

The l3pdfmeta package pdf-standards and XMP-metadata

The L^AT_EX3 Project*

Released XXXX-XX-XX

1 l3pdfmeta documentation

This module sets up some tools and commands needed for pdf standards in general. The goal is to collect the requirements and to provide code to check and fulfill them.

In future it will probably also contain code to setup XMP-metadata needed by the various standards in special. Currently XMP-metadata can be added by two mutual incompatible packages: `hyperxmp` and `pdfx`. Both packages are currently also incompatible with the pdf resource management, but for `hyperxmp` some patches are provided.

`\g_pdfmeta_standard_prop`

This public property is filled with the requirements of the requested standard(s). packages wanting to test a requirement can query the variable. For some of the values it also makes sense to change them. Currently this is only relevant for the colorprofile: With one of the pdf/A standards a colorprofile is embedded and added to the `/OutputIntent`. The default value is `sRGB.icc`, this can be changed by setting the key `outputintent_profile`. Currently the only alternative is `FOGRA39L_coated.icc`. Interfaces to add more colorprofiles will be added later.

`\g_pdfmeta_standard_pdf/A-1b_prop`
`\g_pdfmeta_standard_pdf/A-2b_prop`
`\g_pdfmeta_standard_pdf/A-3b_prop`

These three properties contain the settings for the three pdf/A-standards.

2 l3pdfmeta implementation

```
1 <*package>
2 <@@=pdfmeta>
3 \ProvidesExplPackage {l3pdfmeta} {2020-05-17} {0.2}
4   {XMP-Metadata and PDF-Standards}
```

Message for unknown standards

```
5 \msg_new:nnn {pdf }{unknown-standard}{The~standard~'#1'~is~unknown~and~has~been~ignored}
```

```
\l__pdfmeta_tmpa_tl
\l__pdfmeta_tmpa_str
```

*E-mail: latex-team@latex-project.org

```

6 \tl_new:N\l__pdfmeta_tmpa_tl
7 \str_new:N \l__pdfmeta_tmpa_str

(End definition for \l__pdfmeta_tmpa_tl and \l__pdfmeta_tmpa_str.)

```

3 Standards (work in progress)

3.1 Tools and tests

This public property will contain the settings for the document.

`\g_pdfmeta_standard_prop`

```

8 \prop_new:N \g_pdfmeta_standard_prop

(End definition for \g_pdfmeta_standard_prop. This variable is documented on page 1.)

9 \hook_gput_code:nnn {begindocument} {pdf}
10 {
11   \prop_item:Nn \g_pdfmeta_standard_prop { annot_flags }
12 }

```

3.2 pdf/A

`\g_pdfmeta_standard_pdf/A-1b_prop`

`\g_pdfmeta_standard_pdf/A-2b_prop`

`\g_pdfmeta_standard_pdf/A-3b_prop`

```

13 \prop_new:c { g_pdfmeta_standard_pdf/A-1b_prop }
14 \prop_set_from_keyval:cn { g_pdfmeta_standard_pdf/A-1b_prop }
15 {
16   ,name           = pdf/A-1b
17   ,type           = A
18   ,year           = 2005
19   ,pdf_version    = 1.4           %minimum
20   ,encryption     = false
21   ,external_content = false % no F, FFilter, or FDecodeParms in stream dicts
22   ,embed_content  = false, % no EF key in filespec, no /Type/EmbeddedFiles
23   ,max_string_size = 65535
24   ,max_array_size = 8191
25   ,max_dict_size  = 4095
26   ,max_obj_num    = 8388607
27   ,max_nest_qQ    = 28
28   ,named_actions  = {NextPage, PrevPage, FirstPage, LastPage}
29   ,annot_flags    = {
30     \bitset_set_true:Nn \l_pdfannot_F_bitset {Print}
31     \bitset_set_false:Nn \l_pdfannot_F_bitset {Hidden}
32     \bitset_set_false:Nn \l_pdfannot_F_bitset {Invisible}
33     \bitset_set_false:Nn \l_pdfannot_F_bitset {NoView}
34     \pdfannot_dict_put:nnn {link/URI}{F}{ \bitset_to_arabic:N \l_pdfann
35     \pdfannot_dict_put:nnn {link/GoTo}{F}{ \bitset_to_arabic:N \l_pdfan
36     \pdfannot_dict_put:nnn {link/GoToR}{F}{ \bitset_to_arabic:N \l_pdfa
37     \pdfannot_dict_put:nnn {link/Launch}{F}{ \bitset_to_arabic:N \l_pdfa
38     \pdfannot_dict_put:nnn {link/Named}{F}{ \bitset_to_arabic:N \l_pdfa
39   }
40   %booleans. Only the existence of the key matter.
41   %If the entry is added it means true (so in most cases "don't use ...")
42   %

