PDF/A and Archiving

I've heard about PDF. What's PDF/A?

Digital, PDF documents are rapidly replacing paper ones, as many government and private industries will attest. But digital documents have become saddled with technical challenges that make it difficult for the files to remain readable, renderable, and accessible over the long term. The new PDF/A standard was created to solve this problem.

Why do I need PDF/A?

PDF files are unsuitable for long-term preservation of digital documents for two reasons. First, PDF documents are not self-contained; they draw upon external fonts and other content not stored in the original file. Second, the lack of standardization among the many PDF development tools on the market created inconsistencies in the implementation of the file format. For example, software incompatibilities can easily render old file formats obsolete.

PDF/A addresses the growing need to archive documents electronically in a way that will ensure both the preservation of their contents over an extended period of time and that those documents can be retrieved and rendered with consistent and predictable results in the future.

PDF/A, the standard

In 2005, the International Organization for Standardization (ISO) approved the PDF/A standard, which specifies requirements for archiving and preserving digital documents in PDF. This is the first in a new family of ISO standards to address the need to preserve the visual appearance of digital documents, independent of the tools and systems used for creating, storing, or tending the files.

PDF/A in action

The new standard's activity was initiated in the United States through the joint sponsorship of the Association for Information and Image Management (AIIM) and the Association for Suppliers of Printing, Publishing, and Converting Technologies (NPES). This new standard will have a broad-reaching effect on record-keeping around the world.

The U.S. National Archives and Records Administration (NARA) is one of the first organizations to state its intent to accept PDF/A. Susan Sullivan, an IT specialist from Electronic Records for NARA, states that agency will accept transfer of permanent records in PDF/A format, according to current transfer requirements for electronic records.*

Long-term preservation needs

There are many factors that make PDF/A an ideal solution for the long-term preservation of electronic documents. This new standard addresses many long-term preservation needs, including the following:

1. Device Independent - Can be reliably and consistently rendered without regard to the hardware or software platform

2. Self-contained - Contains all resources necessary for rendering

3. Self-documenting -Contains its own description

4. Unfettered - Absence of technical file protection mechanisms

5. Available - Authoritative specification publicly available

6. Adoption - Widespread use may be the best deterrent against preservation risk

What does PDF/A allow or disallow?

A key difference between PDF and PDF/A is the restrictions that PDF/A places on PDF.

PDF/A-1 files must include:
- Embedded fonts
- Device-independent color
- XMP metadata

PDF/A-1 files may not include:
- Encryption
- LZW compression
- Embedded files
- External content references
- PDF transparency
- Multimedia
- JavaScript

Vendors supporting PDF/A include:

Adobe Systems Inc. – www.adobe.com
Apago – www.apago.com
Callas - www.callassoftware.com
Compart Systemhaus GmbH – www.compart.net
PDF Tools AG – www.pdf-tools.com
Visioneer – www.visioneer.com

For more information:

PDF/A standard ISO 19005-1:2005, Document management — Electronic document file format for long-term preservation — Part 1: Use of PDF 1.4 (PDF/A-1)
http://www.digitalpreservation.gov/formats/fdd/fdd000125.shtml

"The Long-Term Preservation of Authentic Electronic Records: Findings of the InterPARES Project"
www.interpares.org/book/index.cfm

Digital Audio-Visual Preservation Prototyping Project
www.loc.gov/rr/mopic/avprot/avprhome.html

U.K. Records Management for Central Government
www.pro.gov.uk/recordsmanagement

Victorian Electronic Records Strategy Standards and Guides
http://www.prov.vic.gov.au/vers/standard/spec_04/3-2.htm

School of Information Management and Systems at the University of California, Berkeley. Berkeley study:
www.sims.berkeley.edu/research/projects/how-much-info/

* Betsy Fanning, "PDF/Archive An Update," AIIM E-DOC Magazine 30 Jun. 2005

For more information about Apago, call 770-619-1884 or visit www.apago.com.