

INSTRUCTIONS FOR CREATING PDF/A-COMPLIANT FILES FOR ONLINE PUBLISHING AT THE TU BERLIN

We only accept PDF documents meeting the criteria for one of following archiving formats:

- **PDF/A-1b:2005 (RGB)** or **PDF/A-2b:2005 (RGB)** (PDF optimized for online publishing)
- **PDF/X-3:2002** (PDF optimized for print production)

Since 2005, PDF/A is a stable, internationally accepted ISO standard for long-term archiving. A PDF/A- or PDF/X-compliant document may reproduce its visually identical content on different computers with different operating systems and under different printing environments. For example, all fonts and color information used will be embedded in the file.

Free as well as proprietary tools for production, processing and validation of PDF/A- and PDF/X-compliant files are available. Examples are Adobe Acrobat Professional (not to be confused with Adobe Reader), PDF-T-Maker/Ghostscript and PDF-24. Many text processing and layout software suites already provide direct exporting these formats, such as Open Office, Adobe InDesign (PDF/X only) or MS Office 2010 and newer.

This document gives short instructions for the creation and validation of PDF files using a few selected tools:

1. MS Word 2010 and newer
2. MS Word 2007
3. Adobe Acrobat Distiller
4. Open Office/Libre Office
5. LaTeX
6. Software-independent PDF/A creation
7. Validating the PDF/A
8. Converting a file to PDF/A subsequently

ATTENTION

Please validate your created PDF via simple Copy & Paste: you should be able to copy most parts of the PDF into another program (i. e. MS Word or a text editor) without rendering the text unreadable.

Keep in mind that every conversion to another format may cause changes to the document. Please double-check your document after conversion to PDF/A. This is especially important for non-raster graphics (such as diagrams) and scientific formulae.

