

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 and 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.

This package should not replace both packages. Regarding XMP-metadata its goal is to create a skeleton metadata stream, add some core default values and to define interfaces that allows other packages to add data to this metadata and so to extend them.

The problems to solve here are

- which tree structure is sensible
- how to escape if needed the input (or which tools are needed to allow the users to correctly escape their input)
- how interface to input data should look

Regarding the pdf standards the goal is to collect the requirements and to set check how to set them.

2 l3pdfmeta implementation

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

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

3 XMP-metadata

```
8 % we need a command to write a BOM to the pdf:
9 \str_const:Nx\c__pdfmeta_bom_str
10 {
11   \bool_if:nTF
12     {
13       \sys_if_engine luatex_p: || \sys_if_engine_xetex_p:
14     }
15     {
16       \char_generate:nn {65279}{12}
17     }
18     {
19       \char_generate:nn {239}{12}
20       \char_generate:nn {187}{12}
21       \char_generate:nn {191}{12}
22     }
23 }
24 % writing the stream:
25 \str_new:N \g__pdfmeta_xmppacket_str
26
27 \str_gset:Nx \g__pdfmeta_xmppacket_str {<?xpacket~begin="\c__pdfmeta_bom_str"~id="W5M0M"}
28 \str_gput_right:Nx \g__pdfmeta_xmppacket_str {<x:xmpmeta~xmlns:x="adobe:ns:meta/">\iow_newline}
29
30 %here more contents ...
31 \str_gput_right:Nx \g__pdfmeta_xmppacket_str { </x:xmpmeta> \iow_newline: }
32 %here padding??
33 \str_gput_right:Nn \g__pdfmeta_xmppacket_str { <?xpacket~end="w"??> }
34 %
```

3.1 The stream object

```
35
36 \hook_gput_code:nnn {shipout/lastpage}{pdf}
37 {
38   \bool_if:NT \g_pdfmeta_xmp_bool
39   {
40     \pdf_object_new:nn {__pdf_Metadata} {stream}
41     \pdf_object_write:nx {__pdf_Metadata}
42     {
43       { /Type /Metadata /Subtype /XML }
44       {\g__pdfmeta_xmppacket_str}
45     }
46   }
47 }
```

4 Standards (work in progress)

4.1 Tools and tests

This public property will contain the settings for the document.

\g_pdfmeta_standard_prop

```
48 \prop_new:N \g_pdfmeta_standard_prop
```

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

This conditional tests if the value for `/N` in a named action is allowed (pdf/A restrict the names).

```

49 \prg_new_protected_conditional:Npnn \pdfmeta_if_named_action_allowed:n #1 { F, T , TF }
50 {
51   \prop_get:NnNTF \g_pdfmeta_standard_prop {named_actions} \l__pdfmeta_tmpa_tl
52   {
53     \tl_if_in:NnTF \l__pdfmeta_tmpa_tl { #1 }
54     {
55       \prg_return_true: %restricted but ok
56     }
57     {
58       \prg_return_false: %restricted and false
59     }
60   }
61   {
62     \prg_return_true: %not restricted
63   }
64 }

65 \hook_gput_code:nnn {begindocument} {pdf}
66 {
67   \prop_item:Nn \g_pdfmeta_standard_prop { annot_flags }
68 }

