

1. Introduction

PDF/X and PDF/A are umbrella terms used to denote several ISO standards that define different subsets of the PDF standard. The objective of PDF/X is to facilitate graphics exchange between document creator and printer and therefore, has all requirements related to printing. For instance, in PDF/X-1a, all fonts need to be embedded and all images need to be CMYK or spot colors. PDF/X-2 and PDF/X-3 accept calibrated RGB and CIELAB colors along with all other specifications of PDF/X-1a.

PDF/A defines a profile for archiving PDF documents which ensures the documents can be reproduced the exact same way in years to come, a key element to achieve this is that the PDF/A documents shall be 100% self contained. All the information needed to display the document in the same manner every time, is embedded in the file. PDF/A document is not permitted to be reliant on information from external sources. Other restrictions include avoidance of audio/video content, JavaScript and encryption. Mandatory inclusion of fonts, color profile and standards based metadata are absolutely essential for PDF/A.

This package currently supports generation of PDF/X-1a and PDF/A-1b compliant documents using PDF_T_EX. More and more standards will be included in future.

2. Usage

The file, namely `pdfx.dtx` is a composite document of program code and documentation in \LaTeX format in the tradition of *literate programming*. You can extract the program code alone by stripping off the documentation part by running \LaTeX or \TeX over the installer namely, `pdfx.ins` which is also provided with this file. To get the documentation which you are reading now, you need to run (PDF) \LaTeX over the file, `pdfx.dtx`.

2.1. Data file for XMP metadata

As you are aware, standards compliant PDF documents need XMP metadata to be included. In order to create XMP in the prescribed XML format, a simple data file holding the meta information of the document needs to be created either through a program or by hand. For our purposes, we name it as `\jobname.xmpdata`, a simple example of which will look like the following:

```
\Keywords{pdfTeX</rdf:li><rdf:li>PDF/X-1a</rdf:li><rdf:li>PDF/A-b}
\Title{Sample LaTeX input file}
\Author{LaTeX project team}
\Org{TeX Users Group}
```

You may note that the keywords are separated by XML elements `</rdf:li><rdf:li>` instead of comma character. This is the correct format required by the XMP metadata which is in XML format. Similarly, several other kinds of data can be captured using the following commands:

1. `\Subject`
2. `\Creator`
3. `\Producer`
4. `\Volume`
5. `\Issue`
6. `\CoverDisplayDate`

7. `\CoverDate`
8. `\Copyright`
9. `\Doi`
10. `\Lastpage`
11. `\Firstpage`
12. `\Journaltitle`
13. `\Journalnumber`
14. `\CreatorTool`
15. `\AuthoritativeDomain`

The above commands are self-explanatory. Users can resort to alternate ways to create `xmp` file for inclusion in PDF. However, minimal `\jobname.xmpdata` shall be created with `\Title` and `\Author` commands along with their corresponding values for `pdfx` package to work correctly. You may check [Adobe XMP Development Center](#) for more exhaustive information about Extensible Metadata Platform (XMP). An XMP Toolkit SDK which supports GNU/Linux, Macintosh and Windows operating systems is also provided under BSD licence.

`pdfx` makes use of `xmpincl` package to include `xmp` data into the PDF. A good look at the documentation of `xmpincl` package will greatly help the users to understand the process of `xmp` data inclusion.

2.2. Limitations and dependencies

`pdfx.sty` works only with PDF_TE_X. It further depends on the following packages:

1. `xmpincl` for insertion of metadata into PDF.
2. `hyperref` for hyperlinking, bookmarks, etc.
3. `glyphtounicode.tex` maps glyph names to corresponding Unicode.
4. `glyphtounicode-cmr.tex` does the same for `cmr` fonts.

Necessary color profile files may be obtained from International Color Consortium. Please take a look at <http://www.color.org/iccprofile.xalter>.

2.3. Files included

Following files are included in the archive:

1. `pdfx.dtx` — composite package and documentation.
2. `pdfx.ins` — installer batch file.
3. `pdfx-1a.xmp` — specimen `xmp` template for PDF/X-1a.
4. `pdfa-1b.xmp` — specimen `xmp` template for PDF/A-1b.
5. `small2e.xmpdata` — specimen data file to provide values relating to the document to generate metadata.

