described above.

(The “/n/” part (the host part) of the URL is optional for files and can be added in order to comply with the URL specification: *clerk://n/MyBinder.pdf*)

## Create Index

You can include a table of contents at the same time as you create the binder using the [Add Table of Contents](#) in the *Output* tab. If you want to link to multiple binders, add a table of contents to an existing binder or for other reasons need full control over the output, you can use the Create Index function found under the Tools tab.

This function works by combining entries from an Excel sheet into a Word or HTML template. If you want to create a table of contents for one or more existing binders with bookmarks you can export the bookmarks to Excel using the [save to Excel function](#) found under the Load/Save tab.

The *Create Index* function creates an editable .docx or HTML file with *Clerk Links* on the relevant entries. This allows you to edit the file before you proceed to convert it to PDF using the “Convert Clerk Links” function.

### The Excel datasheet

The Excel datasheet should have a datasheet named *Catalog* with the columns with the following column headers

- *file* – the file that the entry should link to. If empty it links to the stated page in the same file
- *page* – the page in the document that the entry should link to
- *level* – The hierarchy level of the entry. If the level is not provided entries will be handled as level 1
- *title* – the text used for the link. Could be *title* or any other value set in the template

Any other columns are handled as metadata and inserted into the document if referenced by the template by the column title as the name.

PdfClerk saves and reads all values either as strings, booleans or numbers. In order to display dates consistently, PdfClerk requires that dates should be saved as text in the Excel sheet. If you have a date stored as a date you need to convert it to text format to properly display it in PdfClerk. The easiest is probably to copy the column containing the date values to Notepad, change the format of the column in Excel to *Text* and copy them back from Notepad. Alternatively, you can use [Excel's Text function](#).

If the Excel workbook only has one datasheet, this datasheet is used.



## Using a Word template

The Word template should have a table in this format:

Column title (optional)	Column title (optional)
<title><#1#>	<page>
<title><#2#>	<page>

- In the table properties, Alt Text, the Title of the table should be set to *index*
- The row with <#1#> is the template for level 1 entries, the row with <#2#> is the template for level 2 entries, and further - up to a maximum of 10 levels.
- The text enclosed in <> is replaced with the corresponding text from the column with the same name in the Excel datasheet. Any number of columns can be added.
- By default, all text becomes clickable links to the relevant entry. If you do not want a particular entry to be a link you can either make sure that there is no data in the file or page columns or explicitly set the item as no-linking by adding “:0” to the property name like <title:0>.
- An example template file is provided in the [example files](#).
- The text will be styled with the style that is applied to the relevant cell in the template. Only one style can be applied to text in a table cell. The style should be based on a paragraph style.
- If the Excel workbook has a datasheet named *Meta*, the first two columns of this datasheet are handled as key and value pairs that are accessible in the template using <m-key> where *key* is the name assigned in the first column and the value written to the template is the corresponding value in the second column.

## Using an HTML template

PdfClerk also supports creating index files using HTML templates. You can edit the supplied standard templates or create your own templates. Note the following:

- Data is filled into the templates using the [Liquid template language](#).
- Data loaded from the Excel sheet is exposed to the template in the object “catalog” with all lowercase letter properties.
- By adding “\_L” after the name of the property it will be translated into a link. Alternatively, the link uri is available as “uri” for each item.
- If the Excel workbook has a datasheet named *Meta*, the first two columns of this datasheet are handled as key and value pairs that are accessible in the template using *meta.key* where *key* is the name assigned in the first column and the value written to the template is the corresponding value in the second column.
- PdfClerk supports full-page coloured background without margins using the custom in-file CSS property “-clerk-page-background-color” which accepts heximal color codes.
- PdfClerk supports extracting headings as bookmarks when converting html files to pdf by adding the custom in-file CSS property -clerk-extract-headings-as-outline: *all*.



- You can use ordinary HTML, CSS and Javascript to render the templates, and load external stylesheets, images and fonts.

## General

The *Create Index* function creates a Word/HTML document based on the provided template with [Clerk Links](#) on all items. The Word/HTML file can be edited before you convert it to pdf using the function *Convert Clerk Links*. The format of the Clerk Links is described [above](#).

Use the option to *Adjust target page number* if you need to offset the page numbers provided in the Excel file. This could be handy if you plan to insert the created index at the beginning of the binder that the index links to.

## Multiple links per row

PdfClerk supports alternative links for the same entry ( excel sheet row). Add additional link columns in the Excel sheet (page, file) by adding a number after the name (page2, file2).

In table-based Word-templates the alternative links are available by adding :X where X is the number used as postfix for the name in the Excel sheet. Like <title:2>.

In HTML-templates the alternative links are available in two separate ways:

- by adding “\_Lx” to the property name where x is the number used at the end of the column name in the Excel sheet. Like “title\_L2.
- as additional .uri properties (uri2, uri3 etc)

## Reconnect links in binders

If the setting *Tag links and documents to enable binder recognition* under the tab *Linking* is enabled when references are recognized and converted to links in a document, information about the link's target is embedded in the document and the target files. If documents are combined in binders via software other than PdfClerk, this information will remain intact. However, the links will need to be reactivated. By using this function, you can reactivate the links in the binder.

You can also use this function to reconnect links if you have changed the name of one or more of the target file a autolink link refers to.

## Transform links

Use this function to update link target and link type of links in an existing pdf document.

The function searches for external links starting with the pattern provided in the *Link address starts with* input box and changes this pattern to the pattern provided in the *Replacement* input box. Links that do not start with the pattern are not changed.

If you select an input file without selecting any changes PdfClerk will look through the file for links and list out their addresses. This is handy when determining what patterns should be changed.



You can specify that only links of a certain type should be transformed. PdfClerk will then only search for links of the type specified. You can specify that links should be converted to another type of links. PdfClerk maintains direct page links when converting between GotoR links and web links.

By selecting the option *Include bookmarks*, PdfClerk will also search in the bookmarks.

PdfClerk currently supports transforming links of the types *GotoR* (for pdfs), *Launch* and *Web* (URI).

You can process multiple search and replace patterns at once by separating them using the “|”-symbol. As an example putting “folder1|folder2” in the search box and “folderA|folderB” in the replace box will replace all links starting with *folder1* with links starting with *folderA* and links starting with *folder2* with *folderB*.

## Global keyboard shortcuts

PdfClerk has the following global keyboard shortcuts

- ctrl+g = Generate
- ctrl+s = Save project
- ctrl+shift-o = Open project
- ctrl+1 = Open Input tab
- ctrl+2 = Open Output tab
- ctrl+3 = Open Linking tab
- ctrl+4 = Open Tools tab
- ctrl+5 = Open Load & Save tab
- ctrl+6 = Open Log tab
- ctrl+7 = Open Help tab

## Legal stuff

### Data privacy

PdfClerk is a desktop application installed on the user’s computer. The application does not share any information about the documents processed with any external systems.

### License

PdfClerk is the property of Protosys AS and its licensors. Third-party software included in the software is listed on the *Help tab* of the application

Use of PdfClerk is only permitted for license holders and is then limited to the usage, the number of users and time period described in the license grant. PdfClerk may only be used for its intended purpose as described in the documentation.

The purpose of a trial license is to test the software, and the output PDFs are marked as not for production use. Use of a trial license for production purposes is not permitted.

*Last updated 2021-08-31 for version 1.6.7*

