PDF/A: THE KEY TO LONG-TERM ARCHIVING

PDF/A is a proven method of archiving documents in a format which preserves accessibility and readability- regardless of which program or system originally generated them.

PDF/A is an ISO standard for long-term archiving of electronic documents. It does not define an archiving strategy, but rather an electronic document format which ensures that documents will be reproducible for years into the future. The PDF/A format preserves the visual structure of electronic documents, regardless of which tool or system was originally used to generate or store them.

PDF/A is based on Adobe Systems Inc.'s Portable Document Format (PDF) technology, but leaves out functions which are not appropriate for long-term archiving. These include audio and video elements as well as transparent objects. An essential element of PDF/A reproducibility is that the documents are 100% self-contained. This function ensures that information required for consistent document presentation is embedded in the file. All content, including text, raster images and vector graphics as well as fonts and colour information, is therefore included in every PDF/A file.

Why archive documents with PDF/A

"Adopting LuraDocument allows us to archive cover images in colour. The cover image archive has become an important element of our day-to-day work."

– Knut Külsen,
Publishing Representative at Bauer

Long-term access
PDF/A files are designed to be accessible independently of platform, technology or manufacturer. Documents are therefore guaranteed to be viewable in the future, without proprietary technology.

Standard
PDF/A is based on the recognised PDF standard and is itself an ISO standard. The long-term nature and easy access to content saved in this format is therefore guaranteed.

Information optimization
PDF/A allows metadata saving and automatic data classification based on this information, as well as the ability to carry out full text searching across the entire archive.

Compression
PDF/A supports compression of data to a fraction of the original size without influencing document quality. The smaller files reduce storage requirements and facilitate transfer over FTP, as email attachments etc.

LuraTech and PDF/A
LuraTech is among the pioneers of PDF/A standards development and the development of products using PDF/A. LuraTech was an active participant in the ISO committee responsible for establishing the standard. Furthermore, LuraTech is among the founding members of the PDF/A Competence Center, an association which provides support for all questions regarding PDF/A and which is dedicated to supporting the format. As one of the first providers, LuraTech presented PDF/A products at CeBIT 2006. Today, LuraTech's product portfolio includes numerous standards-compliant solutions with which PDF/A files can be generated and validated. Of all products on the market today, these solutions provide the most balanced relationship between compression rate, image quality and performance.

LuraTech uses award-winning Mixed Raster Content (MRC) compression technology, in part to reduce document sizes a hundred fold. In the process, image quality and the readability of the text remain preserved. In order to provide a complete conversion solution, we have integrated additional leading technologies such as ABBYY OCR technology for full text searching.

LuraTech

For more information
www.luratech.com
info@luratech.com

LuraDocument ®PDF/A-Solutions

LuraDocument PDF Validator
This tool tests and confirms that all PDF files conform to the PDF/A standard. It ensures that files are correctly structured and can be reliably reproduced in a manner faithful to the original version.

LuraDocument PDF Compressor
This product family is easy to integrate into workflows and uses our award-winning mixed raster content compression technology to convert colour, greyscale or black/white scanned documents into high-quality, highly-compressed PDF and PDF/A files. Our products additionally include the ABBYY OCR technology for full text searching in all PDF and PDF/A files.

System Requirements
Windows® Server 2003 / Server 2008* / XP / Vista / Windows 7 / Windows 8
CPU Speed: 2GHz or greater (recommended)
RAM: 1GB minimum, 2 GB per processor core (recommended)
250MB free hard drive space

** Born Digital Module requires Server 2003 or Server 2008*

Compliance
ISO 19005-1 compatible PDF/A-1a and -1b
ISO 19005-2 compatible PDF/A-2u and -2b
ISO 19005-3 compatible PDF/A-3u and -3b
Reader 5-11 compatible PDF 1.4-1.7