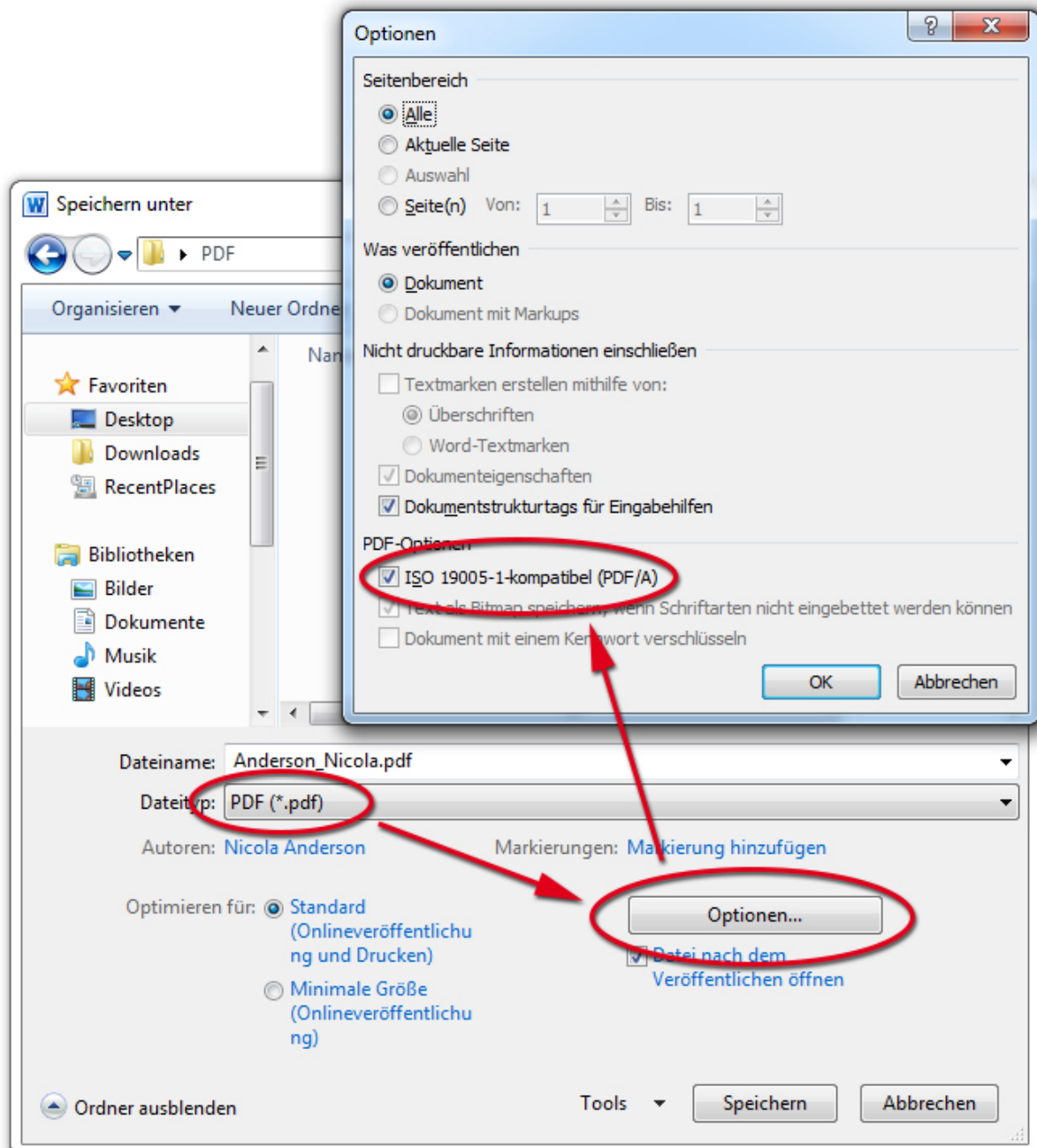


This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

1. MS WORD 2010 AND NEWER

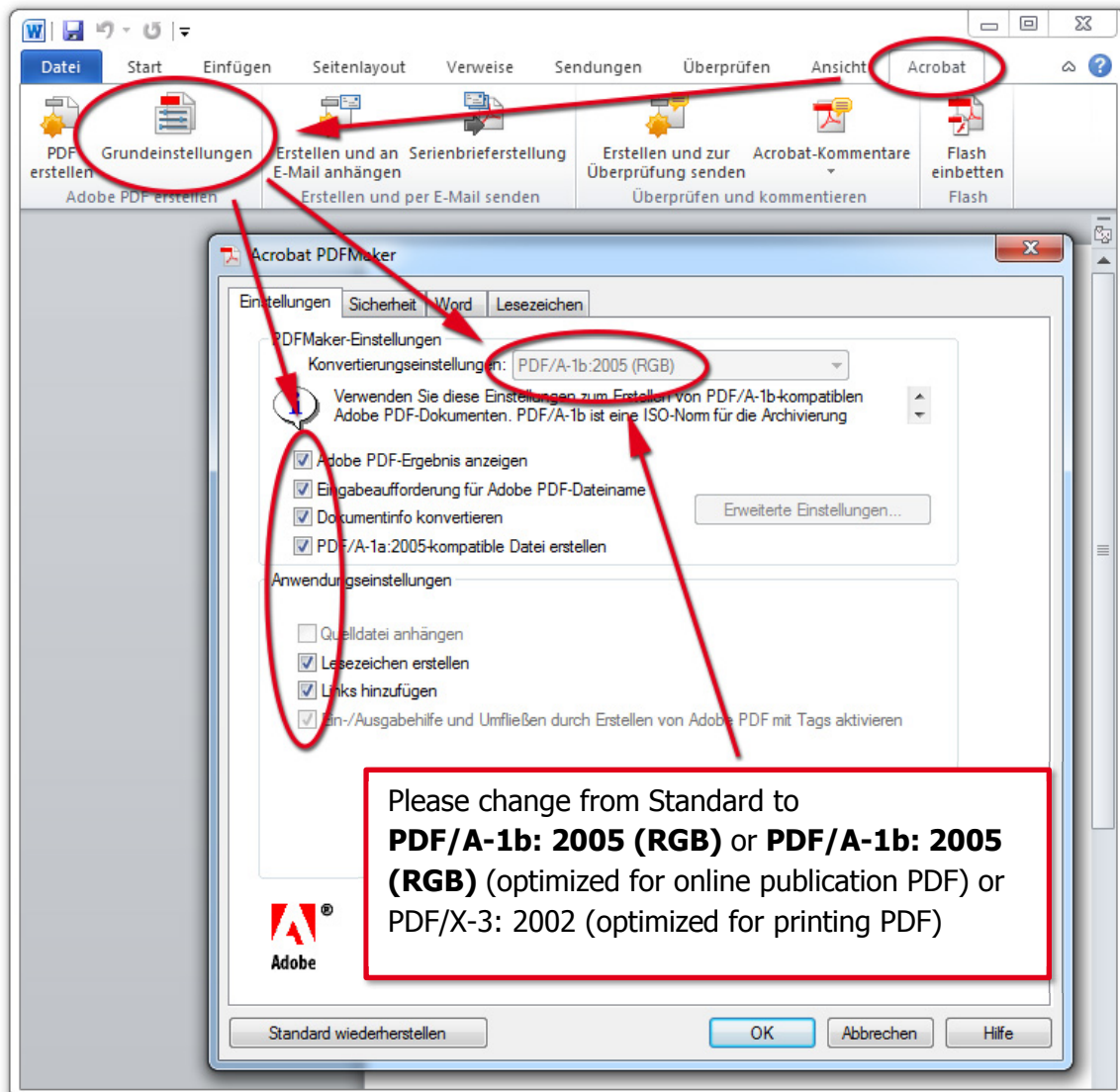
1.1. Save as PDF/A

Word 2010 does not require installation of external programs for saving PDF/A-1 files. However, when saving the file (Go to File, **Save As** and **select PDF as the file format**), make sure the box **"ISO 19005-1 (PDF/A)"** is checked.



