

Neuer Standard für Archivierung: PDF/A?

War in der Vergangenheit Tiff in vielen Firmen und Behörden in aller Welt das bevorzugte Standardformat für Archivierungszwecke, so dreht sich in den letzten Jahren vieles um PDF/A

PDF kann Texte, Vektorgraphiken, Rasterbilder u.a. speichern. Der Text kann durchsucht werden. Also einige Vorteile gegenüber dem Raster-/Pixelformat Tiff.

Um eine Volltextsuche durchzuführen, müssen erst OCR-Programme zum Einsatz gebracht werden

PDF Dateien können kleiner in der Dateigröße sein. Eine kleinere Dateigröße bietet Vorteile bei der elektronischen Übertragung.

Einträge wie Titel, Autor, Erstellungs- und Modifikationsdatum, Inhalt, Schlagworte usw. (sogenannte Metadaten) lassen sich in PDF-Dokumente einbetten. Damit werden sie leichter und kostengünstiger indexierbar.

PDF/A

PDF/A (festgelegt in der ISO 19005-1) basiert auf PDF 1.4.

In der ISO-Norm werden formale Anforderungen an elektronische Dokumenten beschrieben:

So müssen z. B. sämtliche notwendigen Informationen im Dokument selbst enthalten sein: Texte, Vektorgraphiken, Rasterbilder, Schriftarten, Farbräume usw..

Es dürfen keine Referenzen auf externe Quellen, die von der Datei geladen werden müssen.

Es gibt also sowohl Erweiterungen als auch Einschränkungen gegenüber der PDF 1.4. Referenz

So ist z.B. Transparenz sowie Ton- und Videoreproduktion nicht erlaubt, im Gegensatz zur PDF 1.4. Referenz.

Alle verwendeten Schriften müssen eingebettet sein.

Die **Internationale Organisation für Standardisierungen (ISO)** hat 2005 dem neuen Standard für die Regelung der Archivierung elektronischer Dokumente zugestimmt:

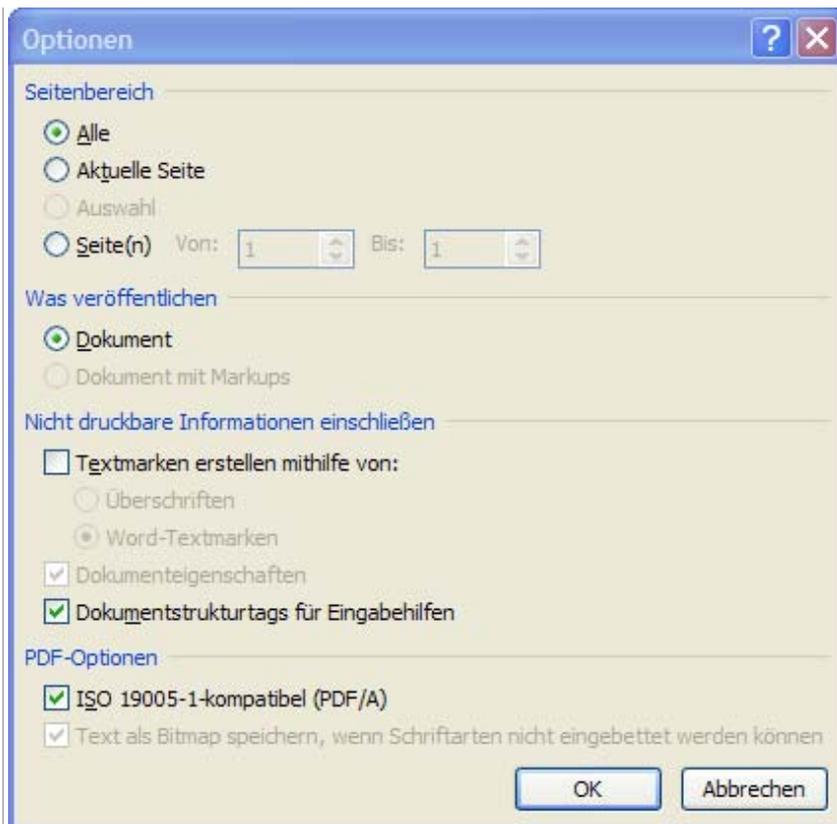
ISO-19005-1 - Document Management – Electronic document file format for long-term preservation – Part 1: Use of PDF 1.4 (PDF/A-1).