A directory named `pdfx` may be created under `$TEXMF/tex/latex` and all `*.sty` and `*.xmp` may be moved to the same. `TEX`'s file database may need to be updated by suitable command depending on your distribution and operating system.

2.4. Options

The package can be loaded with the command:

```
\usepackage[<option>]{pdfx}
```

where the options are:

x-1a generates PDF/X-1a compliant PDF.

a-1b generates PDF/A-1b compliant PDF.

2.5. Useful production notes

We have included some useful notes about production problems which we have encountered while generating PDF/A-1b compliant documents and the fixes recommended at: http://support.river-valley.com/wiki/index.php?title=Generating_PDF/A_compliant_PDFs_from_pdftex.

2.6. Miscellaneous information

The package is released under \LaTeX Project Public Licence. Bug reports, suggestions, feature requests, etc., may be sent to the authors at cvr@river-valley.org and/or thanh@river-valley.org.

3. Implementation

3.1. Various auxiliary macros

Two booleans are defined to switch between two options, **a-1b** and **x-1a**. PDF/X-1a further demands PDF version 1.3 and properly placed mediabox, bleedbox and trimbox (innermost) in that order. The MediaBox defines the size of the entire document, either the ArtBox or the TrimBox, defines the extent of the printable area. If the file is to be printed with bleed, a BleedBox, which must be larger than the TrimBox/ArtBox, but smaller than the MediaBox, must be defined.

```
1 \<package>
2 %
3 % $Id: pdfx.dtx,v 1.1 2008/12/01 10:28:01 cvr Exp cvr $
4 %
5 \NeedsTeXFormat{LaTeX2e}
6 \ProvidesPackage{pdfx}
7 [2008/12/01 v1.1 PDF/X and PDF/A support (CVR/HTH)]
8
9 \newif\ifpdfxonea \pdfxoneafalse
10 \newif\ifpdfaoneb \pdfaonebfalse
11
12 \DeclareOption{a-1b}{\global\pdfaonebtrue}
13 \DeclareOption{x-1a}{\global\pdfxoneatrue}
14 \ProcessOptions
15
16 \ifpdfxonea
17 \pdfminorversion=3
18 \pdfpageattr{/MediaBox[0 0 595 793]
19 /BleedBox[0 0 595 793]
20 /TrimBox[25 20 570 773]}
```

```
21 \else
22 \pdfminorversion=4
23 \fi
```

Several macros were defined to capture data for the XMP metadata to be inserted into the PDF during generation.

```
24 \def\hash{\expandafter\@gobble\string\#}
25 \def\amp{\expandafter\@gobble\string\&}
26 \def\xmpAmp{\amp\hash x0026;}
27 \def\TextCopyright{\amp\hash x00A9;}
28 \def\Title#1{\gdef\xmpTitle{#1}}
29 \let\xmpTitle\@empty
30 \def\Author#1{\gdef\xmpAuthor{#1}}
31 \let\xmpAuthor\@empty
32 \def\Keywords#1{\gdef\xmpKeywords{#1}}
33 \let\xmpKeywords\@empty
34 \let\xmpSubject\xmpKeywords
35 \def\Creator#1{\gdef\xmpCreator{#1}}
36 \def\xmpCreator{\@pdfcreator}
37 \def\Producer#1{\gdef\xmpProducer{#1}}
38 \def\xmpProducer{pdfTeX}
39 \def\Volume#1{\gdef\xmpVolume{#1}}
40 \let\xmpVolume\@empty
41 \def\Issue#1{\gdef\xmpIssue{#1}}
42 \let\xmpIssue\@empty
43 \def\CoverDisplayDate#1{\gdef\xmpCoverDisplayDate{#1}}
44 \let\xmpCoverDisplayDate\@empty
45 \def\CoverDate#1{\gdef\xmpCoverDate{#1}}
46 \let\xmpCoverDate\@empty
47 \def\Copyright#1{\gdef\xmpCopyright{#1}}
48 \let\xmpCopyright\@empty
49 \def\Doi#1{\gdef\xmpDoi{#1}}
50 \let\xmpDoi\@empty
51 \def\Lastpage#1{\gdef\xmpLastpage{#1}}
52 \let\xmpLastpage\@empty
53 \def\Firstpage#1{\gdef\xmpFirstpage{#1}}
54 \let\xmpFirstpage\@empty
55 \def\Journaltitle#1{\gdef\xmpJournaltitle{#1}}
56 \let\xmpJournaltitle\@empty
57 \def\Journalnumber#1{\gdef\xmpJournalnumber{#1}}
58 \let\xmpJournalnumber\@empty
59 \def\Org#1{\gdef\xmpOrg{#1}}
60 \let\xmpOrg\@empty
61 \def\CreatorTool#1{\gdef\xmpCreatorTool{#1}}
62 \def\xmpCreatorTool{\xmpProducer}
63 \def\AuthoritativeDomain#1{\gdef\xmpAuthoritativeDomain{#1}}
64 \let\xmpAuthoritativeDomain\@empty
```

