



1, 2 oder 3? a, b oder u? – Der PDF/A-Standard und seine verschiedenen Versionen

Fabienne Metzger M.A.

Digitale Geisteswissenschaftlerin und Sachbearbeiterin für die
digitale Langzeitarchivierung

- 2005: PDF/A-1 ISO 19005-1
- 2008: PDF 1.7 ISO-Standard
- 2011: PDF/A-2 ISO 19005-2
- 2012: PDF/A-3 ISO 19005-3
- 2020: PDF/A-4 ISO 19005-4

Vergleich der PDF/A-Versionen

Eigenschaften	PDF/A-1	PDF/A-2	PDF/A-3
basiert auf PDF 1.4	✓		
basiert auf PDF 1.7		✓	✓
alle Ressourcen (Bilder, Grafiken, Schriftzeichen) müssen in der Datei eingebettet sein	✓	✓	✓
fester Standard für Farbprofile (ICC-Farbprofile)	✓	✓	✓
Angaben wie zum Beispiel Autor, Dokumententitel, Erstelldatum und Erzeugerprogramm müssen XMP-Framework entsprechen	✓	✓	✓
transparente Elemente sind verboten	✓		
LZW- und JPEG2000-Komprimierung sind nicht zulässig	✓	JPEG 2000 erlaubt	JPEG 2000 erlaubt
Kompressionsverfahren wie JBIG2 und JPEG sind erlaubt	✓	✓	✓
Java-Script ist verboten	✓	✓	✓
Sicherheitseinstellungen sind verboten (Passwort)	✓	✓	✓
Einbindung von Daten beliebiger Dateiformate ist erlaubt			✓
Audio- und Videodateien dürfen nicht eingebettet sein	✓		
Einbettung von digitalen Signaturen unterstützt	✓	✓	✓
Verwendung von Verweisen (Hyperlinks) möglich	✓	✓	✓
Aktionen sind nicht zulässig	✓	✓	✓
Einbindung von PDF/A-1 und 2 konformen Dateien ist erlaubt		✓	✓
Einbettung von Open Type-Fonts ist erlaubt		✓	✓
keine Gewährleistung, dass die Anhänge zu einem späteren Zeitpunkt verarbeitet werden können			✓

Die Auswahl der Konformitätsstufe ergibt sich aus dem Ursprungsdokument und dem zukünftigen Verwendungszweck.

a = accessible

eindeutige visuelle Reproduzierbarkeit

gewährleistet Extrahierbarkeit von Textinhalten, Struktur muss die Leseabfolge klar darlegen

Barrierefreiheit

Bevorzugt bei der Konvertierung aus genuin digital entstandenen Dokumenten

b = basic

eindeutige visuelle Reproduzierbarkeit

gewährleistet Extrahierbarkeit von Textinhalten optional

optional Barrierefreiheit

OCR mit Einbettung in die PDF/A-Datei möglich

Besonders für die Konvertierung von digitalisierten Unterlagen geeignet

u = unicode

eindeutige visuelle Reproduzierbarkeit

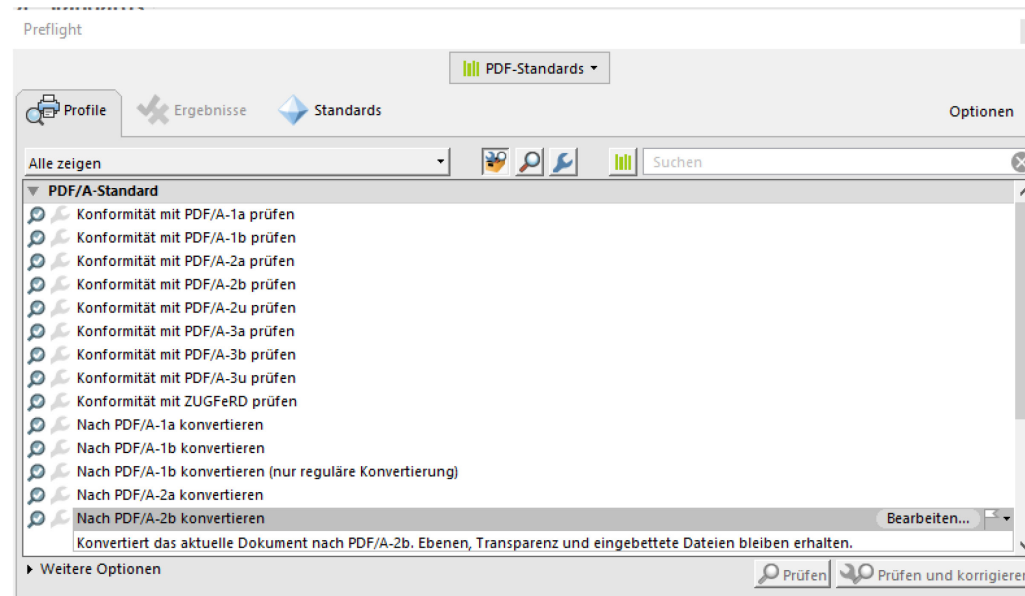
Textinhalte = Unicode, daher durchsuchbar und extrahierbar

Optionale Barrierefreiheit

Für digitalisierte Dokumente mit OCR-Erkennung sinnvoll

PDF/A-Dateien sollten vor einem Ingest ins Digitale Magazin validiert werden.
Nicht jede Datei, die die Konformität vorgibt, ist schlussendlich eine valide Datei.
Eine Validierung trägt zur Qualitätssicherung bei.

