

Guidelines for PDF files

Instructions for files to be published online via the TU Berlin institutional repository

Overview of requirements

The aim is to ensure the long-term readability and usability of online publications. PDF files must meet the following requirements:

- **Content can be copied and indexed:** The text-based content is machine-readable and accessible, i. e. the content must not be embedded as an image, it is searchable, it can be selected and copied to other applications.
- **No password protection:** The file must not be password protected.
- **Fonts embedded:** The fonts used in the file are embedded.
- **PDF properties:** The document properties must at least contain information on the title, author and keywords; further information is desirable (e. g. language and licence information). The bookmark bar is displayed. The page layout is set as single-sided and continuous.
- **PDF bookmarks:** At least the top-level headings must be bookmarked to allow readers to easily navigate the file.
- **PDF/A:** To ensure long-term digital preservation, the publication must be submitted in PDF/A format. It is recommended to submit as PDF/A-2a; the variants PDF/A-2b, PDF/A-2u, PDF/A-1b or PDF/A-1a are also acceptable.

You will find more information on how to implement these requirements below. (Please note that this is not a step-by-step guide; where possible, tips for different word processing applications are included).

Further notes on files

- **File name:** (Re)name the file as follows: name_firstname.pdf (in case of multiple authors or editors: name_et al_year.pdf or similar). The file name shall not contain spaces, umlauts or other diacritics.
- **File upload:** The file is uploaded to the repository directly; to do so a [registration](#) is needed. Users with a TU account can self-register with their email address. Please do not send files via email.
- **Personal data:** Please handle personal data carefully and check whether certain information is necessary. For theses, it is not necessary to include a CV, statutory declaration or similar. WE advise against including a scan of your signature.

Overview of requirements	1
Further notes on files	1
(1) Presets & PDF properties	3
Accessibility (recommendation)	3
Content can be copied and indexed	3
Document properties	3
Embedded fonts	5
No password protection	6
Page layout	7
PDF bookmarks	8
(2) PDF/A	9
Create PDF/A with LaTeX	10
Create PDF/A with Word	11
Create PDF/A with LibreOffice	13
PDF/A validation	14
Converting to PDF/A	17

(1) Presets & PDF properties

The PDF requirements can usually be met by changing settings or certain options in your word processing application, when creating the document and/or when exporting it as a PDF. Some of the settings can be changed later, notably with Adobe Acrobat Pro (licence required).

Accessibility (recommendation)

The file should be accessible (i. e. also usable for people with visual impairments). Please use the regular features of your word processing application to specify alternative texts for content that is not text (illustrations, charts and graphs, tables, etc.) to ensure accessibility.

LaTeX	Information by the TeX User Group: https://www.tug.org/twg/accessibility/ List of packages with accessibility support: https://ctan.org/topic/accessible
Libre Office	Recommendation: tab "Tools" → "Accessibility Check"
Word	Recommendation: Tab "Review" → "Check Accessibility"
Adobe Acrobat Pro	Check: "Tools" → "Accessibility"

Content can be copied and indexed

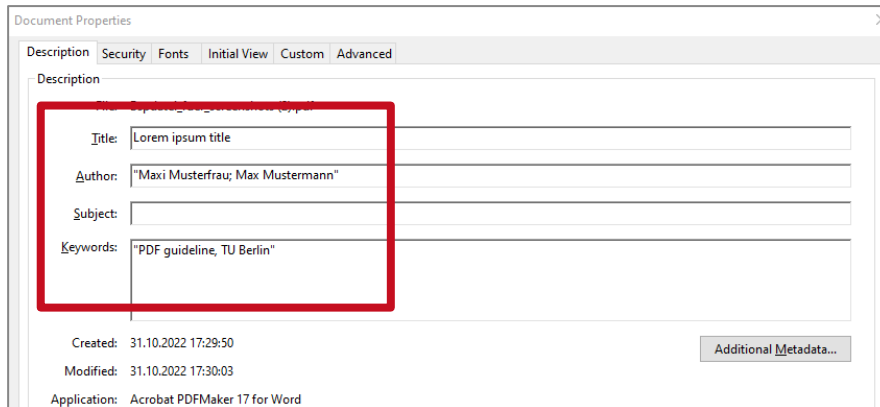
The text-based content is machine-readable, i. e. the content must not be embedded as an image, it is searchable, it can be selected and copied to other applications.

Document properties

In most word processing application, the most important metadata can be entered directly. At least the following must be included: title, author and keywords. It is recommended to state the main language of your document to ensure accessibility.

LaTeX	Use hyperref package: add entries for (at least) "pdfauthor", "pdftitle" and "pdfkeywords" in "hypersetup" Recommendation: additional entry "pdflang"
Libre Office	Tab "File" → "Properties" → "Description" → entries for "title" and "keywords" ("author" not available)
Word	Tab "File" → "Info" → "Properties" → entries for "title", "tags" (corresponds to keywords) and "author"
Adobe Acrobat Pro	Check/Add entries: tab "File" → "Properties" (or: Ctrl + D) Recommendation: state the document language in tab "File" → "Properties" → "Advanced" → "Reading Options" (at the bottom)

Check available entries – e. g. in Adobe Acrobat Pro resp. Reader: “File” → “Properties” → tab “Description” (resp. open “Additional metadata”)



Embedding rights information is recommended (e. g. free license; for information on Creative Commons licenses see <https://www.tu.berlin/en/go11005/>).

LaTeX

Use [hyperxmp](#) + [hyperref](#) packages: add entries for “pdfcopyright” and “pdflicenseurl” in “hypersetup”

```
pdfcopyright={CC BY 4.0},  
pdflicenseurl={https://creativecommons.org/licenses/by/4.0/}
```