1.2. PDF Maker

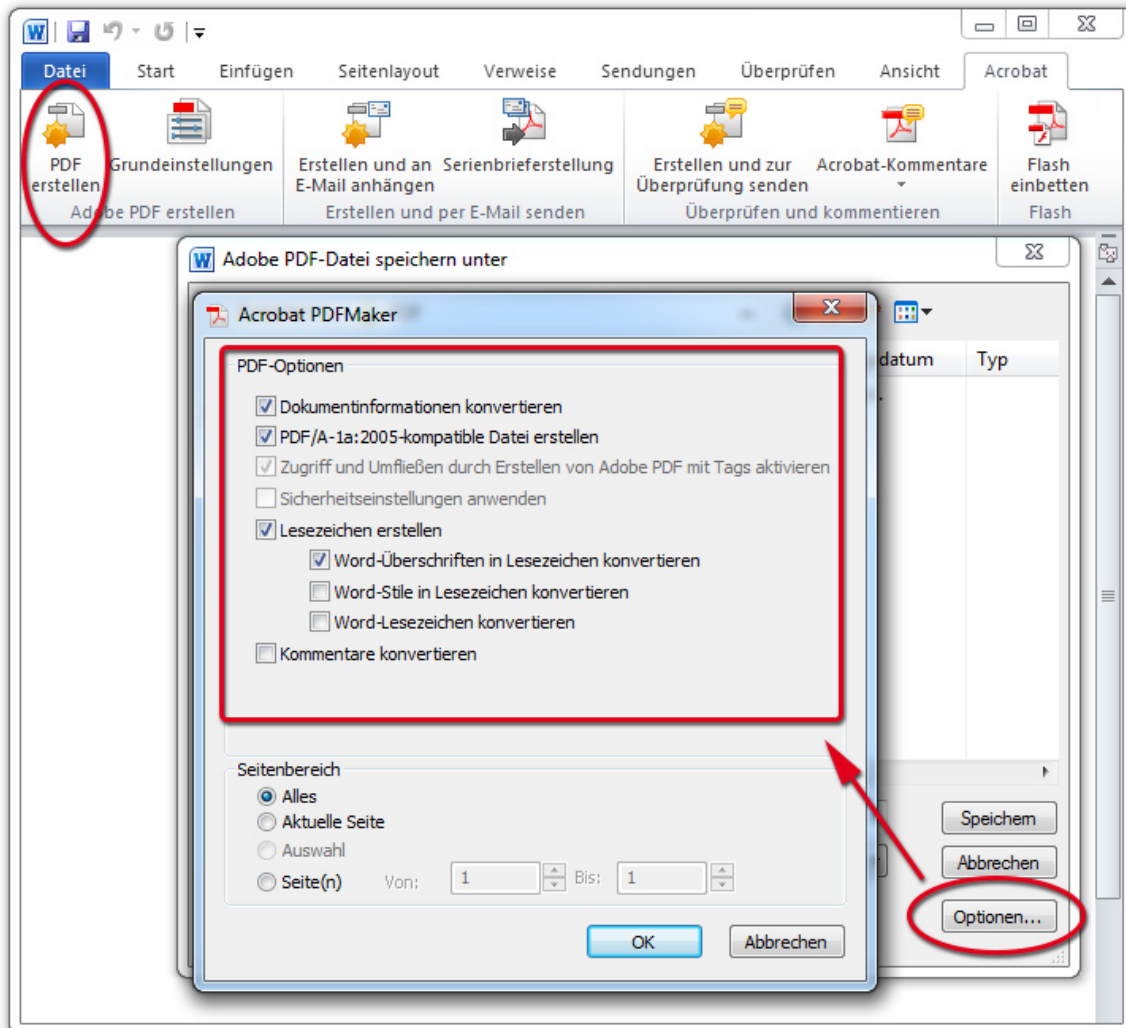
If both MS Word and Acrobat are installed on your system, you can use the Acrobat PDF Maker tab and check the following settings:



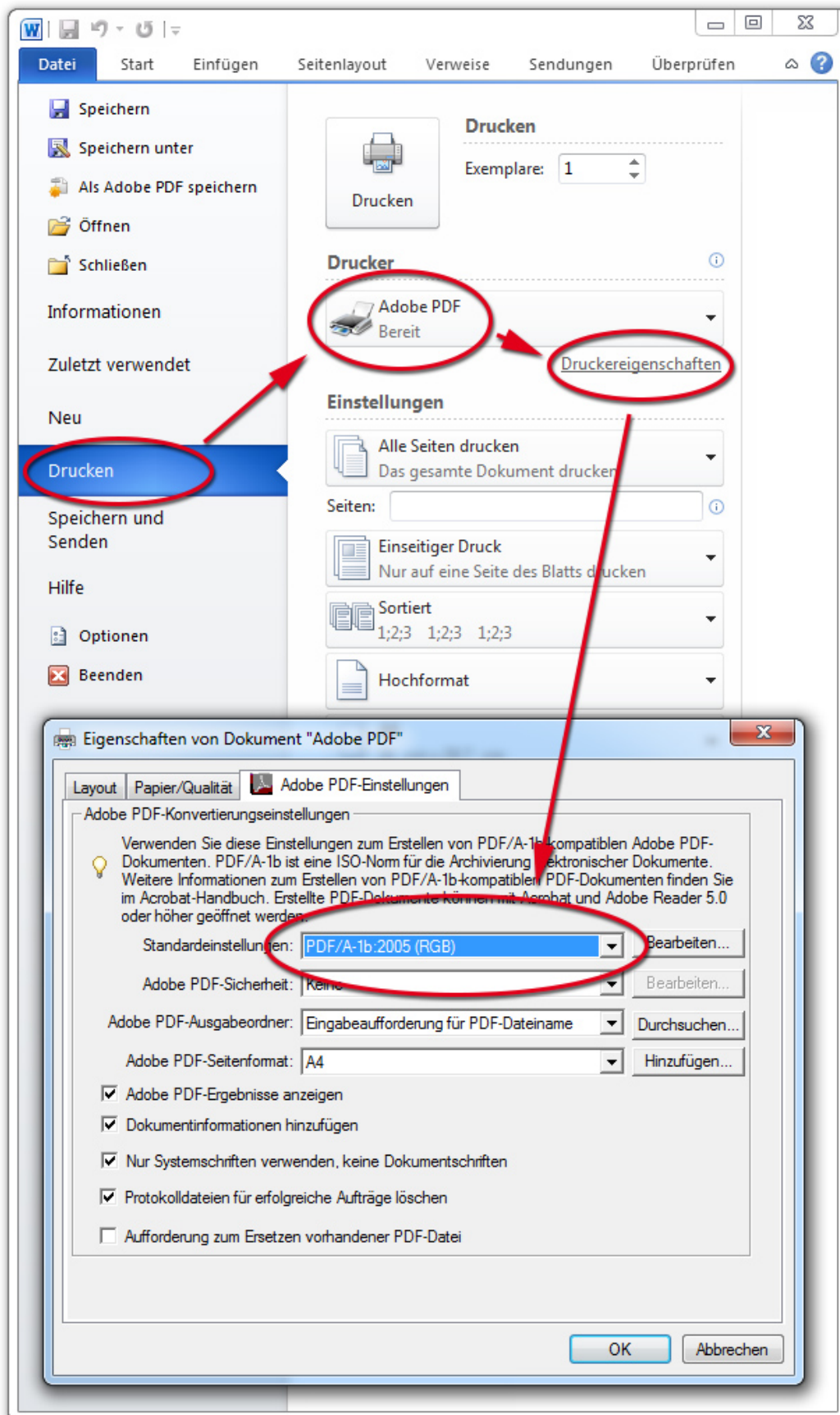
Note regarding bookmarks:

If you have used heading styles in MS Word as generally recommended, make sure to check "Create bookmarks". This way, the formatted headings are included in the bookmarks of the PDF file.

