

PDF/A-3 Overview

Introduction

Since its release in the fall of 2005, the **ISO Standard PDF/A** has prevailed as the accepted standard for archiving PDF documents in the long term. The first part of the standard was based on PDF Version 1.4 and contained a range of rules such as safeguarding the unambiguous visual reproduction of PDF documents, independence from display and printing technology, and the availability of PDF documents to people with limited visual acuity.

It was followed in July 2011 by the ISO standard PDF/A-2 (ISO 19005-2), which was extended by various functions that were made possible by later PDF versions (up to and including PDF Version 1.7). This extended PDF standard is no longer based on a PDF version by Adobe, but on the ISO standard 32000-1 (PDF 1.7).

The ISO Committee released the third edition of the standard (ISO 19005-3) in October 2012; its official title is:

“Document management – Electronic document file format for long-term preservation – Part 3: Use of ISO 32000-1 with support for embedded files (PDF/A-3)”

It is based on ISO 32000-1 (PDF 1.7), as was PDF/A-2. Analog to PDF/A-2, the PDF/A-3 standard also defines three conformity grades:

- **PDF/A-3a**
is concerned with accessibility.
- **PDF/A-3b**
is concerned with visual integrity, i.e. the consistently identical display of the document.
- **PDF/A-3u**
facilitates the searchability of texts and the copying of Unicode text for digitally created PDF documents and those scanned using optical character recognition (OCR).

What is the difference between PDF/A-2 and PDF/A-3?

PDF/A-3 contains just one change that is necessary but controversial: PDF/A-2 already enabled the embedding of PDF/A-compliant documents as attachments. PDF/A-3, however, makes it possible to embed any document format such as Excel, Word, HTML, CAD or XML files for the first time ever.

The purists among the experts opine that this amendment is contradictory to the original idea behind the PDF/A standard. However, pragmatists in companies from various segments, such as the pharmaceutical industry or the banking and financial sector, have a concrete need to keep the original file format alongside the converted PDF/A file. Files that belong together are compiled to form a “collection”. This construct has been known since the days of PDF/A-2. Typical applications include archiving emails and their attachments that can consist of many different file formats.

The standard only assures the representation of PDF/A documents viewed via a compliant viewer. The presentation of non-compliant embedded documents is implemented via a separate action using the tools that support the document formats in question.

Which PDF/A format is the right one?

PDF/A-3 should only be used if you plan to embed documents that do not comply with the PDF/A standard. PDF/A-2 is the right choice in all other cases as it makes it quite clear that no other formats are embedded. PDF/A-1 is still good enough for anyone who does not need all the functionality offered by PDF/A-2. There is no need to migrate existing archives as a PDF/A-3 compliant viewer can display all PDF/A compliant files.

Conclusion

PDF/A-3 meets an important user requirement, namely an option to embed file formats that do not comply with the PDF/A standard. Because this amendment is desirable but controversial, it remains the only change to the PDF/A-2 standard. The user can therefore choose between a pure PDF/A collection and a mix of the various standards that is easily differentiated thanks to the “PDF/A-3” label.

Documentation for further perusal

- White Paper: “PDF/A – The standard for long-term archiving”
www.pdf-tools.com/pdf/pdf-whitepapers-pdf-a-english.aspx
- PDF/A-2 Overview
www.pdf-tools.com/pdf/pdf-whitepapers-pdf-a-english.aspx