Erweiterungen dieser Standards werden folgen in den nächsten Jahren

Von PDF/A ist bisher nur der Teil **PDF/A-1 (Teil 1)** gültig.

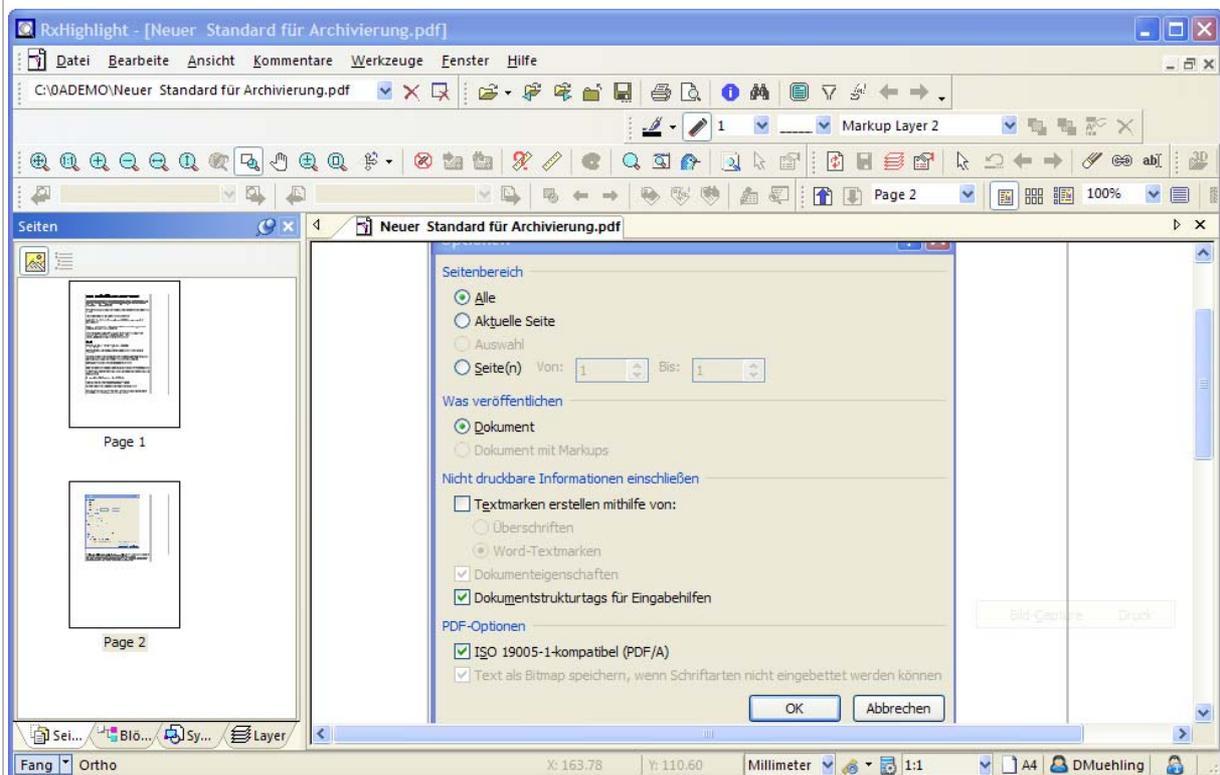
PDF/A in der Praxis

PDF/A-konforme Dokumente können z.Zt. mit der Version 8 von Adobe Acrobat oder Microsoft Word 2007 generiert werden. (Siehe Bild unten). Andere Hersteller werden folgen oder haben schon entsprechende Schnittstellen in neue Produkte eingebaut.



PDF/A und die Viewer von Rasterex: RxView|RxHighlight

Für die Viewer von Rasterex ist es kein Problem, eine so mit Winword 2007 erzeugte Datei darzustellen. Woran gearbeitet wird, sind die Feinheiten: Auslesen von Metadaten usw.



Dieser Text als PDF/A dargestellt in RxHighlight R9