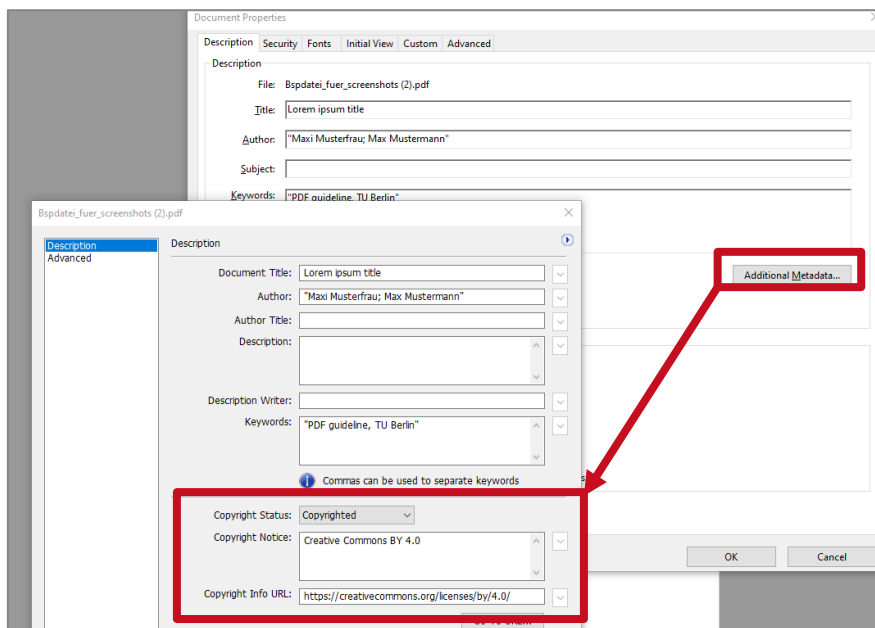
Adobe
Acrobat Pro

Add/Change metadata: “File” → “Properties” → tab “description” → open “Additional metadata”

The following information should then be available:

- Copyright Status (always: “Copyrighted”)
- Copyright Notice (e. g. “CC BY 4.0”)
- Copyright Info URL (e. g. <https://creativecommons.org/licenses/by/4.0/>)

Check or add rights information in Adobe Acrobat Pro: “File” → “Properties” → tab “Description” → open “Additional metadata”



Embedded fonts

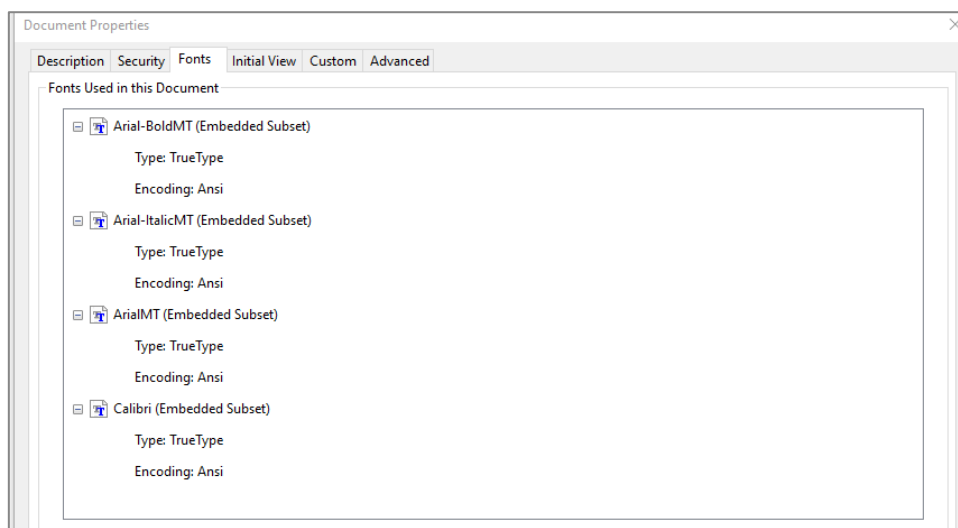
You can usually choose to embed fonts during PDF export (cf. [\(2\) PDF/A](#)). The following presets for your word processing application are recommended:

LaTeX	Fonts are embedded by default when using PDFLaTeX. Recommendation: use fontenc package to make sure ligatures and the like are displayed correctly (cf. Create PDF/A with LaTeX)
Libre Office	Recommendation preset: tab "File" → "Properties" → "Font" → tick box for "Embed fonts in the document"
Word	Recommendation preset: tab "File" → "Options" → "Save" → at the bottom tick box for "Embed fonts in the file" (Mac: "Preferences" → "Save" → "Font Embedding")

Difficulties may arise when including graphic elements (charts, diagrams, etc.) that were created with help of other applications. Check the export settings in your graphics application.

Check the PDF's document properties whether all used fonts are listed with the remark "embedded".

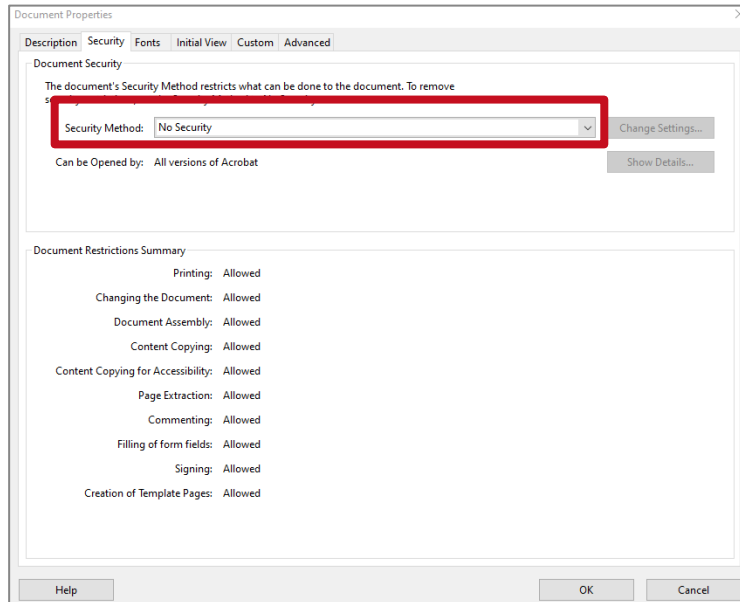
Check in Adobe Acrobat Pro resp. Reader: "File" → "Properties" → tab "Fonts"



No password protection

Check the document properties to make sure there are no restrictions on the file's use.

