

7 Minutes with PDF/A

Leonard Rosenthol
PDF Architect · Principal Scientist
Adobe Systems



PDF/A-1 – PDF for Archiving

- **ISO PDF/A 19005-1 (based on PDF 1.4) - Published October 2005**
- Standard defines what is required, recommended & prohibited
- For more information visit: http://www.aiim.org/pdf_a

- ✓ An archival replacement for paper
- ✓ Customized to offer consistent, long-term retrieval and rendering
- ✓ Includes document metadata (XMP) for search and retrieval
- ✓ Engineered for reliable migration and future compatibility
- ✓ Digital signatures are for verification and not authentication
- ✓ Developed and maintained by PDF/A ISO committee

PDF/A Details:

- ☒ All data must be self-contained
- ☒ All color is managed
- ☒ No encryption
- ☒ No multimedia
- ☒ No JavaScript
- ☒ No Transparency or Layers
- ☒ No JPEG2000
- ☒ No Attachments
- ☒ Basic Digital Signatures

PDF/A-2 – PDF for Archiving

- **ISO PDF/A 19005-2** (based on **ISO 32000-1**) - **Published June 2011**
- Standard defines what is required, recommended & prohibited
- For more information visit: http://www.aiim.org/pdf_a

- ✓ An archival replacement for paper
- ✓ Customized to offer consistent, long-term retrieval and rendering
- ✓ Includes document metadata (XMP) for search and retrieval
- ✓ Engineered for reliable migration and future compatibility
- ✓ Digital signatures are for verification and authentication
- ✓ Developed and maintained by PDF/A ISO committee

PDF/A Details:

- ✗ All data must be self-contained
- ✗ All color is managed
- ✗ No encryption
- ✗ No multimedia
- ✗ No JavaScript
- ✗ Supports Transparency & Layers
- ✗ Supports JPEG2000
- ✗ Only PDF/A Attachments
- ✗ Advanced Digital Signatures

PDF/A-3 – PDF for Archiving

- **ISO PDF/A 19005-3** (based on ISO 32000-1) – **TO BE RATIFIED IN MAY 2012**
- Standard defines what is required, recommended & prohibited
- For more information visit: http://www.aiim.org/pdf_a

- ✓ An archival replacement for paper
- ✓ Customized to offer consistent, long-term retrieval and rendering
- ✓ Includes document metadata (XMP) for search and retrieval
- ✓ Engineered for reliable migration and future compatibility
- ✓ Digital signatures are for verification and not authentication
- ✓ Developed and maintained by PDF/A ISO committee

PDF/A Details:

- ☒ All data must be self-contained
- ☒ All color is managed
- ☒ No encryption
- ☒ No multimedia
- ☒ No JavaScript
- ☒ Supports Transparency & Layers
- ☒ Supports JPEG2000
- ☒ **Any type of Attachment**
- ☒ Advanced Digital Signatures

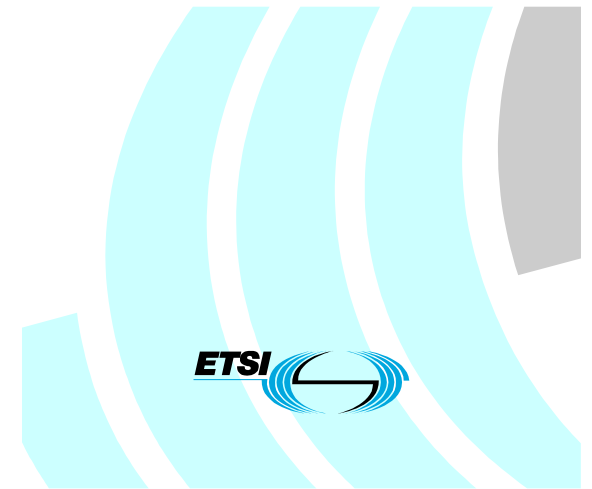
Noteworthy New PDF Features in PDF/A-2 & PDF/A-3

- Improved compression technology/Smaller files
 - JPEG2000
 - Compressed XRefs & Streams (aka "Full Compression")
- Transparency
- PDF Layers (aka Optional Content)
 - Many views may be stored, but only one at a time can be seen
 - And whatever you view on screen, must print!
- PDF Packages/Collections
 - May only contain other PDF/A documents
- Improved tagging/accessibility
 - Including rich support for table-like structures
- Digital signature enhancements
 - Certified Documents (aka DocMDP Signatures)
 - Improved revocation checking
 - Use of ETSI TS 102778 (aka PAdES) as normative requirement



PAdES (PDF Advanced Electronic Signatures) Compliance

- Normative Annex B – “Requirements for Digital Signatures in PDF/A”
- Incorporates all of the requirements and recommendations from PAdES Part 2
 - Signing
 - Verification
- **Every digitally signed PDF/A-2 file WILL comply with PAdES!**



Level U Conformance

- The Level U conformance requirements are intended to be those necessary to ensure that
 - not only is the rendered visual appearance of a conforming file preservable over the long term,
 - but that any text contained in the document can be reliably extracted as a series of Unicode codepoints.
- However, Level U conforming files might not have sufficiently rich internal information to allow for the preservation of the document's logical structure and content text stream in natural reading order, which is provided by Level A conformance.



- Attachments & Embedded Files – That's it!!
 - Introduces a new concept called “Associated Files”
 - Connects various types of content objects (Images, Pages, etc.) to File Specification dictionaries while identifying the reason that the file is associated with it
 - Source
 - Data
 - Alternative



■ ISO 32000-2



- Other committees are working on key issues:
 - PDF/E studying archivability of the new features of 32K-2 for PDF/E-2.
 - PDF/UA and 32K are jointly studying tagging needs for improved semantics.
- My belief is that the committee will take a break for a bit then come back and see how they can align with the work of those other committees to determine if another version of PDF/A is necessary/required/warrented.



Adobe