```

4.2 pdf/A

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

```

69 \prop_new:c { g_pdfmeta_standard_pdf/A-1b_prop }
70 \prop_set_from_keyval:cn { g_pdfmeta_standard_pdf/A-1b_prop }
71 {
72   ,name           = pdf/A-1b
73   ,type           = A
74   ,year           = 2005
75   ,pdf_version    = 1.4           %minimum
76   ,encryption     = false
77   ,external_content = false % no F, FFilter, or FDecodeParms in stream dicts
78   ,embed_content  = false, % no EF key in filespec, no /Type/EmbeddedFiles
79   ,max_string_size = 65535
80   ,max_array_size = 8191
81   ,max_dict_size  = 4095
82   ,max_obj_num    = 8388607
83   ,max_nest_qQ    = 28
84   ,named_actions  = {NextPage, PrevPage, FirstPage, LastPage}
85   ,annot_flags    = {
86     \bitset_set_true:Nn \l_pdfannot_F_bitset {Print}
87     \bitset_set_false:Nn \l_pdfannot_F_bitset {Hidden}
88     \bitset_set_false:Nn \l_pdfannot_F_bitset {Invisible}
89     \bitset_set_false:Nn \l_pdfannot_F_bitset {NoView}
90     \pdfannot_dict_put:nnn {link/URI}{F}{ \bitset_to_arabic:N \l_pdfann
91     \pdfannot_dict_put:nnn {link/GoTo}{F}{ \bitset_to_arabic:N \l_pdfan
92     \pdfannot_dict_put:nnn {link/GoToR}{F}{ \bitset_to_arabic:N \l_pdfa
93     \pdfannot_dict_put:nnn {link/Launch}{F}{ \bitset_to_arabic:N \l_pdfa
94     \pdfannot_dict_put:nnn {link/Named}{F}{ \bitset_to_arabic:N \l_pdfa

```

```

95         }
96     %booleans. Only the existence of the key matter.
97     %If the entry is added it means true (so in most cases "don't use ...")
98     %
99     %=====
100    % Rule 6.1.13-1 CosDocument, isOptionalContentPresent == false
101        ,Catalog_no_OCProperties =
102    %=====
103    % Rule 6.6.1-1: PDAAction, S == "GoTo" || S == "GoToR" || S == "Thread" || S == "URI" || S
104    % means: no /S/Launch, /S/Sound, /S/Movie, /S/ResetForm, /S/ImportData, /S/JavaScript, /S
105        ,annot_action_no_A =
106    %=====
107    % Rule 6.6.2-1: PDAnnot, Subtype != "Widget" || AA_size == 0
108    % means: no AA dictionary
109        ,annot_widget_no_AA =
110    %=====
111    % Rule 6.9-2: PDAnnot, Subtype != "Widget" || (A_size == 0 && AA_size == 0)
112    % (looks like a tightening of the previous rule)
113        ,annot_widget_no_A_AA =
114    %=====
115    % Rule 6.9-1 PDAcroForm, NeedAppearances == null || NeedAppearances == false
116        ,form_no_NeedAppearances =
117    %=====
118    %Rule 6.9-3 PDFFormField, AA_size == 0
119        ,form_no_AA =
120    %=====
121    % to be continued https://docs.verapdf.org/validation/pdfa-part1/
122    % - Outputintent/colorprofiles requirements
123    % an outputintent should be loaded. There are more requirements
124    % but these are not tested
125        ,outputintent =
126        ,outputintent_subtype = {GTS_PDFA1}
127        ,outputintent_profile = {sRGB.icc}
128    % - no Alternates key in image dictionaries
129    % - no OPI, Ref, Subtype2 with PS key in xobjects
130    % - Interpolate = false in images
131    % - no TR, TR2 in ExtGstate
132    }
133
134    %A-2b =====
135    \prop_new:c { g_pdfmeta_standard_pdf/A-2b_prop }
136    \prop_gset_eq:cc
137        { g_pdfmeta_standard_pdf/A-2b_prop }
138        { g_pdfmeta_standard_pdf/A-1b_prop }
139    \prop_gput:cnn
140        { g_pdfmeta_standard_pdf/A-2b_prop }{name}{pdf/A-2b}
141    \prop_gput:cnn
142        { g_pdfmeta_standard_pdf/A-2b_prop }{year}{2011}
143    % embedding files is allowed (with restrictions)
144    \prop_gremove:cn
145        { g_pdfmeta_standard_pdf/A-2b_prop }
146        { embed_content}
147
148    %A-3b =====