Before you save your file as a PDF/A file, you can check the correct settings at "options" again:



1.3. Using the print menu



2. MS WORD 2007

Using MS Word from version 2007, you have the possibility to export Word documents as PDFs with PDF/A settings directly. For this, the installation of Service Pack 2 or the add-in "Save as PDF or XPS" is required, available at:

http://www.chip.de/downloads/Add-in-fuer-Office-2007-Speichern-unter-PDF-oder-XPS_23066541.html

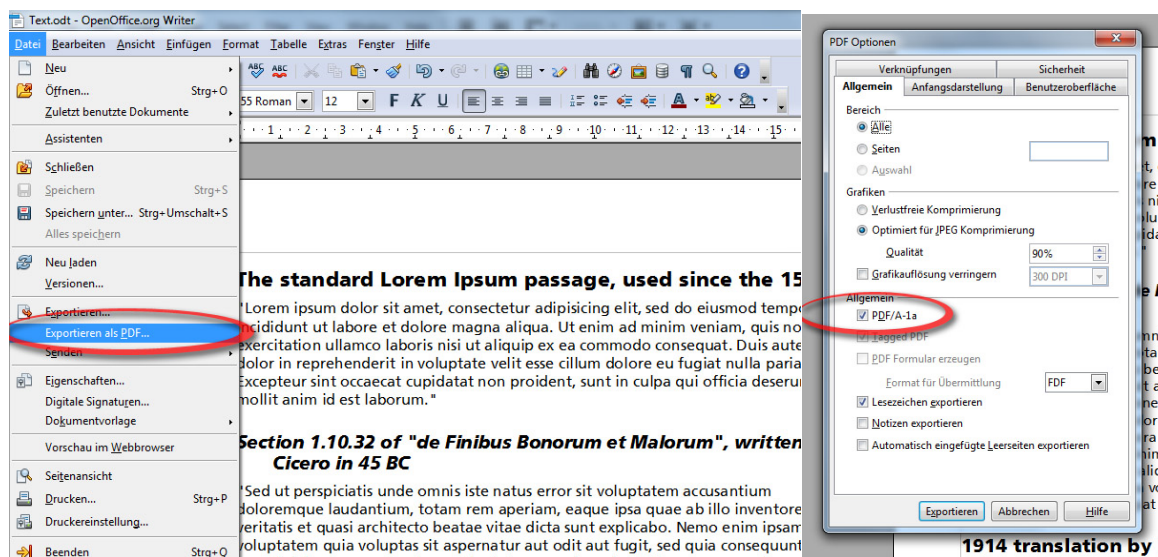
In the "Office" button select (top left) "Save as" -> Set as XPS or PDF publish-> Options: Select for PDF options "ISO 19005-1 compliant (PDF/A)"

3. ADOBE ACROBAT DISTILLER

Edit -> Preferences -> Category: Convert to PDF -> Edit Settings: "Adobe PDF Settings": PDF/A-1b: 2005 (RGB) or PDF/A-2b: 2005 (RGB)

4. OPEN OFFICE/LIBRE OFFICE

OpenOffice/Libre Office supports PDF/A natively since version 3.0. Therefore, it also lends itself to save Word files as PDF, if there is no Word 2010 version or later installed. It reads .docx files and can them as PDF/A Save when "PDF/A" is set in the "Export as PDF" menu.



File -> Export as PDF ... -> check **PDF/A-1** (graphics options: optimized for JPEG compression quality: 90% minimum, resolution not below 300 dpi)

5. LATEX

Creating a PDF/A format may not be possible with every LaTeX setup. Please keep in mind that all fonts and color information have to be embedded when converting LaTeX files to PDF! Always update to the latest TeX distribution. Out-of-date programs are one of the most common sources of error.

Include the package *fontenc* (`\usepackage[T1]{fontenc}`) to make sure that all characters (ff, pf ...) can be copied correctly from the PDF later on.

Do not forget using *hyperref* as well as *bookmarksnumbered=true* (`\usepackage[bookmarksnumbered=true]{hyperref}`) to create PDF bookmarks. Set the correct page margins in *geometry*.

We are glad to pass on some hints about mathematical typesetting provided by one of our authors: The following packages were loaded for typesetting of mathematical formulae: `\usepackage{amsmath}`, `\usepackage{amssymb}`, `\usepackage{dsfont}`, `\usepackage{mathabx}`, `\usepackage{commath}`. The mathematical symbols were displayed correctly in the conventional PDF, but removed when converting to PDF/A. This problem was solved by removing the `{mathabx}` package.

5.1. pdfx

We recommend using the package pdfx via `\usepackage[a-1b]{pdfx}` to create a PDF which requires only little post-editing. You can find more information about the package here: <http://tug.ctan.org/tex-archive/macros/latex/contrib/pdfx>.

When using pdfx you need an ICC profile file which you can download from <http://www.color.org/srgbprofiles.xalter> (e. g. sRGB_IEC61966-2-1_black_scaled.icc). This file has to be saved in the same directory as the main LaTeX file and renamed to sRGBIEC1966-2.1.icm.