3.2. Document and instance ID's

Document ID and instance ID are created from values obtained from `\jobname.pdf` and `\pdfcreationdate` by making use of `\pdfmdfivesum` primitive of PDF_TE_X.

```
\findUUID
\uuid
65 \def\findUUID#1{\edef\tmpstring{\pdfmdfivesum{#1}}
66 \expandafter\eightofnine\tmpstring\end}
67 \def\eightofnine#1#2#3#4#5#6#7#8#9\end{%
68 \xdef\eightchars{#1#2#3#4#5#6#7#8}
```

```
69 \fourroffive#9\end}
70 \def\fourroffive#1#2#3#4#5\end{\xdef\ffourchars{#1#2#3#4}
71 \sfourroffive#5\end}
72 \def\sfourroffive#1#2#3#4#5\end{\xdef\sfourchars{#1#2#3#4}
73 \tfourroffive#5\end}
74 \def\tfourroffive#1#2#3#4#5\end{\xdef\tfourchars{#1#2#3#4}
75 \xdef\laststring{#5}}
76
77 \def\uuid{\eightchars-%
78 \ffourchars-%
79 \sfourchars-%
80 \tfourchars-%
81 \laststring}
82
83 \findUUID{\jobname.pdf}
84 \edef\xmpdocid{\uuid}
85 \findUUID{\pdfcreationdate}
86 \edef\xmpinstid{\uuid}
```

`\jobname.xmpdata` is read if available and the package, `xmpincl` is also loaded which will take care of inserting metadata into the PDF document.

```
87 \InputIfFileExists{\jobname.xmpdata}{}{}
88 \RequirePackage{xmpincl}
```

`\convDate`
`\convertDate`

The date format needed by metadata is different from the value provided by the `\pdfcreationdate`. `\convertDate` macro generates the required date format from `\pdfcreationdate`.

```
89 \def\convertDate{\getYear}
90 {\catcode'\D=12
91 \gdef\getYear D:#1#2#3#4{\edef\xYear{#1#2#3#4}\getMonth}
92 }
93 \def\getMonth#1#2{\edef\xMonth{#1#2}\getDay}
94 \def\getDay#1#2{\edef\xDay{#1#2}\getHour}
95 \def\getHour#1#2{\edef\xHour{#1#2}\getMin}
96 \def\getMin#1#2{\edef\xMin{#1#2}\getSec}
97 \def\getSec#1#2{\edef\xSec{#1#2}\getTZh}
98 \def\getTZh +#1#2{\edef\xTZh{#1#2}\getTZm}
99 \def\getTZm '#1#2' {%
100 \edef\xTZm{#1#2}%
101 \edef\convDate{\xYear-\xMonth-\xDay
102 T\xHour:\xMin:\xSec+\xTZh:\xTZm}}
103 \expandafter\convertDate\pdfcreationdate
```