```

```

43 %=====
44 % Rule 6.1.13-1 CosDocument, isOptionalContentPresent == false
45 ,Catalog_no_OCProperties =
46 %=====
47 % Rule 6.6.1-1: PDAAction, S == "GoTo" || S == "GoToR" || S == "Thread" || S == "URI" || S
48 % means: no /S/Launch, /S/Sound, /S/Movie, /S/ResetForm, /S/ImportData, /S/JavaScript, /S
49 ,annot_action_no_A =
50 %=====
51 % Rule 6.6.2-1: PDAnnot, Subtype != "Widget" || AA_size == 0
52 % means: no AA dictionary
53 ,annot_widget_no_AA =
54 %=====
55 % Rule 6.9-2: PDAnnot, Subtype != "Widget" || (A_size == 0 && AA_size == 0)
56 % (looks like a tightening of the previous rule)
57 ,annot_widget_no_A_AA =
58 %=====
59 % Rule 6.9-1 PDAcroForm, NeedAppearances == null || NeedAppearances == false
60 ,form_no_NeedAppearances =
61 %=====
62 %Rule 6.9-3 PDFFormField, AA_size == 0
63 ,form_no_AA =
64 %=====
65 % to be continued https://docs.verapdf.org/validation/pdfa-part1/
66 % - Outputintent/colorprofiles requirements
67 % an outputintent should be loaded. There are more requirements
68 % but these are not tested
69 ,outputintent =
70 ,outputintent_subtype = {GTS_PDFA1}
71 ,outputintent_profile = {sRGB.icc}
72 % - no Alternates key in image dictionaries
73 % - no OPI, Ref, Subtype2 with PS key in xobjects
74 % - Interpolate = false in images
75 % - no TR, TR2 in ExtGstate
76 }
77
78 %A-2b =====
79 \prop_new:c { g_pdfmeta_standard_pdf/A-2b_prop }
80 \prop_gset_eq:cc
81 { g_pdfmeta_standard_pdf/A-2b_prop }
82 { g_pdfmeta_standard_pdf/A-1b_prop }
83 \prop_gput:cnn
84 { g_pdfmeta_standard_pdf/A-2b_prop }{name}{pdf/A-2b}
85 \prop_gput:cnn
86 { g_pdfmeta_standard_pdf/A-2b_prop }{year}{2011}
87 % embedding files is allowed (with restrictions)
88 \prop_gremove:cn
89 { g_pdfmeta_standard_pdf/A-2b_prop }
90 { embed_content}
91
92 %A-3b =====
93 \prop_new:c { g_pdfmeta_standard_pdf/A-3b_prop }
94 \prop_gset_eq:cc
95 { g_pdfmeta_standard_pdf/A-3b_prop }
96 { g_pdfmeta_standard_pdf/A-2b_prop }

```

```

97 \prop_gput:cnn
98 { g_pdfmeta_standard_pdf/A-3b_prop }{name}{pdf/A-3b}
99 \prop_gput:cnn
100 { g_pdfmeta_standard_pdf/A-2b_prop }{year}{2012}
101 % embedding files is allowed (with restrictions)
102 \prop_gremove:cn
103 { g_pdfmeta_standard_pdf/A-3b_prop }
104 { embed_content}
105
106 % ["pdf/a-1b:2005"] = {
107 %         pdf_version           = 1.4,
108 %         format_name           = "pdf/a-1b:2005",
109 %         xmp_file              = "lpdf-pda.xml",
110 %         gts_flag              = "GTS_PDFA1",
111 %         gray_scale            = true,
112 %         cmyk_colors           = true,
113 %         rgb_colors            = true,
114 %         spot_colors           = true,
115 %         calibrated_rgb_colors = true, -- unknown
116 %         cielab_colors         = true, -- unknown
117 %         include_intents       = true,
118 %         forms                 = true,
119 %         internal_icc_profiles = true,
120 %         include_cidsets       = true,
121 %         include_charsets      = true,
122 %         inject_metadata       = function()
123 %             injectxmpinfo("xml://rdf:RDF", "<rdf:Description rdf:about=''"
124 %                 end
125 %     },

```

(End definition for `\g_pdfmeta_standard_pdf/A-1b_prop`, `\g_pdfmeta_standard_pdf/A-2b_prop`, and `\g_pdfmeta_standard_pdf/A-3b_prop`. These variables are documented on page [1](#).)