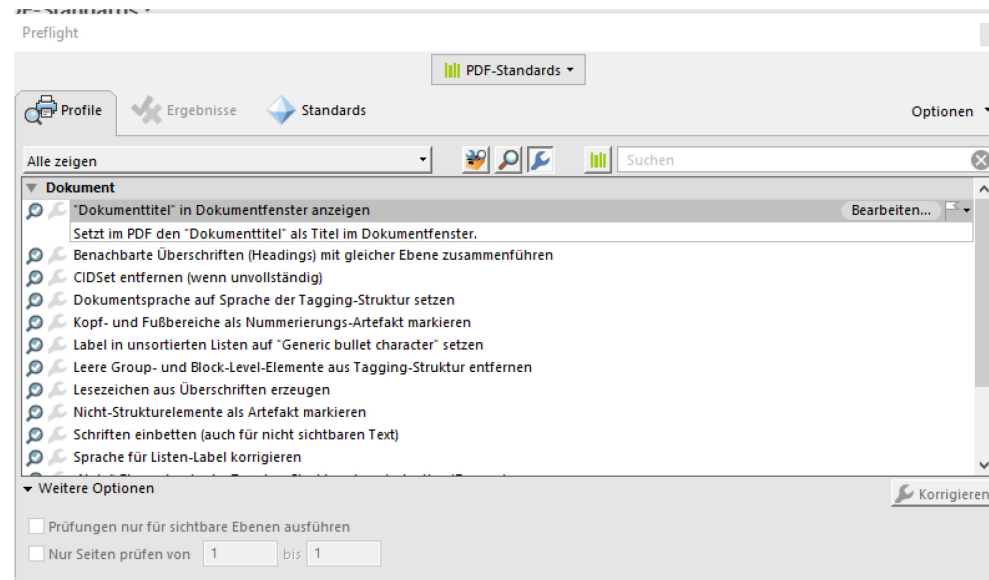
Beispiel Adobe Pro:



Alternativen: veraPDF, KostVal

Liegen Auffälligkeiten vor -> Validierungsbericht

Lösung: Reparatur mittels Software



- Spart Speicherplatz
- In weiten Kreisen akzeptiert
- Zahlreiche Werkzeuge bereits erhältlich
- PDF/A-Dokumenten bleiben fristlos gültig
- Alle wichtigen Informationen sind im Dokument
- Unterstützen Volltextsuche
- uvm.

- PDF/A-Version 2 vor Version 1, da Möglichkeit Ebenen und Transparenz - > geringere Wahrscheinlichkeit für Informationsverluste
- PDF/A-Version 3 = nicht archivfähig, da beliebige Dateiformate eingebunden werden können -> keine uneingeschränkte dauerhafte Lesbarkeit, ggf. aufwendige Migration, Emulation, Abhängigkeit von Software
- Stufe „a“ bevorzugen, da Barrierefreiheit. Aber: Eher nur bei genuin digitalen Dokumenten.
- Stufe „u“ vor Stufe „b“, da Textinhalte in Unicode -> durchsuchbar und extrahierbar. Für digitalisierte Dokumente mit OCR-Texterkennung sinnvoll. Barrierefreiheit optional.
- Stufe „b“ = eindeutige visuelle Reproduzierbarkeit, ohne Durchsuchbarkeit. „Mindestanforderung“
- PDF/A-1-Dateien müssen nicht in die neue Version konvertiert werden.

Immer an die Gegebenheiten und Möglichkeiten im jeweiligen Archiv anpassen!

- [bundesarchiv.de](https://www.bundesarchiv.de)
- [PDF/A kompakt \(pdfa.org\)](https://www.pdfa.org)
- [PDF-Dokumente | Forschungsdatenzentrum \(ianus-fdz.de\)](https://www.ianus-fdz.de)
- [kost-ceco | PDF/A-1](#)
- [kost-ceco | PDF/A-2](#)
- [kost-ceco | PDF/A-3](#)

Vielen Dank für Ihre Aufmerksamkeit!



Fabienne Metzger M.A.

Metzger.Fabienne@LRA-ES.de

0711/3902-43130