3.3. Color profiles

`/OutputIntents`

For better color management, PDF/X-1a and PDF/A-1b need an ICC profile included in the document. An ICC profile is a set of data that characterizes a color input or output device, or a color space, according to standards promulgated by the International Color Consortium (ICC). Profiles describe the color attributes of a particular device or viewing requirement by defining a mapping between the device source or target color space and a profile connection space. For PDF/X-1a, we have included the ICC profile namely, `F0GRA39L.icc` which is for CMYK data and for PDF/A-1b, we have used `sRGBIEC1966-2.1.icm` for RGB data. You can change the value of the `file` attribute in the code below to use different color profile files.

```
104 \ifpdfxonea
```

```
105 \def\pctchar{\expandafter@gobble\string\%}
106 \def\bchar{\expandafter@gobble\string\}
107 \immediate\pdfobj stream attr{/N 4} file{FOGRA39L.icc}
108 \edef\OBJ@CVR{\the\pdflastobj}
109 \pdfcatalog{/OutputIntents [ <<
110 /Type/OutputIntent
111 /S/GTS_PDFX
112 /OutputCondition (FOGRA39)
113 /OutputConditionIdentifier (FOGRA39 \@bchar(ISO Coated v2
114 300\pctchar\space \@bchar(ECI\@bchar)\@bchar))
115 /DestOutputProfile \OBJ@CVR\space 0 R
116 /RegistryName(http://www.color.org)
117 >> ]}
118 \else
119 \immediate\pdfobj stream attr{/N 4} file{sRGBIEC1966-2.1.icm}
120 \edef\OBJ@RVT{\the\pdflastobj}
121 \pdfcatalog{%
122 /OutputIntents [ <<
123 /Type /OutputIntent
124 /S/GTS_PDFA1
125 /DestOutputProfile \OBJ@RVT\space 0 R
126 /OutputConditionIdentifier (sRGB IEC61966-2.1)
127 /Info(sRGB IEC61966-2.1)
128 >> ]}
129 \fi
```

One of the `xmp` files is selectively loaded based on the option chosen.

```
130 \begingroup
131 \let\&=\xmpAmp
132 \ifpdfxonea
133 \includexmp{pdfx-1a}
134 \else
135 \includexmp{pdfa-1b}
136 \fi
137 \endgroup
```

`glyptounicode.tex` and `glyptounicode-cmr.tex` are read. These files contain mapping from glyph names to corresponding unicode for embedded fonts, which are required by PDF/A-1b. `glyptounicode.tex` covers AGL (Adobe Glyph List), names from `texglyphlist.txt` (part of `lcdf-typetools`) and `zapfdingbats.txt`, plus a few exceptions. `glyptounicode-cmr.tex` covers glyphs that are used in CM fonts but not listed in `glyptounicode.txt`; the mappings come from file `goadb998.nam` (part of \TeX Gyre fonts).

```
138 \input glyptounicode.tex
139 \input glyptounicode-cmr.tex
140 \pdfgentounicode=1
```

Active content is not allowed in a PDF/X-1a file. This means that standard PDF features like forms, signatures, comments and embedded sounds and movies are not allowed in PDF/X-1a. So `hyperref` is loaded in `draft` mode and `info` dictionary is defined with `\pdfinfo` command.

```
141 \ifpdfxonea
142 \RequirePackage[draft,pdftex,pdfpagemode=UseNone,bookmarks=false]{hyperref}
143 \pdfinfo{
144 /Title(\xmpTitle)%
145 /Author(\xmpAuthor)%
```

```

146     /Creator(\xmpProducer)%
147     /CreationDate(\convDate)%
148     /ModDate(\convDate)%
149     /Producer(\xmpProducer)%
150     /Trapped /False
151     /GTS_PDFXVersion (PDF/X-1:2001)%
152     /GTS_PDFXConformance (PDF/X-1a:2001)%
153 }
154 \else