```

```

149 \prop_new:c { g_pdfmeta_standard_pdf/A-3b_prop }
150 \prop_gset_eq:cc
151   { g_pdfmeta_standard_pdf/A-3b_prop }
152   { g_pdfmeta_standard_pdf/A-2b_prop }
153 \prop_gput:cnn
154   { g_pdfmeta_standard_pdf/A-3b_prop }{name}{pdf/A-3b}
155 \prop_gput:cnn
156   { g_pdfmeta_standard_pdf/A-2b_prop }{year}{2012}
157 % embedding files is allowed (with restrictions)
158 \prop_gremove:cn
159   { g_pdfmeta_standard_pdf/A-3b_prop }
160   { embed_content}
161
162 % ["pdf/a-1b:2005"] = {
163 %     pdf_version           = 1.4,
164 %     format_name          = "pdf/a-1b:2005",
165 %     xmp_file              = "lpdf-pda.xml",
166 %     gts_flag              = "GTS_PDFA1",
167 %     gray_scale            = true,
168 %     cmyk_colors           = true,
169 %     rgb_colors            = true,
170 %     spot_colors           = true,
171 %     calibrated_rgb_colors = true, -- unknown
172 %     cielab_colors         = true, -- unknown
173 %     include_intents       = true,
174 %     forms                 = true,
175 %     internal_icc_profiles = true,
176 %     include_cidsets       = true,
177 %     include_charsets      = true,
178 %     inject_metadata       = function()
179 %         injectxmpinfo("xml://rdf:RDF","<rdf:Description rdf:about='' xmlns:pdfaid='
180 %     end
181 % },

```

(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 ??.)

4.3 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.

```

182 \pdfdict_new:n {l_pdfmeta/outputintent}
183 \pdfdict_put:nnn {l_pdfmeta/outputintent}
184   {Type}{/OutputIntent}
185 \prop_const_from_keyval:cn { c__pdfmeta_colorprofile_sRGB.icc}
186   {
187     ,OutputConditionIdentifier=IEC~sRGB
188     ,Info=IEC~61966-2.1~Default~RGB~colour~space~~~sRGB
189     ,RegistryName=http://www.iec.ch
190     ,N = 3
191   }
192 \prop_const_from_keyval:cn { c__pdfmeta_colorprofile_FOGRA39L_coated.icc}
193   {

```

```

194     ,OutputConditionIdentifier=FOGRA39L-Coated
195     ,Info={Offset~printing,~according~to~ISO~12647-2:2004/Amd~1,~OFCOM,~ %
196     paper~type~1~or~2~==coated~art,~115~g/m2,~tone~value~increase~curves~A~(CMY)~and~B~(K)}
197     ,RegistryName=http://www.fogra.org
198     ,N = 4
199 }
200
201 \cs_new_protected:Npn \__pdfmeta_embed_colorprofile:n #1#1 file name
202 {
203     \pdf_object_if_exist:nF {__pdfmeta_colorprofile_#1}
204     {
205         \pdf_object_new:nn {__pdfmeta_colorprofile_#1}{fstream}
206         \pdf_object_write:nx {__pdfmeta_colorprofile_#1}
207         {
208             {/N\c_space_tl
209             \prop_item:cn{c__pdfmeta_colorprofile_#1}{N}
210             }
211             {#1}
212         }
213     }
214 }
215
216 \cs_new_protected:Npn \__pdfmeta_write_outputintent:nn #1 #2 % #1 file name, #2 subtype
217 {
218     \group_begin:
219     \pdfdict_put:nnx {l_pdfmeta/outputintent}{S}{/\str_convert_pdfname:n{#2}}
220     \pdfdict_put:nnx {l_pdfmeta/outputintent}
221     {DestOutputProfile}
222     {\pdf_object_ref:n{__pdfmeta_colorprofile_#1}}
223     \clist_map_inline:nn { OutputConditionIdentifier, Info, RegistryName }
224     {
225         \prop_get:cnNT
226         { c__pdfmeta_colorprofile_#1}
227         { ##1 }
228         \l__pdfmeta_tmpa_tl
229         {
230             \pdf_text_convert:nVN {utf8/string-print}\l__pdfmeta_tmpa_tl\l__pdfmeta_tmpa_str
231             \pdfdict_put:nnx
232             {l_pdfmeta/outputintent}{##1}{\l__pdfmeta_tmpa_str}
233         }
234     }
235     \pdf_object_now:nx {dict}{\pdfdict_use:n {l_pdfmeta/outputintent} }
236     \pdfmanagement_add:nnx {Catalog}{OutputIntents}{\pdf_object_last:}
237     \group_end:
238 }
239
240 \AddToHook{begindocument/end}
241 {
242     \prop_if_in:NnT\g_pdfmeta_standard_prop {outputintent}
243     {
244         \exp_args:Nx
245         \__pdfmeta_embed_colorprofile:n
246         {\prop_item:Nn \g_pdfmeta_standard_prop {outputintent_profile}}
247         \exp_args:Nxx