5.2. pdfTeX

Another option for creating PDF/A from LaTeX sources is pdfTeX, which is included in current distributions of MikTeX and TeXLive. You can find more information about the package here:

- http://support.river-valley.com/wiki/index.php?title=Generating_PDF/A_compliant_PDFs_from_pdfTeX

This information regarding LaTeX and PDF/A is based on the guide on creating PDF/A at the edoc server of the Humboldt University Berlin (Author: Dr. S. Henneberger): https://edoc-info.hu-berlin.de/de/zentral/diss_anleitungen_pdf. Further explanations were kindly provided by Roman Alexander (TU Berlin, Institut für Mathematik).

6. SOFTWARE-INDEPENDENT PDF/A CREATION

A common method to store PDF/A is using software which installs a virtual PDF printer on the system, so that regardless of the software with which the document was created, it can be “printed” to PDF. A selection:

Adobe Acrobat Pro (not Adobe *Reader!*) by <http://www.adobe.com>

PDF24 by <http://en.pdf24.org>

FreePDF + GhostScript by <http://freepdfxp.de/>

The usage of most software is largely the same. For example, to use PDF24 after installing it, **PDF24** has to be selected as the printer in your text processing or layout software. This opens a window in which, you can select “**PDF/A**” on the “**Save as ...**” dialogue.

7. VALIDATING THE PDF/A

7.1. Using veraPDF

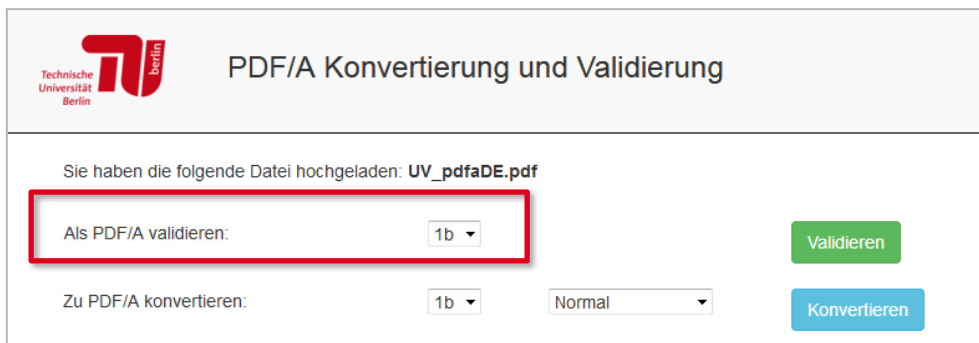
veraPDF is an open source PDF/A validator. veraPDF is currently at the development stage. The installer is available for Windows, Mac and Linux.

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

7.2. Using pdfaPilot by Callas (access for TU members)

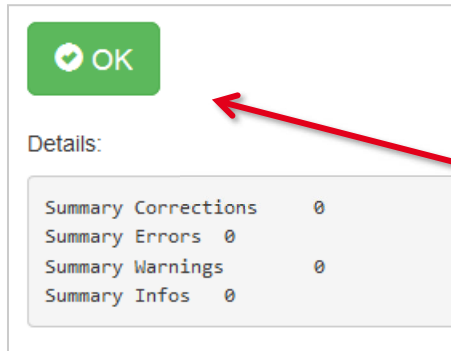
Members of the TU Berlin can use the [pdfaPilot](#) by Callas to check documents on the fulfillment of the PDF/A criteria and/or converting them to PDF/A. Please log in on <https://conversion.ub.tu-berlin.de/> with your TU account and password.

Once logged in upload the file and select -> “Validate as PDF/A” confirm with-> “Validate”. Choose PDF/A-**1b** or PDF/A-**2b** for validate.



Once you have finished editing, remove the file from the server.

Validation OK

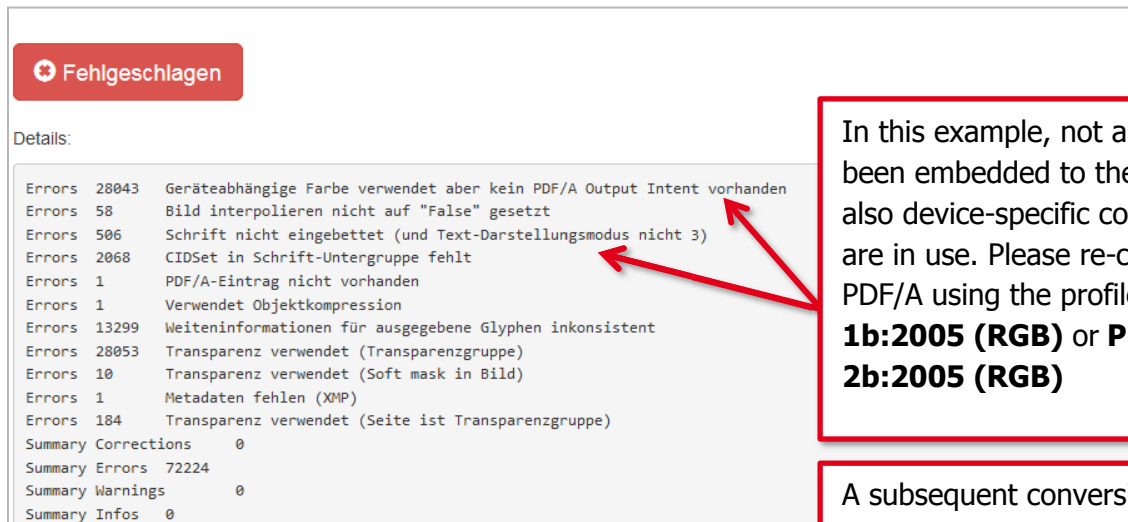


A green button with a checkmark and the text "OK" is displayed. Below it, the word "Details:" is followed by a table showing zero errors, warnings, and infos.