```

For PDF/A-1b, load `hyperref` package with `pdfa` option, so that it will take care of the link annotations correctly. We have slightly modified the `\pdfinfo` by including `/GTS_PDFa1Version (PDA/A-1b:2005)`. Take a look at the modified `\PDF@FinishDoc` macro of `hyprref`.

```

155 \RequirePackage[pdftex,pdfa]{hyperref}
156 \def\PDF@FinishDoc{%
157 \Hy@UseMaketitleInfos
158 \pdfinfo{%
159     /Creator(\xmpProducer)%
160     \ifx\@pdfcreationdate\@empty
161     \else
162     /CreationDate(\@pdfcreationdate)%
163     \fi
164     \ifx\@pdfmoddate\@empty
165     \else
166     /ModDate(\@pdfmoddate)%
167     \fi
168     /Producer(\xmpProducer)%
169     /Trapped /False
170     /GTS_PDFa1Version (PDF/A-1b:2005)%
171 }%
172 \Hy@DisableOption{pdfauthor}%
173 \Hy@DisableOption{pdftitle}%
174 \Hy@DisableOption{pdfsubject}%
175 \Hy@DisableOption{pdfcreator}%
176 \Hy@DisableOption{pdfcreationdate}%
177 \Hy@DisableOption{pdfmoddate}%
178 \Hy@DisableOption{pdfproducer}%
179 \Hy@DisableOption{pdfkeywords}}
180 \fi
181 %
182 \end{package}

```

4. Index

Numbers written in *italic* refer to the page where the corresponding entry is described; numbers underlined refer to the code line of the definition; numbers in **roman** refer to the code lines where the entry is used.

| Symbols | | |
|------------------------|---------------|---|
| <code>\#</code> | 20 | <code>\@pdfcreationdate</code> 156, 158 |
| <code>\%</code> | 101 | <code>\@pdfcreator</code> 32 |
| <code>\&</code> | 21, 127 | <code>\@pdfmoddate</code> 160, 162 |
| <code>\@bchar</code> | 102, 109, 110 | A |
| <code>\@pctchar</code> | 101, 110 | <code>\amp</code> 21–23 |

