Frequently Asked Questions (FAQs)

ISO 19005-1:2005
PDF/A-1
Date: July 10, 2006

Statement:
This FAQ is prepared in support of ISO 19005-1:2005, Document management —
Electronic document file format for long-term preservation — Part 1: Use of PDF 1.4
(PDF/A-1)

This document is published by the PDF/A Joint Working Group working under the
auspices of TC-171 Document Management Applications Subcommittee 2 Application
Issues, with representatives from ISO Technical Committees 42, 46, 130, and 171. It
may be obtained directly from ISO (www.iso.org) or from national body standards
organizations (e.g., ANSI, BSI, DIN, AFNOR, etc.).

The following statement must accompany any distribution of this FAQ:
"This FAQ may be freely distributed and/or translated in its entirety. The current
authoritative version of this FAQ is maintained at both NPES (www.npes.org) and AIIM
(www.aiim.org)."

Frequently Asked Questions (FAQs)
ISO 19005-1:2005 (PDF/A-1)

What is PDF/A?
ISO 19005-1, Document management — Electronic document file format for long-term preservation — Part 1: Use of PDF 1.4 (PDF/A-1) is the first in a new family of ISO Standards to address the growing need to maintain information in electronic documents over archival time spans.

Because this initial version of PDF/A is based on PDF 1.4, the standard is being published in parts so that new parts can be added without obsolescing previous parts. For example: PDF/A-1 refers to the format defined by Part 1 (ISO 19005-1) of the standard (or PDF 1.4) while Part 2 (ISO 19005-2) and later parts may be based on a later version of PDF and/or may define archiving requirements for more complex content types.

Why PDF/A?
The feature-rich nature of PDF can create difficulties in preserving information over the long-term, and some useful features of the PDF file format are incompatible with the demands of long-term preservation. For example, PDF documents are not necessarily self-contained, drawing on system fonts and other content stored external to the original file. As time passes, and especially as technology changes, these external connections can be broken, and the dependencies cause information to be lost. Additionally, because of the lack of standardization among the many PDF development tools on the market, there is inconsistency in the implementation of the file format. This lack of standardization could be chaotic for the information managers of the future, especially as it would be difficult (if not impossible) for them to "get under the hood" of the PDF files unless a format specification were put in place that specifically addressed long-term preservation needs.

Tremendous quantities of valuable information are currently be created and saved all over the world as PDF, and a specification solution is needed to ensure that digital PDF documents remain readable, renderable and accessible for the long-term. PDF/A is designed to be that specification.

What is the difference between PDF and PDF/A?
The PDF/A-1 (ISO 19005-1:2005) standard is based on Adobe's PDF Reference 1.4, and specifies how to use a subset of PDF components to develop software that creates, renders and otherwise process a flavor of PDF that is more suitable for archival preservation than traditional PDF. PDF/A-1 aims to preserve the static visual appearance of electronic documents over time and also aims to support future access and future migration needs by providing frameworks for: 1) embedding metadata about electronic documents, and 2) defining the logical structure and semantic properties of electronic documents. The result is a file format, based on PDF 1.4 that is more suitable for long term preservation. PDF/A-1 files will be more self-contained, self-describing and more device-independent than traditional PDF 1.4 files.

Frequently Asked Questions (FAQs)
ISO 19005-1:2005 (PDF/A-1)

What does PDF/A-1 allow/disallow?
One of the key differences 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
- Multi-media
- JavaScript

What does PDF/A mean by "long term"?
PDF/A defines long-term as: "the period of time long enough for there to be concern about the impacts of changing technologies, including support for new media and data formats, and of a changing user community, on the information being held in a repository, which may extend into the indefinite future."

Does PDF/A-1 replace other archival file formats?
No. PDF/A-1 does not replace other archival file formats. Rather, PDF/A-1 was developed to allow PDF to be used as an archival format in a well-defined and robust manner.

What long-term preservation needs does PDF/A-1 address?
Characteristics identified as objectives for PDF/A were:

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

How does PDF/A-1 address these long-term preservation needs?

Device Independent. PDF/A-1 requires device independent components so that the static visual appearance can be reliably and consistently rendered and printed without regard to the hardware or software platform used. The graphics clause, for example, incorporates requirements to ensure predictable color rendering. PDF/A-1 also prohibits the use of components not defined in PDF Reference 1.4.