Summary Corrections	0
Summary Errors	0
Summary Warnings	0
Summary Infos	0

The checked file is PDF/A compliant. Please also check validity by copying and pasting text from the PDF/A into a Word or a text editor.

Validation not OK



A red button with a star and the text "Fehlgeschlagen" is displayed. Below it, the word "Details:" is followed by a list of errors and a summary table.

Errors	28043	Geräteabhängige Farbe verwendet aber kein PDF/A Output Intent vorhanden
Errors	58	Bild interpolieren nicht auf "False" gesetzt
Errors	506	Schrift nicht eingebettet (und Text-Darstellungsmodus nicht 3)
Errors	2068	CIDSet in Schrift-Untergruppe fehlt
Errors	1	PDF/A-Eintrag nicht vorhanden
Errors	1	Verwendet Objektkompression
Errors	13299	Weiteninformationen für ausgegebene Glyphen inkonsistent
Errors	28053	Transparenz verwendet (Transparenzgruppe)
Errors	10	Transparenz verwendet (Soft mask in Bild)
Errors	1	Metadaten fehlen (XMP)
Errors	184	Transparenz verwendet (Seite ist Transparenzgruppe)
Summary Corrections	0	
Summary Errors	72224	
Summary Warnings	0	
Summary Infos	0	

In this example, not all fonts have been embedded to the PDF and also device-specific color proof sets are in use. Please re-convert to PDF/A using the profile **PDF/A-1b:2005 (RGB)** or **PDF/A-2b:2005 (RGB)**

A subsequent conversion to PDF/A is also required if document text cannot be copied and pasted into other documents.

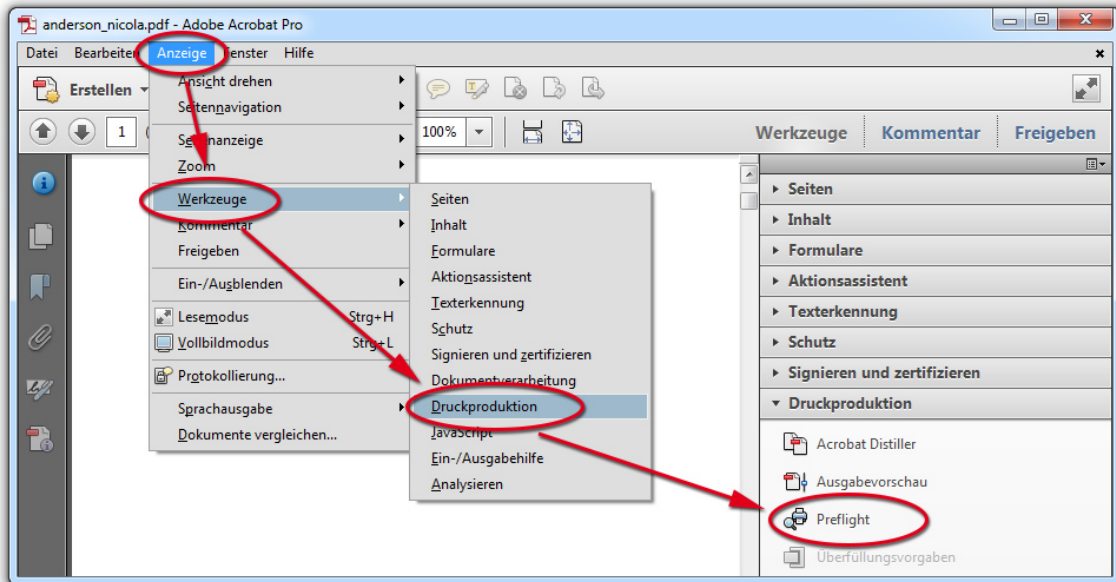
A detailed how-to-guide is available at:

<https://conversion.ub.tu-berlin.de/pdfa/help.php?lang=en>

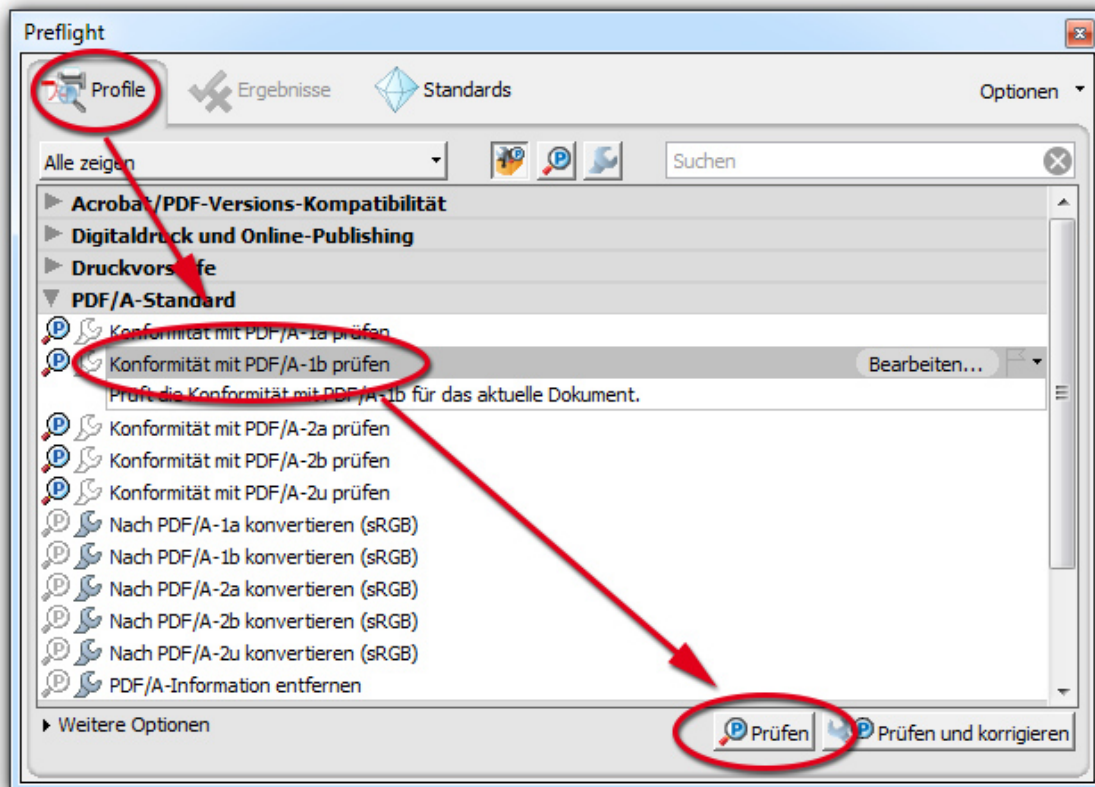
7.3. Using Adobe Acrobat Professional

Using Adobe Acrobat, documents can be checked on the fulfillment of the PDF/A criteria. The following section describes the steps for Acrobat X Professional.

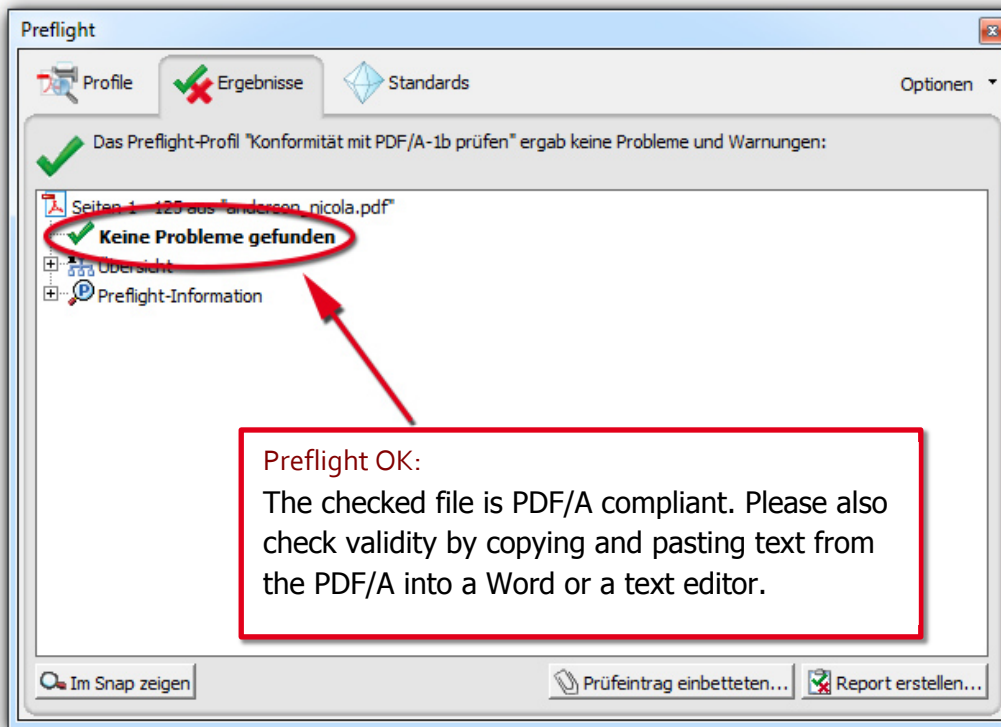
To do this, open the PDF file in Acrobat and choose the right point tools -> Print Production -> Preflight



This opens the "Profile" dialogue -> select **"Check conformity with PDF/A-1b"** or **"Check conformity with PDF/A-2b"** here.

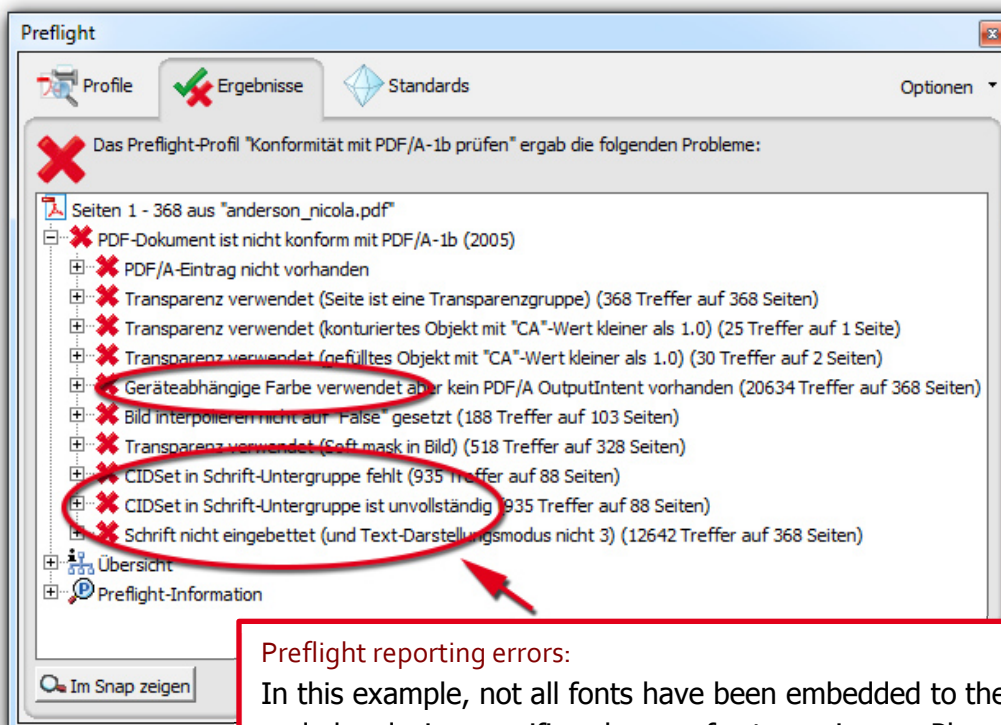


No errors during Preflight



Preflight OK:
 The checked file is PDF/A compliant. Please also check validity by copying and pasting text from the PDF/A into a Word or a text editor.