Check security settings in Adobe Acrobat Pro resp. Reader: "File" → "Properties" → tab "Security" → "No security" is selected for "Security method"



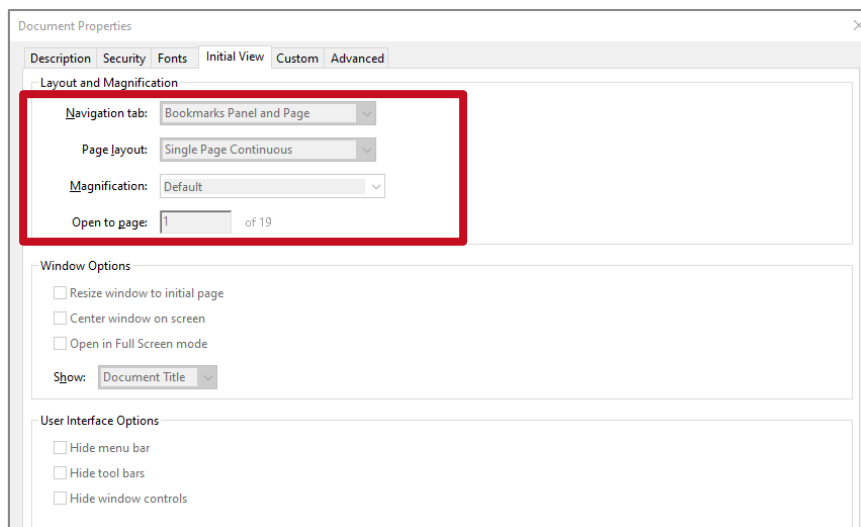
Page layout

The following presets are recommended for opening files: The bookmark bar is shown. The page layout is set as single-sided and continuous. The file is opened in standard size and on the first page.

LaTeX	Use hyperref package <code>\usepackage[bookmarksopen=true]{hyperref}</code>
Libre Office	Option for PDF export "Initial View" (cf. Create PDF/A with LibreOffice)
Word	Setting unknown/ not possible
Adobe Acrobat Pro	Add/Edit options: "File" → "Properties" → tab "Initial View"

Check in Adobe Acrobat Pro: "File" → "Properties" → tab "Initial View"

- Navigation tab: "Bookmarks Panel and Page"
- Page layout: "Single Page Continuous"
- Magnification: "Default"
- Open to page: "1"
- "Window Options" → Show: "Document Title"



PDF bookmarks

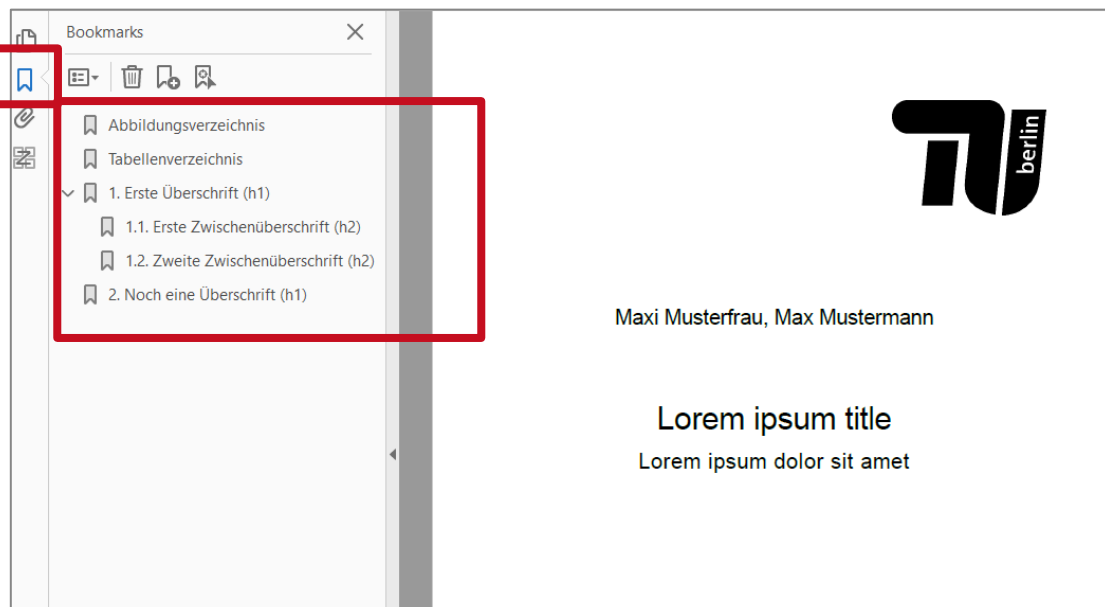
At least the top-level headings must be bookmarked to allow readers to easily navigate the file. Use styles (formatting templates) in your word processing application to create headings and generate the bookmark bar automatically.

LaTeX	Use hyperref package <code>\usepackage[bookmarks=true,bookmarksopenlevel=1]{hyperref}</code>
LibreOffice	Use styles for headings, when exporting PDF use the option "Export outlines" (cf. Create PDF/A with LibreOffice)
Word	Use styles for headings, when exporting PDF use the option "Create Bookmarks" (cf. Create PDF/A with Word)
Adobe Acrobat Pro	Set bookmarks manually afterwards: <ol style="list-style-type: none">1. Open bookmark bar (bookmark icon in left navigation pane)2. Mark the desired heading in your document3. Right-click "Add bookmark" (or: Ctrl + B)4. Continue for all headings (at least top-level headings)5. Use drag & drop to indent subheadings to the correct level; then collapse the open bookmarks to the top level

Check in Adobe Acrobat Pro resp. Reader: bookmark icon in left navigation pane, or "View" → "Show/Hide" → "Navigation Panes" → "Bookmarks"

The bookmarks should be collapsed to show the top-level headings.

Bookmarks can be edited (renamed, collapsed, created, change order, ...) in Adobe Acrobat Pro.



(2) PDF/A

PDF/A is an ISO standard (ISO 19005-1 or 19005-2) specialized for use in archiving and to ensure the long-term digital preservation of documents. A PDF/A-compliant document can visually reproduce the content in the same way on different computers with different operating systems and in different printing environments.

Submission as PDF/A-2a is recommended; the following variants are also accepted: PDF/A-2b, PDF/A-2u, PDF/A-1b or PDF/A-1a (PDF/A-3 is currently not accepted.)

- PDF/A-1b and PDF/A-2b (Level B, "Basic") provide reliable visual reproducibility; they ensure the minimum requirements for an archivable document.
- PDF/A-1a und PDF/A-2a (Level A, "Accessible"), besides reliable visual reproducibility, ensure character mappings to Unicode and a hierarchical document structure and tagging.

Free and proprietary tools are available for creating, editing and validating PDF/A-compliant files. Some, but not all, word processing applications offer direct export in PDF/A format.

- Notes on creating a PDF with [LaTeX](#) | [Word](#) | [Libre Office](#)
- Notes on [PDF/A validation](#) (checking standard compliance)

If you convert a PDF file (retrospectively) into PDF/A format: Please note that every conversion can produce errors. Therefore, please carefully check the file you converted to PDF/A: Are figures/graphs, formulas and special characters (e. g. umlauts, diacritics) reproduced correctly? Can the text be copied without errors?