Self-Contained. Everything that is necessary to render or print a PDF/A-1 file must be contained within the file. The fonts clause requires that all fonts used for rendering content are embedded in the file. A PDF/A-1 conforming writer must always embed fonts, and a conforming reader must always use the embedded fonts. This means that the file will be rendered using the fonts intended and not using fonts residing on the local workstation.

Self-Documenting. PDF/A-1 requires Adobe Extensible Metadata Platform (XMP) be used for embedding metadata in PDF files. To allow flexibility of implementation, PDF/A-1 provides recommendations for documenting file attributes such as: File identifier, File provenance, Font metadata, and allows non-XMP schemas to be included, as long as they are embedded. Implementers can use XMP in a variety of ways to include information about electronic records within the file itself. Having metadata embedded in the file can increase the informational value of electronic documents and enhance the future researcher's understanding of the document.

Unfettered. PDF/A-1 prohibits encryption. This prohibition means that User IDs and/or Passwords are not needed to do anything with a PDF/A-1 file. PDF/A-1 files are open and available to anyone or any software that processes the file. Implementers that require access controls can provide these access controls outside of the file format.

Available. PDF/A-1 is based on an authoritative specification that is publicly available. Anyone can use the PDF Reference and XMP Specification in conjunction with PDF/A-1 to create applications that read, write, or process PDF/A-1 files. Adobe has granted a general royalty free license to use certain of its patents to create applications that process PDF/A-1 files. Additionally, Adobe has granted AIIM and NPES the rights to publish these specifications on their respective Internet sites into the indefinite future.

Adoption. PDF/A-1 was designed for flexibility of implementation to promote its wide adoption. If widely adopted, PDF/A software tools will proliferate, and the market will support the file format as long as the demand exists. Market support of PDF/A will help ensure the viability of PDF/A and extend the length of time that PDF documents can be maintained as PDF/A.

What are special font considerations?

Many fonts have restrictions on use, embedding and exchange. PDF/A requires fonts to be embedded. Therefore, organizations using PDF/A-1 must take extra precautions to be sure that the fonts they use are properly licensed to allow embedding.

When should PDF/A be used?

PDF/A should be used as a way to standardize the use of PDF for electronic document storage and ensure that these documents will be available well into the future. This is important to support business needs that require reliable rendering of electronic documents over the long term.

As a file format specification, users will need to establish their own capture methodology that meets domain specific policies and procedures (e.g., for reliability, integrity, compliance, comprehensiveness).

For example, for permanent records in PDF, US Federal agencies will need to implement PDF/A-1 in conjunction with additional requirements identified in guidance from the National Archives and Records Administration (NARA) for transferring permanent PDF records to NARA,
http://www.archives.gov/records_management/initiatives/pdf_records.html.

It is important to be aware that:
- PDF/A-1 alone does not guarantee preservation
- PDF/A-1 alone does not guarantee exact replication of source material

How many conformance levels of PDF/A-1 are there?

PDF/A-1 supports two conformance levels to promote the creation of PDF/A-1 files with rich semantic and structural information, and to also allow less complex files such as scanned images. The two levels of conformance are referred to as Level A and Level B.

- PDF/A-1a – Conformance level 1a shall adhere to all of the requirements of the PDF Reference as modified by the ISO 19005 specification and requires structural and semantic properties to be preserved. Level 1a uses "Tagged PDF" and Unicode character maps to preserve the document's logical structure and content text stream in natural reading order. For some applications, users may need to use a PDF/A-1 conformant viewer to take maximum advantage of embedded metadata.

- PDF/A-1b – Conformance level 1b requirements are intended to be those minimally necessary to ensure the visual appearance of electronic documents. Level 1a preserves document structure from the original document that allows the consumer to the archived document to view and manipulate the document as they could the original.

Will there be revisions or new parts of PDF/A?

Work has already begun on PDF/A Part 2 and is planned to be based on PDF 1.6. Implementers have requested that the following features be considered for inclusion in future parts.

- JPEG 2000 image compression
- More sophisticated digital signature support
- OpenType fonts
- 3D graphics
- Audio/video content
- Consistency with other PDF-based standards

Will new parts be compatible with older parts?