3.3 Colorprofiles and Outputintents

The following provides a minimum of interface to add a color profile and an outputintent need for PDF/A for now. There will be need to extend it later, so we try for enough generality.

```

126 \pdfdict_new:n {l_pdfmeta/outputintent}
127 \pdfdict_put:nnn {l_pdfmeta/outputintent}
128 {Type}{/OutputIntent}
129 \prop_const_from_keyval:cn { c__pdfmeta_colorprofile_sRGB.icc}
130 {
131     ,OutputConditionIdentifier=IEC~sRGB
132     ,Info=IEC~61966-2.1~Default~RGB~colour~space~~~sRGB
133     ,RegistryName=http://www.iec.ch
134     ,N = 3
135 }
136 \prop_const_from_keyval:cn { c__pdfmeta_colorprofile_FOGRA39L_coated.icc}
137 {
138     ,OutputConditionIdentifier=FOGRA39L-Coated
139     ,Info={Offset~printing,~according~to~ISO~12647-2:2004/Amd~1,~OFCOM,~ %
140     paper~type~1~or~2~==coated~art,~115~g/m2,~tone~value~increase~curves~A~(CMY)~and~B~(K)}
141     ,RegistryName=http://www.fogra.org

```

```

142     ,N = 4
143 }
144
145 \cs_new_protected:Npn \__pdfmeta_embed_colorprofile:n #1#1 file name
146 {
147     \pdf_object_if_exist:nF {__pdfmeta_colorprofile_#1}
148     {
149         \pdf_object_new:nn {__pdfmeta_colorprofile_#1}{fstream}
150         \pdf_object_write:nx {__pdfmeta_colorprofile_#1}
151         {
152             {/N\c_space_tl
153             \prop_item:cn{c__pdfmeta_colorprofile_#1}{N}
154             }
155             {#1}
156         }
157     }
158 }
159
160 \cs_new_protected:Npn \__pdfmeta_write_outputintent:nn #1 #2 %#1 file name, #2 subtype
161 {
162     \group_begin:
163     \pdfdict_put:nnx {l_pdfmeta/outputintent}{S}{/\str_convert_pdfname:n{#2}}
164     \pdfdict_put:nnx {l_pdfmeta/outputintent}
165     {DestOutputProfile}
166     {\pdf_object_ref:n{__pdfmeta_colorprofile_#1}}
167     \clist_map_inline:nn { OutputConditionIdentifier, Info, RegistryName }
168     {
169         \prop_get:cnNT
170         { c__pdfmeta_colorprofile_#1}
171         { ##1 }
172         \l__pdfmeta_tmpa_tl
173         {
174             \pdf_text_convert:nVN {utf8/string-print}\l__pdfmeta_tmpa_tl\l__pdfmeta_tmpa_str
175             \pdfdict_put:nnx
176             {l_pdfmeta/outputintent}{##1}{\l__pdfmeta_tmpa_str}
177         }
178     }
179     \pdf_object_now:nx {dict}{\pdfdict_use:n {l_pdfmeta/outputintent} }
180     \pdfmanagement_add:nnx {Catalog}{OutputIntents}{\pdf_object_last:}
181     \group_end:
182 }
183
184 \AddToHook{begindocument/end}
185 {
186     \prop_if_in:NnT\g_pdfmeta_standard_prop {outputintent}
187     {
188         \exp_args:Nx
189         \__pdfmeta_embed_colorprofile:n
190         {\prop_item:Nn \g_pdfmeta_standard_prop {outputintent_profile}}
191         \exp_args:Nxx
192         \__pdfmeta_write_outputintent:nn
193         {\prop_item:Nn \g_pdfmeta_standard_prop {outputintent_profile}}
194         {\prop_item:Nn \g_pdfmeta_standard_prop {outputintent_subtype}}
195     }

```

```
196 }
```

4 XMP-metadata (work in progress)

We need a command to write a BOM to the pdf:

```
197 %
198 \str_const:Nx\c__pdfmeta_bom_str
199 {
200   \bool_if:nTF
201     {
202       \sys_if_engine luatex_p: || \sys_if_engine_xetex_p:
203     }
204     {
205       \char_generate:nn {65279}{12}
206     }
207     {
208       \char_generate:nn {239}{12}
209       \char_generate:nn {187}{12}
210       \char_generate:nn {191}{12}
211     }
212   }
213 % writing the stream:
214 \str_new:N \g__pdfmeta_xmppacket_str
215
216 \str_gset:Nx \g__pdfmeta_xmppacket_str
217 {<?xpacket~begin="\c__pdfmeta_bom_str"-id="W5M0MpCehiHzreSzNTczkc9d"?>\iow_newline:}
218 \str_gput_right:Nx \g__pdfmeta_xmppacket_str
219 {<x:xmpmeta~xmlns:x="adobe:ns:meta/">\iow_newline:}
220
221 %here more contents ...
222 \str_gput_right:Nx \g__pdfmeta_xmppacket_str { </x:xmpmeta> \iow_newline: }
223 %here padding??
224 \str_gput_right:Nn \g__pdfmeta_xmppacket_str { <?xpacket~end="w"?> }
225 %
```

