

PDF-Archive
Draft Meeting Minutes
December 12-13, 2002

Mr. Steve Levenson called the meeting to order at 10:15 a.m. The meeting participants included:

Stephen Abrams	Harvard University Library
Lee Altman	Glaxo Smith Kline
Christine Bontempo	NPES
Robert Borochoff	Administrative Office, U.S. Courts
Robert Breslawski	Eastman Kodak Company
John Brinkema	Administrative Office, U.S. Courts
Jane Cohen	DTIC
Betsy Fanning	AIIM International
Mark Gavin	Appligent, Inc.
Macduff Hughes	Adobe Systems
John Janick	Merck & Company, Inc.
Dwight Kelly	Apago, Inc.
Bill LeFurgy	Library of Congress
Steve Levenson	Administrative Office, U.S. Courts
Rich Lysakowski	CENSA and GERA
Joe McConnell	Proquest Co.
David McDowell	NPES
John Mancini	AIIM International
Pamela Mason	NARA
Basil Manns	Library of Congress
Ed Medvid	SRA International Inc.
Leonard Rosenthol	PDF Sages, Inc.
Lou Sharpe	Picture Elements Inc.
Paul Showalter	Internal Revenue Service
Mary Beth Smartt	Global Graphics Software
Dana Stone	Merck & Company
Susan Sullivan	NARA
David Van Driessche	Enfocus Software
Melonie Warfel	Adobe Systems

The December 12-13, 2002 meeting agenda (PDF-A 2002-002) was approved with the addition of New Business Item including compression, new standards functions, and extensibility/archivability and a philosophical discussion on the structure for the standard after approving the minutes. (Smartt/Cohen)

The draft meeting minutes for the October 3-4, 2002 (PDF-A 2002-001) meeting were approved as amended with the removal of the table listing the working group listserv addresses that appears on the last page. (Mason/Levenson)

The secretariat will send draft version of the next revision of ISO 15930 (PDF-X) to the committee and will distribute the published version to the committee upon receiving approval from ISO.

AIIIM will initiate paperwork to formalize the ISO Joint Working Group comprising ISO TC 130, Graphic Technology; TC 46, SC11, Information and documentation, Archives/records management; TC42, Photography and TC 171, SC2, Document Imaging Applications, Application Issues to work on the project internationally. Upon approval of the Joint Working Group, paperwork will be submitted to ISO to formally register the project with ISO.

Each of the working groups provided status reports highlights follow.

The Rendering Working Group reviewed the scope and work of the working group. They identified the following issues:

- virtual colorspace,
- high level color fidelity is not essential,
- questioned why character spacing is included in rendering
- overlap with other working groups
- expressed interest in following tagged PDF guidelines for natural reading order
- hairlines and drawings have not been addressed

Regarding the Rendering section of the document:

- 4.2-4.4 exclude external data
- 4.5 device dependent
- 4.6 allowed default of TR2 key
- Transfer and half-tone device specific; can put in PDF or exclude – decided to treat as in PDF-X an put it in but ignore it
- 4.7 excluded the same in PDF-X
- Interpolation excluded from image dictionary (4.2)
- Thumbnails – one suggested to keep
- PDF-X excluded RGB
- Operators – an error for any content stream to include

Document must be rendered according to the specification. We may want to dictate there shall be no undefined keys.

The Rendering Working Group recommended that the Rendering, Fonts and Tagged groups potentially may be best merged.

The Fonts Working Group reported that semantic requirements add stress on the producer to do more than just reproduce the symbol that could potentially hurt the ability to bring

PDF-X into PDF-A. The group looked at fonts but did not discuss operators. We need to make sure PDF-A allows future systems to read the files and specify semantic level to give information about characters. We will use additional informative notes in the document to justify decisions.

- 6.2.1, Mapping fonts to glyphs, PDF-X used manually. We shall embed fonts.
- 6.7, last sentence on note, may not be true. We must map to Unicode character or provide mapping.
- Library community feels it needs to maintain this data.
- 6.6, Glyph with discrepancy. We need to define some tolerance. An alternate approach would be to use embedded font and ignore width. This places more burden on the checking tool, though.

The Multimedia Working Group, discussed policy issues. Excluding multimedia can be explicit if other sections are as well. If not, then it needs to be extensible. The document should include a list of multimedia types it will include. The group agreed to exclude multimedia (movie annotation and sound annotations particularly). NARA is developing procedures for 'as is' and 'to be' for use in PDF-A.

The introduction of the document should include a statement that this version of PDF-A is for static documents. Future versions of the document will address compound documents.

Embedded Files Working Group recommended not including embedded files in the PDF-A document file. The committee raised the question about internal digital signatures but feels it needs more discussion prior to including it in the document. They have recommended that incremental saves should be allowed but this was questioned. We may need to warn the users of incremental saves and that deleted pages are not really gone.

The Intellectual Property (IP) Working Group reviewed its concerns with copyright with regard to needing permission for quotes and obtaining copyright releases from Adobe for the use of PDF and Portable Document Format. PDF/A is not an Adobe copyright or trademarked word. The committee will need a statement from Adobe releasing the use of the term to the committee. PDF/X has a letter on file that allows for the use of it.

Ms. M. Warfel, Adobe Systems, is to obtain a letter with copyright releases as outlined above. Adobe released all patents for PDF-X.

The IP working group reviewed additional issues with viewer and writer intellectual property that must be addressed. With regard to patents, there should be a limit set license granted and Adobe should identify the version to be used for developing the PDF/A standard whether it is PDF 1.3 or earlier but not PDF 1.4. The IP group needed to complete their presentation later but expressed a need to be careful regarding the patents releases that will be required to ensure that archivists will be able to open the file and see what is available. (<http://partners.adobe.com/asn/developer/legalnotices.htm> for

more information) The IP Working Group requested statements from Adobe and patent agreements for 1.4. Mr. Sharpe led a discussion on the issue of royalty free software. The committee has been asked to identify the Intellectual Property requirements for the specification.

