



Adobe[®] Acrobat[®]
PDF CONFERENCE
Orlando 2004

Presented by



How I do things

- You should all have received copies of the presentation before you walked in.
- Electronic versions are also available on my website at <http://www.pdfsages.com>
- I've left time at the end for Q&A, but please feel free to raise your hand at any time.



What is PDF/A?

- A file format for the “Long Term Electronic Preservation” of multi-page documents that may contain a mixture of text, raster images and vector graphics.



Long Term?

- Court documents protect citizens rights
 - Access is assured in trial courts for 20 to 40 years for the Judiciary
- Permanent records held “until the end of the republic” by NARA (National Archives & Records Administration)
 - NARA enables people to inspect for themselves the record of what government has done. NARA enables officials and agencies to review their actions and helps citizens hold them accountable for those actions. And NARA ensures continuing access to essential evidence that documents the rights of American citizens, the actions of Federal officials, and the national experience.



What PDF/A does NOT solve...

- Hardware migration issues
- OS and application migration issues
- Document management system issues
 - Document access tracking
 - Digital Rights Management



Some previous attempts...

- Federal courts began imaging for Central Violations Bureau in 1993 in TIFF
 - Short retention schedule on tickets and citations (5 years)
 - National system allowed for electronic shipment of calendars and tickets
 - Operational savings from reduced number of files



More History

- **Maritime Asbestos cases - 1995**
 - Format important and retention long-term
 - PDF chosen
 - Born electronic & paper
- **New York Southern Bankruptcy Pilot - 1996**
 - Multi-state filings
 - Sophisticated Court and Bar
 - Seven plus years of filings now in PDF
- **SEC/EDGAR allows submissions of PDF - 1999**



So what is needed?

- A document format that
 - Can convey critical information
 - Can be rendered accurately
 - “consistent and predictable”
 - Can incorporate associated “marginalia”
 - Notes scrawled in the margins on documents
 - Today perhaps as version history, general metadata, etc.
- A definition of how retrieval devices (Readers) will behave



What is PDF/A?

- International Standard
 - Initiated by AIIM & the US Gov't (Courts, NARA, IRS)
- Based on Portable Document Format (PDF)
- Long term preservation of black and white and color compound documents as electronic data
 - Combinations of character, raster, vector and other data
 - Preservation and retrieval of appropriate metadata
- Initial version “emulates static paper”
 - Plus annotations, signatures, marginalia and approvals



Details

- **Based on PDF 1.4 & PDF/X (ISO 15930)**
 - All PDF/X documents can potentially be minimally conforming PDF/A documents without any changes
 - Reduces ambiguity between different vendor's implementations
- **More restrictive “coding” of PDF details**
 - Ensures less ambiguity when implementing
 - Improved searching/indexing
- **Removal of any complex or potentially confusing graphic concepts**
 - No transparency
 - Limited colorspace



More Details

- No Security/Encryption
- All data must be self-contained
 - No external references
 - Fonts MUST be embedded(!!)
- Limited annotation support
 - No movies and sounds
 - No JavaScript
 - Links are stored but not to be executed
- Digital Signatures are for verification and not authentication
- Metadata based on Adobe XMP



Two “levels”

■ Minimal Conformance

- Meet the standard/basic requirements

■ Full Conformance

- Improved searchability via Unicode mappings
- Comprehensive metadata requirements
 - Font data
 - Document “pedigree”
 - Audit trail



Some PDF/A history

- NARA announcement March 31, 2003
 - PDF/A defining future technical standards
- Lots of meetings here in the US to define the requirements and then start on the technical points to achieve them
- First presented to ISO in October, 2003
- Revised by US committee
- Represented to ISO, March, 2004 during AIIIM
- Currently back in US committee for (hopefully) final revisions



Where are we going?

- 3rd party tools

- Verifiers
- Creators

- Acrobat Integration

- “Save as PDF/A”
- PDF/A job options for Distiller
- Changes to viewer to meet PDF/A requirements



What can you do now??

- Make good PDFs using high quality tools
 - ALWAYS embed ALL fonts
- Use preflight tools to check for PDF/X compatibility
 - If you meet PDF/X-3 compliance, you are also PDF/A complaint
- If creating PDF from scans, be sure to use OCR to enable searching/indexing



Wrap Up

- PDF/A is

- A file format for the “Long Term Electronic Preservation” of multi-page documents that may contain a mixture of text, raster images and vector graphics.

- It is in final review and should become an international (ISO) standard in the next 6-12 months.

- So now is the time to start thinking about it, esp. if you work with the government or regulated industry





Questions and Answers

Q & A