```

```

248         \__pdfmeta_write_outputintent:nn
249         {\prop_item:Nn \g_pdfmeta_standard_prop {outputintent_profile}}
250         {\prop_item:Nn \g_pdfmeta_standard_prop {outputintent_subtype}}
251     }
252 }
253 \end{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		P	
\AddToHook	240	pdf commands:	
B		\pdf_object_if_exist:nTF	203
bitset commands:		\pdf_object_last:	236
\bitset_set_false:Nn	87, 88, 89	\pdf_object_new:nn	40, 205
\bitset_set_true:Nn	86	\pdf_object_now:nn	235
\bitset_to_arabic:N	90, 91, 92, 93, 94	\pdf_object_ref:n	222
bool commands:		\pdf_object_write:nn	41, 206
\bool_if:NTF	38	\pdf_text_convert:nnN	230
\bool_if:nTF	11	pdfannot commands:	
C		\pdfannot_dict_put:nnn	90, 91, 92, 93, 94
char commands:		\l_pdfannot_F_bitset	86, 87, 88, 89, 90, 91, 92, 93, 94
\char_generate:nn	16, 19, 20, 21	pdfdict commands:	
clist commands:		\pdfdict_new:n	182
\clist_map_inline:nn	223	\pdfdict_put:nnn	183, 219, 220, 231
cs commands:		\pdfdict_use:n	235
\cs_new_protected:Npn	201, 216	pdfmanagement commands:	
E		\pdfmanagement_add:nnn	236
exp commands:		pdfmeta commands:	
\exp_args:Nx	244	\pdfmeta_if_named_action_	
\exp_args:Nxx	247	allowed:n	49
G		\g_pdfmeta_standard_pdf/A-1b_	
group commands:		prop	69
\group_begin:	218	\g_pdfmeta_standard_pdf/A-2b_	
\group_end:	237	prop	69
H		\g_pdfmeta_standard_pdf/A-3b_	
hook commands:		prop	69
\hook_gput_code:nnn	36, 65	\g_pdfmeta_standard_prop	48, 51, 67, 242, 246, 249, 250
I		\g_pdfmeta_xmp_bool	38
iow commands:		pdfmeta internal commands:	
\iow_newline:	27, 28, 31	\c_pdfmeta_bom_str	9, 27
M		__pdfmeta_embed_colorprofile:n	201, 245
msg commands:		\l_pdfmeta_tmpa_str	6, 230, 232
\msg_new:nnn	5	\l_pdfmeta_tmpa_tl	6, 51, 53, 228, 230

<code>__pdfmeta_write_outputintent:n</code>	<code>\prop_set_from_keyval:Nn</code>	70
..... 216, 248	<code>\ProvidesExplPackage</code>	3
<code>\g__pdfmeta_xmppacket_str</code>		
..... 25, 27, 28, 31, 33, 44		
prg commands:	S	
<code>\prg_new_protected_conditional:Npnn</code>	<code>str</code> commands:	
..... 49	<code>\str_const:Nn</code>	9
<code>\prg_return_false:</code>	<code>\str_convert_pdfname:n</code>	219
..... 58	<code>\str_gput_right:Nn</code>	28, 31, 33
<code>\prg_return_true:</code>	<code>\str_gset:Nn</code>	27
..... 55, 62	<code>\str_new:N</code>	7, 25
prop commands:	sys commands:	
<code>\prop_const_from_keyval:Nn</code>	<code>\sys_if_engine luatex_p:</code>	13
..... 185, 192	<code>\sys_if_engine xetex_p:</code>	13
<code>\prop_get:NnNTF</code>		
..... 51, 225		
<code>\prop_gput:Nnn</code>		
..... 139, 141, 153, 155		
<code>\prop_gremove:