It is planned that future parts of PDF/A will be written in such a way that older PDF/A files will be compliant with newer parts of the standard. Decisions about what goes into newer parts are ultimately decided by the voting members of ISO.

It is important to note that existing parts of PDF/A will continue to exist regardless of the introduction of supplemental parts, and therefore, will support rendering of archived files into the indefinite future.

Does PDF/A-1 support digital signatures?

Yes, PDF/A-1 allows the use of digital signatures embedded in PDF as defined in the PDF Reference Manual. Because PDF digital signatures include a visual appearance, the appearance must conform to all the requirements of PDF/A-1 including font embedding and use of device independent color. Not all commercial digital signature tools follow these requirements.

What organizations intend to accept PDF/A files?

Many organizations worldwide are evaluating the advantages of using PDF/A and will provide guidance on their criteria. Some organizations that have stated their intent to accept PDF/A includes:

The US National Archives and Records Administration

The US National Archives and Records Administration (NARA) may accept transfers of permanent records in PDF/A format that additionally meet the current transfer requirements for electronic records in PDF.

The Swedish National Archives

The Swedish National Archives will refer to the PDF/A-standard as an appropriate and approved archiving format for certain types of documents. This will be announced in the technical and format specific regulation, issued by the Swedish National Archives, which is mandatory for the Swedish government authorities.

The Swedish National Archives doesn't recommend any specific product; supporting the principle that t independency from supplier-specific products is

necessary in order to attain stable and secure long-term preservation. The Swedish National Archive's policy is to refer primarily to international recognized standards, secondly to other standards (Swedish standards etc.) and finally to specific regulations prepared by the Swedish National Archive. PDF/A-1 enables preservation of certain types of documents in their original form. PDF/A-1 is not primarily to be used as a format for producing or editing documents, but as a format in which to establish documents. One of the main criteria when prescribing about document formats is independency from specific hardware and/or software.

Who are some of the vendors that report PDF/A-1 support?

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

Why does the standards development process take so long?

Developing a standard according to National Standards bodies and ISO policies ensure that the process is an open and consensus driven process. Several participants with expertise in the area come together to contribute to development of the standard.
In order to ensure that there is ample opportunity for all interested parties to provide input and express opinions there is a very formal balloting process determined by the standards organization.

How did PDF/A get started and who is involved?

The PDF/A activity was initiated in the United States through the joint sponsorship of AIIM, Association for Information and Image Management and NPES, The Association for Suppliers of Printing, Publishing, and Converting Technologies. Under the auspices of TC-171 Document Management Application ns Subcommittee 2 Application Issues, a Joint Working Group (WG5) was formed with representatives from ISO Technical Committees 42, 46, 130, and 171. A diverse group of librarians, archivists, PDF software developers, government agencies, imaging experts, graphics experts and others collaborated to develop PDF/A. Initial meetings were held in mid-2002 and the standard was approved in June 2005. Technical experts from 15 national standards bodies provided input throughout the development process.

Currently actively involved organizations are: AIIM, NPES, Adobe, Appligent, Callas, EMC/Documentum, Global Graphics, Harvard University, Merck & Co., Inc., PDF Sages, the Administrative Office of the US Courts, the Internal Revenue Service and the National Archives and Records Administration.

Countries currently involved in this ISO activity include: Australia, China, Germany, France, Japan, the Netherlands, Switzerland, the United Kingdom, the United States and Sweden.

Where can I get more information?

There is a considerable amount of information about PDF/A listed on the AIIM website at http://www.aiim.org/standards. There are copies of initial requirements, meeting minutes, meeting notes and presentation materials at this site. In addition both AIIM and NPES have posted the PDF 1.4 Reference and the XMP Specification:

www.aiim.org/standards

www.npes.org/standards

Where can I purchase a copy of the PDF/A standard?

The published PDF/A-1 standard may be purchased directly from ISO or from national standards bodies around the world.

http://www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=38920&ICS1=37&ICS2=100&ICS3=99

www.ansi.org

How can I get involved?

AIIM and NPES welcomes representatives from vendors, user organizations and users themselves. If you think you can help contribute to the development of this standard please contact AIIM or NPES

www.aiim.org

www.npes.org