4.1 The stream object

This code is currently not activated.

```
226
227 %\hook_gput_code:nnn {shipout/lastpage}{pdf}
228 % {
229 %   \bool_if:NT \g_pdfmeta_xmp_bool
230 %   {
231 %     \pdf_object_new:nn {__pdf_Metadata} {stream}
232 %     \pdf_object_write:nx {__pdf_Metadata}
233 %     {
234 %       { /Type /Metadata /Subtype /XML }
235 %       {\g__pdfmeta_xmppacket_str}
236 %     }
237 %   }
238 % }
239 </package>
```

Index

The italic numbers denote the pages where the corresponding entry is described, numbers underlined point to the definition, all others indicate the places where it is used.

A	
\AddToHook	184
B	
bitset commands:	
\bitset_set_false:Nn	31, 32, 33
\bitset_set_true:Nn	30
\bitset_to_arabic:N	34, 35, 36, 37, 38
bool commands:	
\bool_if:NTF	229
\bool_if:nTF	200
C	
char commands:	
\char_generate:nn	205, 208, 209, 210
clist commands:	
\clist_map_inline:nn	167
cs commands:	
\cs_new_protected:Npn	145, 160
E	
exp commands:	
\exp_args:Nx	188
\exp_args:Nxx	191
G	
group commands:	
\group_begin:	162
\group_end:	181
H	
hook commands:	
\hook_gput_code:nnn	9, 227
I	
iow commands:	
\iow_newline:	217, 219, 222
M	
msg commands:	
\msg_new:nnn	5
P	
pdf commands:	
\pdf_object_if_exist:nTF	147
\pdf_object_last:	180
\pdf_object_new:nn	149, 231
\pdf_object_now:nn	179
\pdf_object_ref:n	166
\pdf_object_write:nn	150, 232
\pdf_text_convert:nnN	174
pdfannot commands:	
\pdfannot_dict_put:nnn	34, 35, 36, 37, 38
\l_pdfannot_F_bitset	30, 31, 32, 33, 34, 35, 36, 37, 38
pdfdict commands:	
\pdfdict_new:n	126
\pdfdict_put:nnn	127, 163, 164, 175
\pdfdict_use:n	179
pdfmanagement commands:	
\pdfmanagement_add:nnn	180
pdfmeta commands:	
\g_pdfmeta_standard_pdf/A-1b_-prop	1, 13
\g_pdfmeta_standard_pdf/A-2b_-prop	1, 13
\g_pdfmeta_standard_pdf/A-3b_-prop	1, 13
\g_pdfmeta_standard_prop	1, 8, 11, 186, 190, 193, 194
\g_pdfmeta_xmp_bool	229
pdfmeta internal commands:	
\c__pdfmeta_bom_str	198, 217
__pdfmeta_embed_colorprofile:n	145, 189
\l__pdfmeta_tmpa_str	6, 174, 176
\l__pdfmeta_tmpa_tl	6, 172, 174
__pdfmeta_write_outputintent:nn	160, 192
\g__pdfmeta_xmppacket_str	214, 216, 218, 222, 224, 235
prop commands:	
\prop_const_from_keyval:Nn	129, 136
\prop_get:NnNTF	169
\prop_gput:Nnn	83, 85, 97, 99
\prop_gremove:Nn	88, 102
\prop_gset_eq:NN	80, 94
\prop_if_in:NnTF	186
\prop_item:Nn	11, 153, 190, 193, 194
\prop_new:N	8, 13, 79, 93
\prop_set_from_keyval:Nn	14
\ProvidesExplPackage	3
S	
str commands:	
\str_const:Nn	198

<code>\str_convert_pdfname:n</code>	163	<code>\sys_if_engine_xetex_p:</code>	202
<code>\str_gput_right:Nn</code>	218, 222, 224		
<code>\str_gset:Nn</code>	216		T
<code>\str_new:N</code>	7, 214	tl commands:	
sys commands:		<code>\c_space_tl</code>	152
<code>\sys_if_engine luatex_p:</code>	202	<code>\tl_new:N</code>	6