Technical description and areas of application for PDF/A

PDF/A is based on the PDF 1.4 specification published by Adobe in 2001. The PDF/A standard - formally known as ISO 19005-1:2005 or PDF/A-1 – was recognised by the ISO in September 2005. ISO 19005-1 defines “a PDF-based file format named PDF/A which provides a means to display electronic documents in a manner which preserves visual appearance over time, regardless of which tools and systems were used for production and saving”.
PDF/A is guided by four basic principles which are briefly illustrated with practical examples:

- Device-/Software-/Version-independent: Content is always displayed the same way.
  PDF/A can be displayed on e.g. Windows, Macintosh or UNIX systems.
- Self-contained: A PDF/A-compatible file contains all components necessary for displaying it. The simplest example of this is the demand that fonts must be embedded, as it cannot be foreseen whether the fonts will be available on future computers.
- Self-documented: A PDF/A-compatible file describes and documents itself (metadata). Necessity aside, metadata is an important means of describing documents.
- Transparent: A PDF/A-compatible file can be analysed by simple means. Even in a “worst-case” scenario in an unforeseeable future, all that is needed is the ISO standard and some form of computer in order to compile a viewer to reproduce PDF/A documents.

PDF/A is designed as a multi-element standard range.

PDF/A is further subdivided into the following conformance levels:

- PDF/A-1a (Level A conformance)
  Ensures conformity of the logical document structure and semantics, including the natural reading system. All Level B requirements must also be fulfilled in Level A.
- PDF/A-1b (Level B conformance)
  Ensures reproduction of documents without visual ambiguities, as well as colour and text consistency.
- PDF/A-2a (Level A conformance)
  Satisfies level 1a in accordance with the PDF/A-2 standard. Level A is also the basis for complete accessibility, which will be further detailed in the PDF/UA standard.
- PDF/A-2u (Level U conformance)
  Was recently introduced with the PDF/A-2 standard as an intermediate level. As well as the Level B requirements, Unicode is used for unambiguous character encoding. Documents in the scanning and output sector benefit in particular from this.
- PDF/A-2b (Level B conformance)
  Ensures reproduction of documents without visual ambiguities, as well as colour and text consistency.
- PDF/A-3
  PDF/A-3 allows the embedding of any source data.

Forwards compatibility is guaranteed in line with PDF/A standards. A PDF/A-1 document is also further valid with the PDF/A-2 and also with PDF/A-3 standard. There is therefore no need for migration if PDF/A-1 documents have already been generated and archived.

Whether PDF/A-1, PDF/A-2 or PDF/A-3 should be adopted depends on individual project requirements.

Areas of application for PDF/A
Fundamentally, PDF/A can be used to archive all electronic documents that can also be printed out, as well as all non-electronic documents that can be scanned in. Adopting PDF/A therefore makes sense in many sectors. The following are exemplary outlines of some selected areas of application in which PDF/A is already used for archiving.

- Incoming mail
  However incoming mail reaches an enterprise- by classical post, fax or email. Incoming paper documents must be scanned into an electronic archive, while emails- including any attachments- must be converted to an appropriate archival format. A unified infrastructure is best realised using PDF/A as a foundation..
- Document digitization
  In many organisations, document portfolios (such as credit documents in banks or construction documents in public agencies) must be digitized. PDF/A lends itself to this as a secure standard for long-term archiving with many advantages.
- Archive migration
  For various reasons, some archive users are no longer satisfied with their existing archive infrastructure and must replace their hardware or software. If all documents must be processed in the course of an archive migration, it can be a sensible opportunity to convert old TIFF archives, for example, to the modern PDF/A format at the same time.

LuraTech

For more information
www.luratech.com
info@luratech.com

Wiltshire, UK
+44 (0)1793 / 85 21 00
+44 (0)1793 / 23 82 52 (F)

California, USA
+1 (650) 481 - 8156
+1 (650) 779 - 4644 (F)

Berlin, Germany
+49 (0)30 / 39 40 50 - 0
+49 (0)30 / 39 40 50 - 99 (F)

Remscheid, Germany
+49 (0)2191 / 58 96 0 - 0
+49 (0)2191 / 58 96 0 - 29 (F)