Generation of PDF/X-1a and PDF/A-1b compliant PDF's with PDF_T_EX — pdfx.sty

C. V. RADHAKRISHNAN and HÀN THÊ THÀNH

QUICK LINKS

▶ Introduction

▶ Usage

▶ Implementation

▶ Index

| | | | | | |
|--|----------------------|---|--------------|--|----------|
| <code>\Author</code> | 26 | <code>\laststring</code> | 71, 77 | | |
| <code>\AuthoritativeDomain</code> | 59 | | | | |
| C | | | | | |
| <code>\catcode</code> | 86 | <code>\newif</code> | 5, 6 | | |
| <code>\convDate</code> | 97, 143, 144 | O | | | |
| <code>\convertDate</code> | 85, 99 | <code>\OBJ@CVR</code> | 104, 111 | | |
| <code>\Copyright</code> | 43 | <code>\OBJ@RVT</code> | 116, 121 | | |
| <code>\CoverDate</code> | 41 | <code>\Org</code> | 55 | | |
| <code>\CoverDisplayDate</code> | 39 | P | | | |
| <code>\Creator</code> | 31 | <code>\PDF@FinishDoc</code> | 152 | | |
| <code>\CreatorTool</code> | 57 | <code>\pdfaonebfalse</code> | 6 | | |
| D | | | | <code>\pdfaonebtrue</code> | 8 |
| <code>\D</code> | 86 | <code>\pdfcatalog</code> | 105, 117 | | |
| <code>\DeclareOption</code> | 8, 9 | <code>\pdfcreationdate</code> | 81, 99 | | |
| <code>\Doi</code> | 45 | <code>\pdfgentounicode</code> | 136 | | |
| E | | | | <code>\pdfinfo</code> | 139, 154 |
| <code>\eightchars</code> | 64, 73 | <code>\pdflastobj</code> | 104, 116 | | |
| <code>\eightofnine</code> | 62, 63 | <code>\pdfmdfivesum</code> | 61 | | |
| F | | | | <code>\pdfminorversion</code> | 13, 18 |
| <code>\ffourchars</code> | 66, 74 | <code>\pdfobj</code> | 103, 115 | | |
| <code>\findUUID</code> | 61, 79, 81 | <code>\pdfpageattr</code> | 14 | | |
| <code>\Firstpage</code> | 49 | <code>\pdfxoneafalse</code> | 5 | | |
| <code>\fouroffive</code> | 65, 66 | <code>\pdfxoneatru</code> | 9 | | |
| G | | | | <code>\ProcessOptions</code> | 10 |
| <code>\getDay</code> | 89, 90 | <code>\Producer</code> | 33 | | |
| <code>\getHour</code> | 90, 91 | R | | | |
| <code>\getMin</code> | 91, 92 | <code>\RequirePackage</code> | 84, 138, 151 | | |
| <code>\getMonth</code> | 87, 89 | S | | | |
| <code>\getSec</code> | 92, 93 | <code>\sfourchars</code> | 68, 75 | | |
| <code>\getTZh</code> | 93, 94 | <code>\sfouroffive</code> | 67, 68 | | |
| <code>\getTZm</code> | 94, 95 | T | | | |
| <code>\getYear</code> | 85, 87 | <code>\TextCopyright</code> | 23 | | |
| H | | | | <code>\tfourchars</code> | 70, 76 |
| <code>\hash</code> | 20, 22, 23 | <code>\tfouroffive</code> | 69, 70 | | |
| <code>\Hy@DisableOption</code> | 168–175 | <code>\Title</code> | 24 | | |
| <code>\Hy@UseMaketitleInfos</code> | 153 | <code>\tmpstring</code> | 61, 62 | | |
| I | | | | U | |
| <code>\ifpdfaoneb</code> | 6 | <code>\uuid</code> | 73, 80, 82 | | |
| <code>\ifpdfxonea</code> | 5, 12, 100, 128, 137 | V | | | |
| <code>\includemp</code> | 129, 131 | <code>\Volume</code> | 35 | | |
| <code>\InputIfFileExists</code> | 83 | X | | | |
| <code>\Issue</code> | 37 | <code>\xDay</code> | 90, 97 | | |
| J | | | | <code>\xHour</code> | 91, 98 |
| <code>\Journalnumber</code> | 53 | <code>\xMin</code> | 92, 98 | | |
| <code>\Journaltitle</code> | 51 | <code>\xMonth</code> | 89, 97 | | |
| K | | | | <code>\xmpAmp</code> | 22, 127 |
| <code>\Keywords</code> | 28 | <code>\xmpAuthor</code> | 26, 27, 141 | | |
| L | | | | <code>\xmpAuthoritativeDomain</code> | 59, 60 |
| <code>\Lastpage</code> | 47 | <code>\xmpCopyright</code> | 43, 44 | | |
| | | <code>\xmpCoverDate</code> | 41, 42 | | |
| | | <code>\xmpCoverDisplayDate</code> | 39, 40 | | |
| | | <code>\xmpCreator</code> | 31, 32 | | |

Generation of PDF/X-1a and PDF/A-1b compliant PDF's with PDF_TE_X — pdfx.sty

C. V. RADHAKRISHNAN and HÀN THÊ THÀNH

QUICK LINKS

▶ Introduction

▶ Usage

▶ Implementation

▶ Index

| | | | |
|--------------------------------------|--------|---------------------------------|--------------------------------|
| <code>\xmpCreatorTool</code> | 57, 58 | <code>\xmpOrg</code> | 55, 56 |
| <code>\xmpdocid</code> | 80 | <code>\xmpProducer</code> | 33, 34, 58, 142, 145, 155, 164 |
| <code>\xmpDoi</code> | 45, 46 | <code>\xmpSubject</code> | 30 |
| <code>\xmpFirstpage</code> | 49, 50 | <code>\xmpTitle</code> | 24, 25, 140 |
| <code>\xmpinstid</code> | 82 | <code>\xmpVolume</code> | 35, 36 |
| <code>\xmpIssue</code> | 37, 38 | <code>\xSec</code> | 93, 98 |
| <code>\xmpJournalnumber</code> | 53, 54 | <code>\xTZh</code> | 94, 98 |
| <code>\xmpJournaltitle</code> | 51, 52 | <code>\xTZm</code> | 96, 98 |
| <code>\xmpKeywords</code> | 28–30 | <code>\xYear</code> | 87, 97 |
| <code>\xmpLastpage</code> | 47, 48 | | |