- Notes on [converting to PDF/A](#)

Create PDF/A with LaTeX

Creating a fully valid PDF/A file may not be possible with every LaTeX setup. There is no best practice approach, also due to the large number of TeX packages and options.

Please make sure that all fonts and colour information are embedded. Always update to the latest TeX distribution. Out-of-date distributions are one of the most common sources of error.

We recommend using the [pdfx](#) package (`\usepackage[a-1b]{pdfx}`) to create a PDF which requires only little to no post-editing. Alternatively, you can use the [pdftex](#) package.

The PDF/A standard specifies, among other things, that document properties must be included in a certain format; this can be achieved using e. g. the [hyperxmp](#) package.

To embed colour information, you may need to include an ICC profile file (download e. g. via <http://www.color.org/srgbprofiles.xalter>).

We gladly recommend consulting the helpful manual "[How to create a PDF/A file with LaTeX](#)" by Jarmo Niemelä (dating from October 2021).

Further tips for LaTeX users

To ensure that all characters can later be copied correctly from the PDF (frequent source of errors: ligatures such as ff, fi ...), the [fontenc](#) package should be included (with the **T1** option, most European languages are supported).

We also gladly include hints on mathematical typesetting from one of our authors:

"To typeset mathematical formulae the following packages were used: `\usepackage{amsmath}`, `\usepackage{amssymb}`, `\usepackage{dsfont}`, `\usepackage{mathabx}`, `\usepackage{commath}`. The mathematical symbols were displayed correctly in the regular PDF. However, when converting to PDF/A, they were removed. This problem was solved by removing the package `mathabx`."

Recommended settings & options:

```
\usepackage[T1]{fontenc}

\usepackage{hyperxmp}

\usepackage[bookmarks=true,bookmarksopen=true,bookmarksopenlevel=0]{hyperref}

\hypersetup{
  pdftitle={TITLE},
  pdfauthor={first author, second author, ...},
  pdfkeywords={KEYWORD, KEYWORD, ...},
  pdflang={en},
  pdfcopyright={CC BY 4.0},
  pdflicenseurl={https://creativecommons.org/licenses/by/4.0/}
}
```

Create PDF/A with Word

There are several ways to export a document as PDF. To create a PDF/A file, it is recommended to use the plugin PDFMaker. On Windows, the print menu can also be used.

Add basic information in the document properties; at least the following must be included: title, author and keywords (cf. [Document properties](#)).

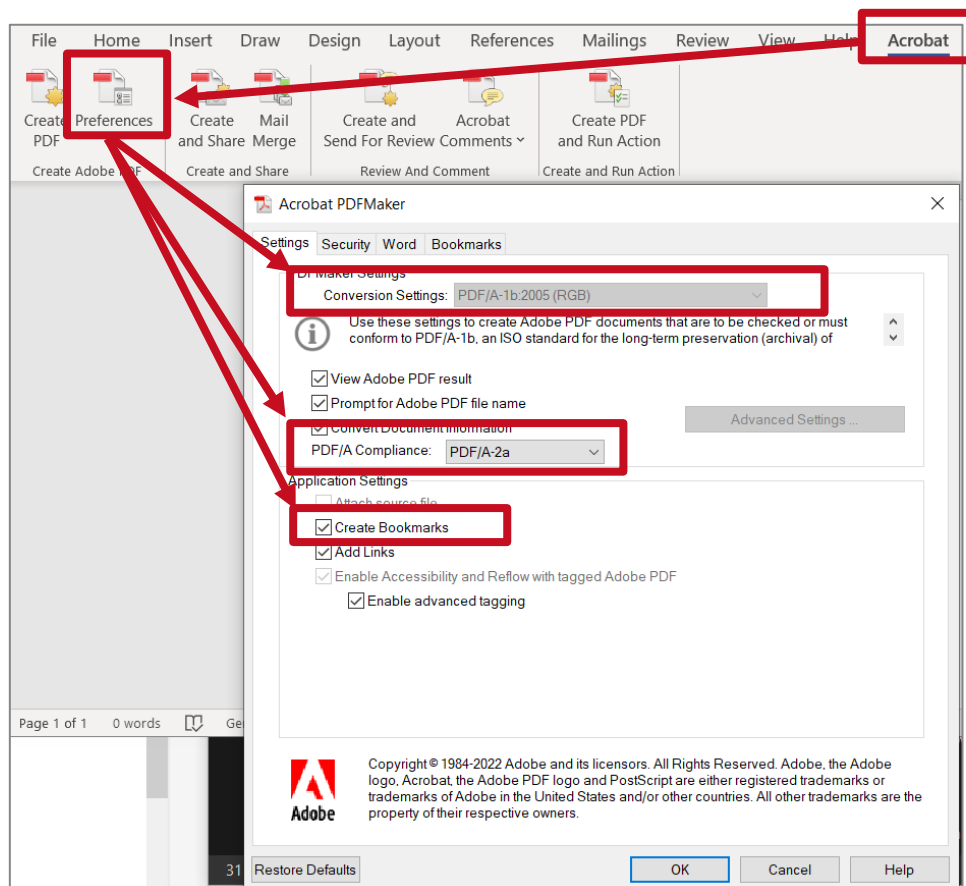
Use styles (formatting templates) to create headings and hence generate the bookmark bar automatically.

(1) PDF Maker

The plugin [Acrobat PDFMaker](#) is free of charge and is available once Adobe Acrobat Reader or Pro is installed (tab "File" → "Options" → "Add-Ins"); when installing Acrobat this plugin is usually activated by default. You can then adjust the default settings for PDF export.