The Security Working Group raised the following issues based on their work and discussions. The committee pointed out that there is not a lot of difference between digital signature, encryption, etc. Adobe uses standard XML Namespace.

Overview of the Security group's discussions:

- Do we allow documents to be encrypted? If so, where do we store the keys?
- Do the archivists want to validate signatures?
- Digital signatures can have embedded all the data regarding the signature but the document cannot encode the signature.
- The archivists thought signatures were being disallowed which opens more patent issues.
- NARA records should not be edited but can be copied from
- Restrictive settings need to be broken out
- If we choose to include encryption, there are solutions for encryption that are not owned by Adobe and work well with Acrobat.
- The group discussed redaction and determined that the onus is on the producer.

Mr. Hughes proposed the use of an encrypted dictionary be prohibited in PDF-A. Mr. LeFurgy seconded the motion that was passed unanimously.

Secretariat is to post on the web site the documents posted by Mr. Gavin as part of the working group discussion.

Mr. McDowell recommended the use of liberal and informative notes throughout the document and posting the application notes on the web site.

The question of which working group should address digital signatures was introduced. The group agreed to restrict the specification of signatures that algorithms can specify. While archivists know how to protect classified document, the group still needs to know what archivists need to have regarding the signatures. Digital signatures appear in forms facility in Acrobat giving them access to Java script. The group had already agreed that they do not want to have Java script in the specification. The Courts will be using digital signatures but they do not know what technology will be used. The group discussed facsimiles of signatures. The process where digital signatures occur needs to be defined. Some CRC checker will be needed. There is a need to identify the signer as the person who is responsible for the records.

The group raised the question of what other objects would be required in PDF documents.

The group reviewed the definition of archival as provided by Mr. Levenson. Mr. Stone recommended that we do not need to define archive and agree upon it. Several committee members like the term long term storage instead of archive. There is a subtle difference between long term and archive. Long-term storage implies that maintaining the bit stream; the library community preserves the usability.

Mr. McDowell recommended a title for the document as ____.

The group agreed to the following as the scope statement for the document:

This International Standard specifies rules for correct use of the Portable Document Format for documents that must meet high standards for preservation of reliable visual reproduction over long periods of time. The Portable Document Format is a specification for the representation of electronic documents. It is designed for multipage documents that may contain a mix of text, raster images and vector graphics.

This standard also enables systems to easily retrieve semantic content where present. Towards this goal, this standard aims at making PDF/A compliant files as self-contained and self-documenting as possible, encapsulating all resources necessary to form an exact reproduction of the document as originally created.

The document will accept digital signatures as specified in the PDF specification but not specify their format and use. (McDowell/LeFurgy) Affirmative 24, Abstentions 2. If new signature specification is introduced, the standard will need to be revised. Digital signatures must be identifiable to a single individual and validated. The validation will verify that the signature claims to be from X but does not authenticate that the signature came from the individual.

The metadata working group summarized their discussions and recommended the format of the metadata to use XML. They noted that we will never know what metadata will be needed but metadata elements will exist for all PDF-A documents. While there will be a defined structure for the metadata it will be loose and not well defined. There will be no mandatory elements or fields of metadata for the documents. The group considered using the Dublin Core Metadata schema. They raised the issue of considering restrictions on the metadata content but the full committee did not discuss the topic other than to say that we do not care about the content, as the standard does not specify the content. The group would like to see the metadata stored as metadata keys without compression.

The XML and Metadata working groups were combined and Mr. LeFurgy and Mr. Lysakowski will co-chair the group. The Secretariat will make the necessary changes to the listserv to merge these working groups.

The Secretariat was asked to make available the list of meeting attendees with phone numbers and a statement across the top of the document "For Use Only by PDF-A Committee Members". The Secretariat was also asked to post the DIS version of PDF-X to the committee list.

A Forms Working Group was created to be chaired by Ms. Warfel, with assistance from L. Rosenthol, M. Gavin, J. Brinkema, L. Altman and R. Lysakowski. A listserv list will be created for the group.

The Hyperlinks Working Group stated that they are divided. Two papers were presented, one allowing nothing and one allowing everything.

The Unrecognized Data/Extensions and Annotations Working Groups will report at the next meeting. The committee also recognized a need to have a section on Compression. Mr. Sharpe will work on the Compression section.

The committee discussed extensibility that is not as much of an issue since we are using specific dated references. Mr. Sharpe wants to disallow the use of private key value pairs. Mr. Gavin said that if we disallowed the use of private key value pairs it would impact their business.

The committee wants to negotiate with Adobe Systems to have a technote for JPEG2000. Adobe intends to add JPEG2000 in the future at a non-disclosed date.

The next committee meeting will be Thursday and Friday, February 27 and 28, 2003 in the Washington, DC area. The meeting will start at 9 a.m. both days.

All revisions/document edits are due to Ms. Fanning by February 10, 2003. (This date was later extended to February 18, 2003.)

The meeting was adjourned at 2:51 p.m. on Friday, December 13. (Smartt/Levenson)

The following are the issues that were identified as requiring additional discussion:

- Hairlines
- Drawings
- Character spacing
- Merging of working groups – Rendering, Fonts, Tags
- Incremental saves
- Possible to add document integrity
- Concerned about patents that have not been applied for.
- Intellectual Property
- Digital Signature
- Define Archive/archivability
- Title for the standard
- XML issues
- General objects and XML embedded objects

- Links / hyperlinks (can make a link that is relevant to the present location of the document)
- XML Data vs. metadata with Acrobat draw as digital information