

Getting Started With PDF/A Output Format

PDF/A is a newly-approved standard from the International Standards Organization (ISO). The new ISO 19005-1 standard that defines PDF/A format is the digital document standard for the long-term preservation of required tax, legal, historical, transactional and other important documents.

Archiving in PDF/A format to satisfy your corporate, HIPAA, Sarbanes-Oxley or other legal requirements is the best way to ensure that you will be able to search, retrieve, view and print your documents in the future with the full, legal fidelity of the original. PDF/A format will remain compatible as PDF technology evolves.

This is why ISO, in conjunction with the US Courts, Library of Congress, ARMA, AIIM, NARA and many more corporations, governments, and other major institutions around the world have approved this new standard for PDF files for long-term archiving of documents. This new standard defines a subset of PDF 1.4 format and composition rules that will withstand the test of time.

PDF/A is designed to be:

- Device and software independent
- Self-contained: Includes all the resources needed to render each page (embedded fonts)
- Self-documented: Capable of including sufficient XMP metadata for auto-classification and indexing.
- A framework for representing the logical document structure (PDF/A-1a)

Basically, it inhibits the use of PDF functionality that may cause problems in future releases. If you generate PDF documents from a product that complies with this new PDF/A standard, your files will be safe and viable for the years to come.

The best way to generate PDF/A documents is directly from Adobe Acrobat V7.08 or above. For those who have applications that cannot generate PDF/A directly, but can only generate print streams for HP PCL printers (and, do not want to print, scan, convert to PDF/A, then OCR for indexing and text searchability), you need one of PageTech's product options to convert your PCL into PDF/A format.

Both PCLTool SDK V8.3 and PCLWorks Program V8.3 can produce PDF/A output. By default, we support PDF/A-1b because we feel confident that we can generate this level of PDF/A format from virtually any PCL file. However, you can switch to PDF/A-1a in order to see if we can achieve this level of compliance with your PCL. PDF/A-1a compliance is more dependent on the structure and composition of the document. Consequently, you may have PCL that we are unable to convert into PDF/A-1a format at this time. We want to see all the PCL that produces PDF/A files that are not in compliance with the ISO standard. Please use the <http://www.pagetech.com/support/supportform.php> to submit any files you think we should review.

The following is a list of the enhancements added to our various programs and other files:

PCLTool SDK (PCLTool.exe) and PCLWorks Program (PCLWorks.exe):

PDF Preferences Dialog:

- Above "Document Info" added "**Compliance Format**": "PDF/A-1a" or "**PDF/A-1b**"
- Change "Subject" to "Subject/Description"
- Comma-delimit all keywords provided when saving in the various PDF metadata fields.
- Added "Description Writer" field
- Added "Copyright Notice" field
- Added "Copyright/WebStatement URL" field
- Added "Copyright Status" field

IMG2PDF.exe:

PDF Preferences Dialog:

Above "Document Info" added "**Compliance Format**": "PDF/A-1a" or "**PDF/A-1b**"
Change "Subject" to "Subject/Description"
Comma-delimit all keywords provided when saving in the various PDF metadata fields.
Added "Description Writer" field
Added "Copyright Notice" field
Added "Copyright/WebStatement URL" field
Added "Copyright Status" field

IMGCvt.dll: Internal changes.

IMGCvt_Net.dll: Internal changes.

PDFWrite.dll: Internal changes.

PTC32.dll: Internal changes.

PTCNet.dll: Internal changes.

PTC32.ini: Added "PdfALevel" preference.

```
[SavedValues]
PdfALevel=0
;      DEF:>=0      0=no PDF/A
;
;                  1=PDF/A Conformance level 1-B
;                  2=PDF/A Conformance level 1-A
```

APP.ini: Added several new preferences.

```
[PCL2PDF]
PdfKeepInputExt=0
;      DEF:>=0
;
PdfDescWriter=
;
;                  =Description Writer
;
PdfCopyrightNotice=
;
;                  =Copyright Notice
;
PdfWebStatement=
;
;                  =Copyright URL
;
PdfCopyrightStatus=
;      DEF:>=0      0=Unknown Copyright Status
;                  1=Copyrighted
;                  2=Public Domain Status
;
PdfLevelx10=PdfLevelx14
;      DEF:>=PdfLevelx14  =PDF V1.4
;      PdfLevelx13      =PDF V1.3
;      PdfLevelx15      =PDF V1.5
;      PdfLevelx16      =PDF V1.6
```

```
;  
PdfA=0  
;DEF:>=0          0=none  
;                  1=PDF/A-1b  
;                  2=PDF/A-1a  
;
```

PCLXForm.exe: Changes internally to interpret new PDF/A script (.TPT) functions and preferences.

```
[PdfParams]  
DescWriter      =Description Writer  
CopyrightNotice =Copyright Notice  
WebStatementURL =URL where copyright information resides  
CopyrightStatus =Unknown  
;                unless marked Copyrighted or Public Domain  
PdfVersion      =PdfLevelx14  
;                V1.4, or 13=V1.3, 15=V1.5 or 16=V1.6  
PdfALevel       =0  
;                None, 1=1b, 2=1a
```

Some of the tools we used to test our PDF/A compliance are listed below:

- Adobe Acrobat Professional V7.08 or V8.0 in the Advanced -> Preflight dialog.

NOTE: Acrobat's PDF/A-1b compatibility test is behind the most current ISO-19005-1 specification and flags all of our PDF/A files with an error message as follows:

"Info does not match metadata according to PDF/A".

We decided to make our PDF/A's according to the most current specification and that pass PDF/A compliance tools testing for the correct metadata formatting, rather than make out-of-spec files that Adobe passes as valid. We found Apago's PDF Appraiser to have the most up-to-date adherence to the PDF/A specifications.

- Apago's PDF Appraiser V1.0 at www.apagoinc.com.

As with any new, complex file format, there are going to be different interpretations regarding how to implement the standard. If we create a PDF/A file that does not pass any of the many compliance-checking programs. Please use the <http://www.pagetech.com/support/supportform.php> to submit the input PCL or TIFF files and the PDF/A files we output.