In Word tab "Acrobat" → "Preferences" → "Settings"

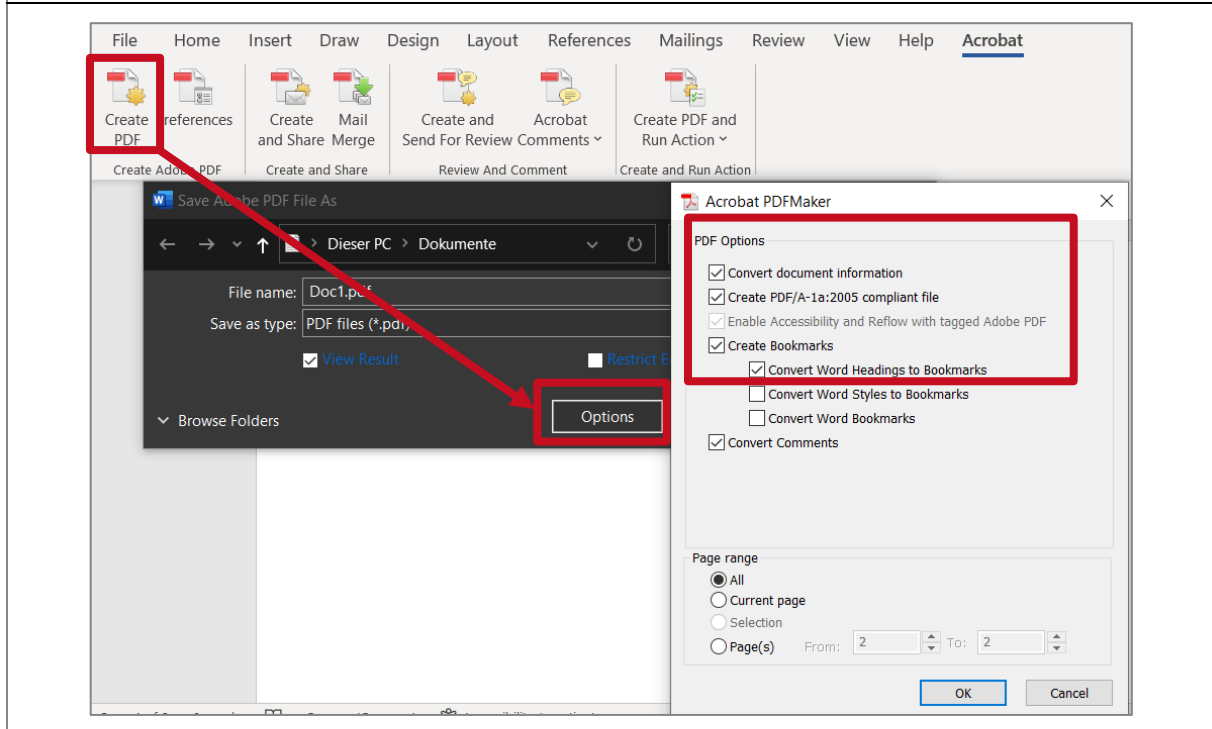
- Conversion settings: "PDF/A-1b:2005(RGB)"
- PDF/A Compliance: "PDF/A-2a" (or: PDF/A-2b, PDF/A-1b, PDF/A-1a)
- Tick box for "Create bookmarks"



When creating a PDF, check the correct settings in the "Options".

In Word tab "Acrobat" → "Create PDF" → "Options"

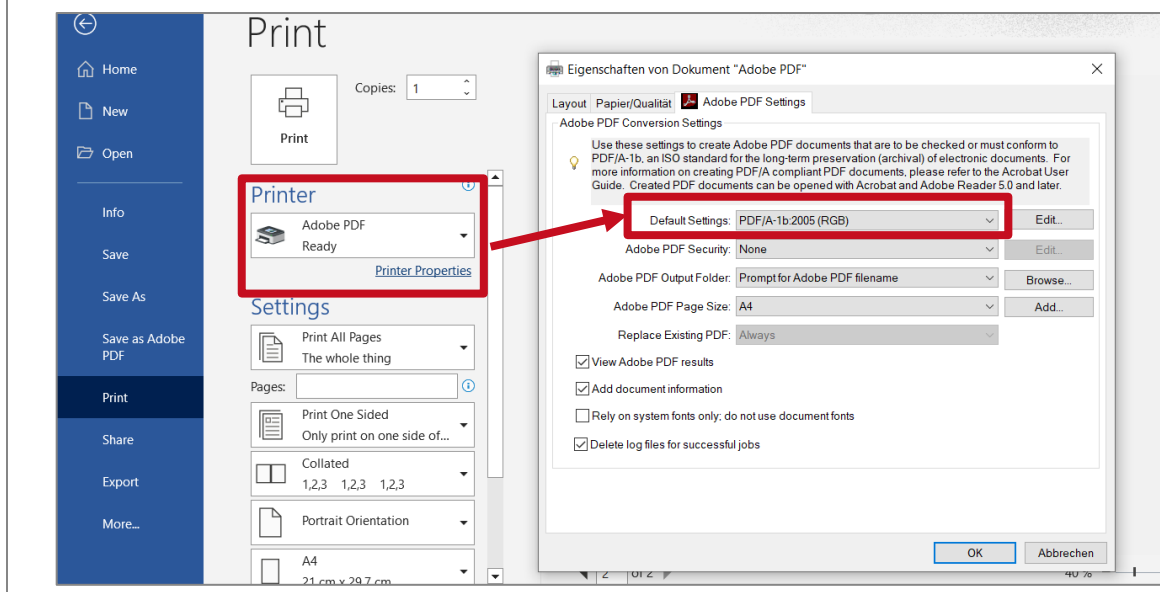
- Tick box for "Create document information"
- Tick box for "PDF/A-1a:2005 compliant file"
- Tick box for "Create Bookmarks" → "Convert Word Headings to Bookmarks"



(2) Using the print menu

In Word tab "File" → "Print" → Printer: "Adobe PDF" → "Printer Properties" → Default Settings: "PDF/A-1b:2005 (RGB)"

Please note: There is no known setting to include formatted headings in the bookmark bar of the PDF file.



Create PDF/A with LibreOffice

From version 3.0, Libre Office supports [direct](#) export as a PDF/A file. It is possible to open/edit files in the MS Word formats DOC or DOCX.