Errors during Preflight



Preflight reporting errors:
 In this example, not all fonts have been embedded to the PDF and also device-specific color proof sets are in use. Please re-convert to PDF/A using the profile **PDF/A-1b:2005 (RGB)** or **PDF/A-2b:2005 (RGB)**

A subsequent conversion to PDF/A is also required if document text cannot be copied and pasted into other documents.

7.4. Using 3-Height™ PDF Validator

You could also use [3-Heights™ PDF Validator](#) to check your PDF. The 3-Heights™ PDF Validator is a free online service to validate PDF files and check whether PDF/A criteria are met.

The final validation of the PDF file is conducted by the staff of the Dissertation Office.

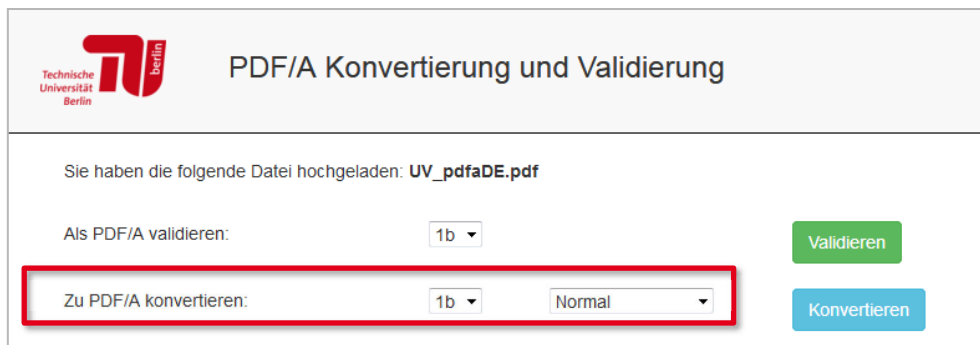
8. CONVERTING A FILE TO PDF/A SUBSEQUENTLY

Note: a new conversion of the original file with the profile PDF/A-1b:2005 (RGB) or PDF/A-2b:2005 (RGB) is always preferable. Subsequent conversion of the PDF file as described below may possibly lead to loss of information.

8.1. Using pdfaPilot by Callas (access for TU members)

Members of the TU Berlin can use the [pdfaPilot](#) by Callas to check documents on the fulfillment of the PDF/A criteria and/or converting them to PDF/A. Please log in on <https://conversion.ub.tu-berlin.de/> with your TU account and password.

Once logged in upload the file and select in the row -> "Convert to PDF/A" PDF/A-**1b** or PDF/A-**2b**. Keep the default setting ("Normal") for now. Confirm with -> "Convert".



The screenshot shows the 'PDF/A Konvertierung und Validierung' interface. At the top left is the TU Berlin logo. The main heading is 'PDF/A Konvertierung und Validierung'. Below this, it states 'Sie haben die folgende Datei hochgeladen: UV_pdfaDE.pdf'. There are two rows of options. The first row is 'Als PDF/A validieren:' with a dropdown menu set to '1b' and a green 'Validieren' button. The second row is 'Zu PDF/A konvertieren:' with a dropdown menu set to '1b' and another dropdown menu set to 'Normal', both of which are highlighted with a red rectangular box. To the right of these options is a blue 'Konvertieren' button.

Which of the possible options can lead to success depends on the source file and the occurring errors. This makes it difficult to provide a general recommendation. Are "transparencies" found faulty during validation, good experience were made with conversion to PDF/A-2b.

In addition to the conversion you could add metadata in the corresponding fields. Please note that existing data will be overwritten.

Zu PDF/A konvertieren: 1b Normal Konvertieren

Optional - ergänzende Metadaten bei Konvertierung (Überschreibt evtl. vorhandene Metadaten):

Autor(en) mit ';' getrennt: Musterfrau, Maxi

Titel: Test-Dokument

Keywords mit ';' getrennt: Test; PDF/A; Validierung; Konvertierung

Kurzbeschreibung: Dissertation, DDC 570: Biowissenschaften, Biologie

Conversion successful

OK

Sie können nun die Datei herunterladen. UV_pdfaDE_processed.p

The file was converted into PDF/A. A subsequent conversion to PDF/A is required if document text cannot be copied and pasted into other documents.

Conversion unsuccessful

Wenn Sie die Datei fertig bearbeitet haben, löschen Sie die Datei bitte vom Server. Löschen

Fehlgeschlagen

Details:

FixFailure	Nach PDF/A-1b konvertieren
Errors	1 Syntax-Problem: Reeller Wert
Summary Corrections	3750
Summary Errors	1
Summary Warnings	0
Summary Infos	0

The file could not be converted to PDF/A properly. In "Details" error comments can be found. Please attempt to solve the problem in the source document. You could also try a conversion with other parameters again.

If necessary test to convert the file with other parameters again:

- 1) In the first box select another PDF/A standard (e. g. 2b instead of 1b).
- 2) In the second box select another option. In further trials use the following order:
 - a. "Normal"
 - b. "Reconversion"
 - c. "Pages to images"
 - d. "Docs to images"

A detailed how-to-guide is available at:

<https://conversion.ub.tu-berlin.de/pdfa/help.php?lang=en>

8.2. Using Adobe Acrobat