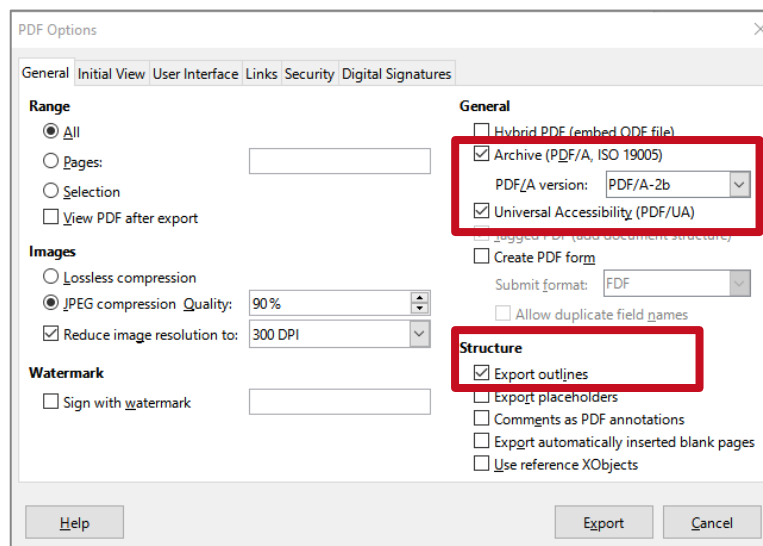
Add basic information in the document properties; at least the following must be included: title, author and keywords (cf. [Document properties](#)).

Use styles (formatting templates) to create headings and hence generate the bookmark bar automatically.

In LibreOffice tab "File" → "Export as" → "Export as PDF"

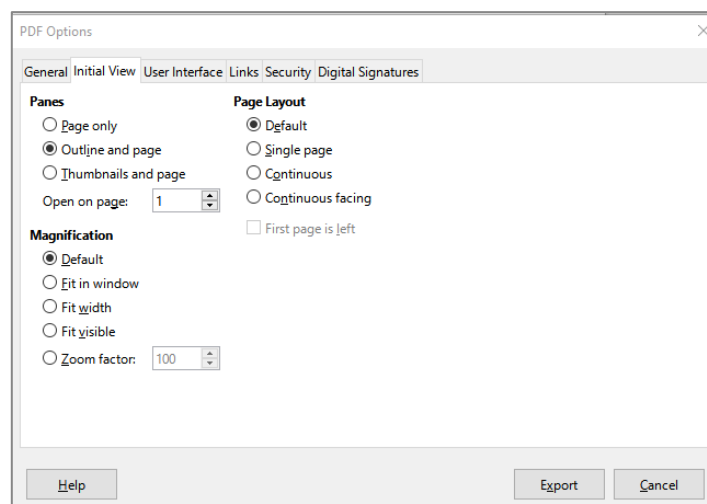
(1) Tab "General"

- Tick box for "Archive (PDF/A, ISO 19005)" (recommendation: PDF/A-2b)
- Recommendation: tick box for "Universal Accessibility (PDF/UA)"
- Structure: tick box for "Export outlines", to include formatted headings in the PDF bookmark bar
- Images: optimized for JPEG compression (quality at least 90 %, resolution not less than 300 dpi)



(2) Recommended settings in tab "Initial View"

- Panes: "Outline and page"
- Page layout: "Default"
- Magnification: "Default"



PDF/A validation

During PDF/A validation the file is checked for compliance with the ISO standards for PDF and PDF/A. Different validation tools may produce different results because the underlying standards are interpreted differently. The final validation of your file is done by the University Library staff.

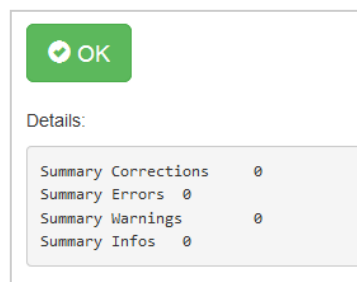
PDF/A validation with PDF/A Conversion (option for members of TU Berlin)

Members of TU Berlin can use [PDF/A Conversion](#) (based on the software [pdfaPilot](#) by Callas) to check files for compliance with the PDF/A standard and/or to convert them to PDF/A.

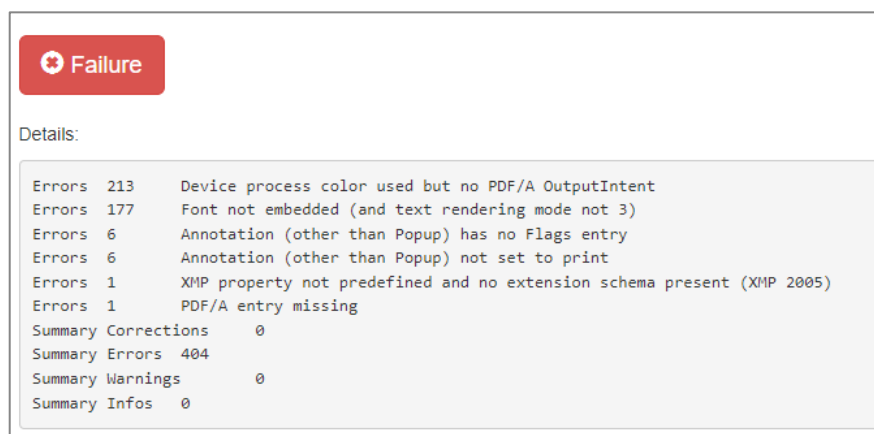
- Login with TU account: <https://conversion.ub.tu-berlin.de/>
- See also detailed info: <https://conversion.ub.tu-berlin.de/pdfa/help.php?lang=en>

Log in to [PDF/A Conversion](#) and upload a file → “Validate as PDF/A” → choose respective PDF/A standard → “Validate”

The result is “OK” if the validated file is a long-term archivable PDF/A file. Additionally, please check whether the text can be copied correctly to other applications (e.g. copy and paste to text editor or Word).



The result is “Failure” if the validated file is not compliant with the selected PDF/A standard. Details are provided for errors (e. g. font not embedded, device-dependent colours used). In this case, it is recommended to convert the file to PDF/A (cf. [Converting to PDF/A](#)).



After you have finished processing the file, please delete it.

PDF/A validation with veraPDF

veraPDF is an open source software to validate PDFs. veraPDF is still being developed, but can already be installed (Windows, Mac, Linux) and used.

- About veraPDF <http://verapdf.org/>
- Download prototype <http://verapdf.org/software/>

PDF/A validation with Adobe Acrobat Professional

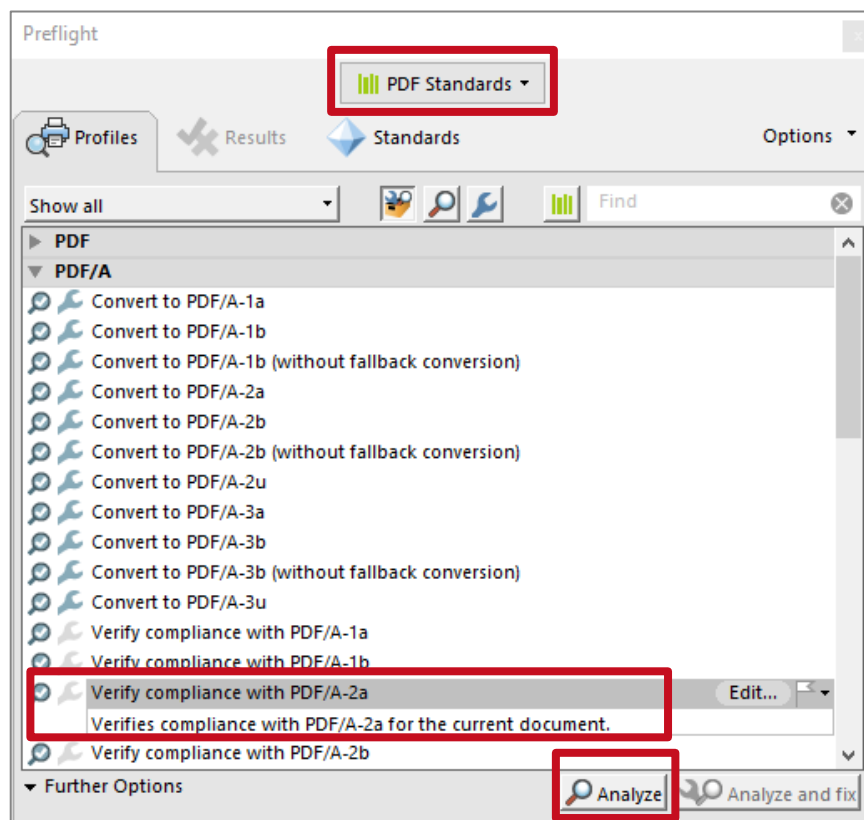
In Adobe Acrobat Pro (licence required) the Preflight tool can be used to validate files and correct errors.

The following description refers to the version Acrobat Pro 2017.

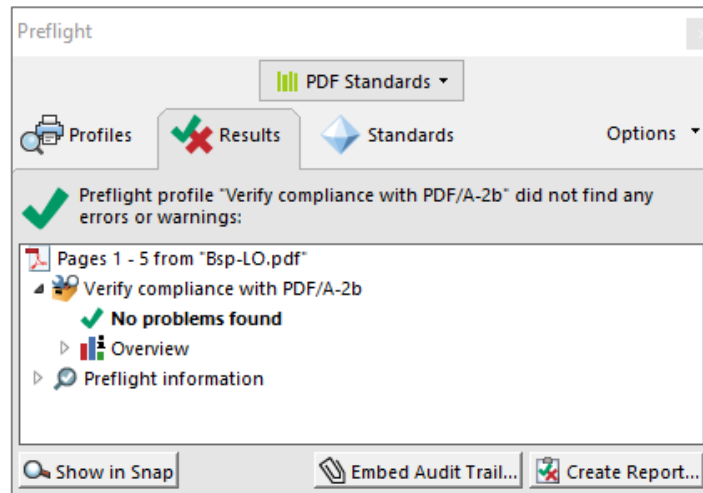
Open file in Adobe Acrobat Pro → “View” → “Tools” → “PDF Standards” → “Preflight” (or: Shift + Ctrl + X)

In tab “Profiles” and section “PDF/A”, choose the respective standard for validation → “Analyze”

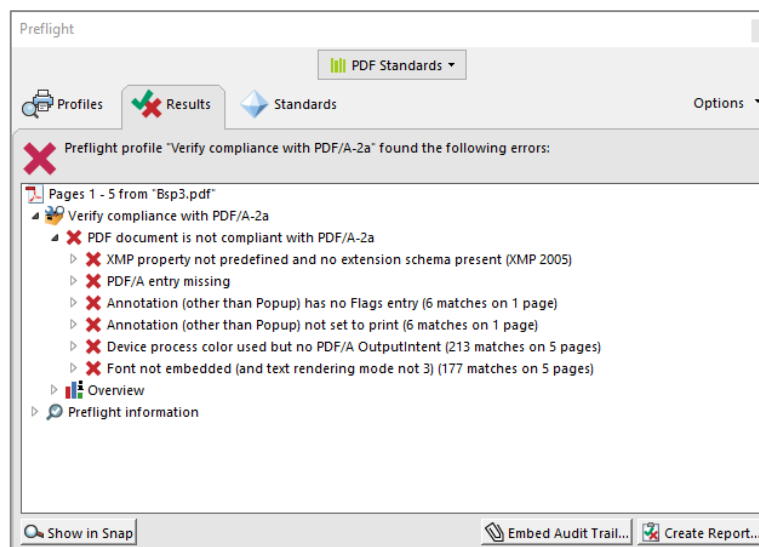
(“Essentials” might be pre-selected in the upper menu resulting in numerous options missing in the list → change to “PDF Standards”)



The result is “No problems found” if the validated file is a long-term archivable PDF/A file. Additionally, please check whether the text can be copied correctly to other applications (e.g. copy and paste to text editor or Word).



The result is „PDF document is not compliant with ...“ if the validated file is not compliant with the selected PDF/A standard. The detected errors are displayed (e. g. font not embedded, device-dependent colours used); if needed details on certain errors can be expanded (using the triangular arrow leading the respective error). In this case, it is recommended to convert the file to PDF/A (cf. [Converting to PDF/A](#)).



Converting to PDF/A

Generally speaking, a direct export in PDF/A format in your word processing applications is advisable.

If you convert a PDF file (retrospectively) into PDF/A format: Please note that every conversion can produce errors. Therefore, please carefully check the file you converted to PDF/A: Are figures/graphs, formulas and special characters (e. g. umlauts, diacritics) reproduced correctly? Can the text be copied without errors?

Using PDF/A Conversion (option for members of TU Berlin)

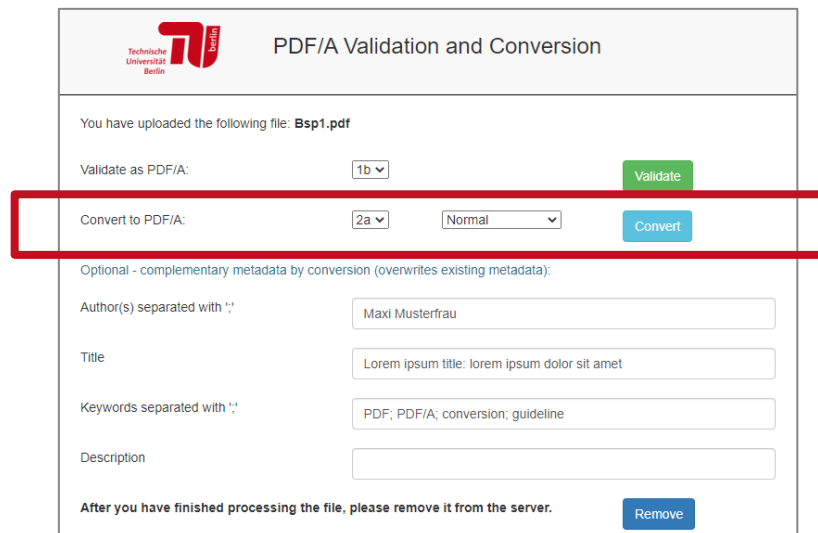
Members of TU Berlin can use [PDF/A Conversion](#) (based on the software [pdfaPilot](#) by Callas) to check files for compliance with the PDF/A standard and/or to convert them to PDF/A.

- Login with TU account: <https://conversion.ub.tu-berlin.de/>
- See also detailed info: <https://conversion.ub.tu-berlin.de/pdfa/help.php?lang=en>

Which of the available conversion options will be successful depends on the original document and the type of errors. Therefore, a general recommendation is hardly possible. If "transparencies" are one of the errors during validation, there has so far been a high success rate with conversion to PDF/A-2b.

Log in to [PDF/A Conversion](#) and upload a file → "Convert to PDF/A" → choose respective PDF/A standard → "Convert"

In addition, metadata can be embedded in the document properties (information on authors, title, keywords, description). Please note that any existing metadata will be replaced.



Technische Universität Berlin

PDF/A Validation and Conversion

You have uploaded the following file: **Bsp1.pdf**

Validate as PDF/A:

Convert to PDF/A:

Optional - complementary metadata by conversion (overwrites existing metadata):

Author(s) separated with ";":

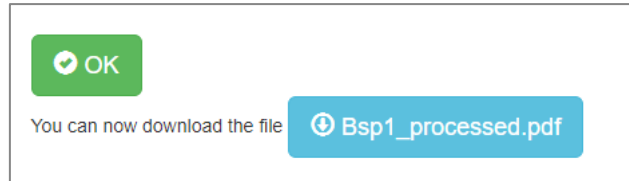
Title:

Keywords separated with ";":

Description:

After you have finished processing the file, please remove it from the server.

The result is "OK" if the validated file is a long-term archivable PDF/A file. The suffix "_processed" is appended to the file name and the file can be downloaded.



The result is "Failure" if the validated file is not compliant with the selected PDF/A standard. Error details are provided. Please try to correct the sources of error in the original document



If the errors cannot be corrected in the original file and your usual word processing application, try a new conversion with different parameters if necessary:

- 1) Select a different PDF/A standard in the first menu (e.g. 2b or 1b instead of 2a).
- 2) If that fails, select a different option in the second menu. Try the options in this order for further conversion attempts:
 - a. "Normal"
 - b. "Reconversion"
 - c. Not recommended: "Pages to images"
 - d. Not recommended: "Docs to images"

(The options "Pages to images" and "Docs to images" should be avoided – and if used, this should be followed by OCR handling (text recognition) – as they have an impact on how text can be copied from / searched for in the file.

After you have finished processing the file, please delete it.

Converting to PDF/A using Adobe Acrobat Pro

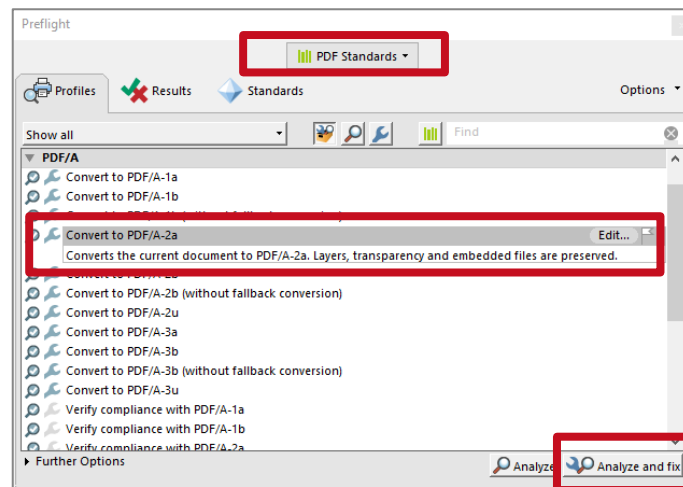
In Adobe Acrobat Pro (licence required) the Preflight tool can be used to validate files and correct errors.

The following description refers to the version Acrobat Pro 2017.

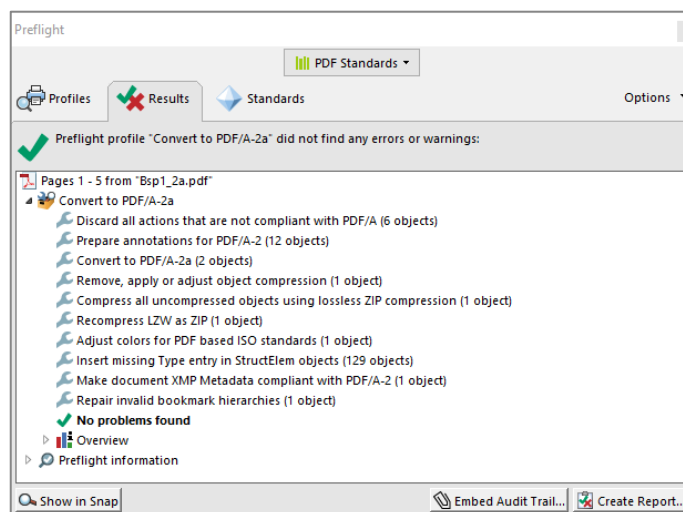
Open file in Adobe Acrobat Pro → “View” → “Tools” → “PDF Standards” → “Preflight” (or: Shift + Ctrl + X)

In tab “Profiles” and section “PDF/A”, choose the respective standard for validation → “Analyze and fix”

(“Essentials” might be pre-selected in the upper menu resulting in numerous options missing in the list → change to “PDF Standards”)



You will be prompted to save the file under a new name. The preflight report gives information on performed corrections:



The preflight report also provides information about sources of errors if the file could not be converted to the selected PDF/A standard.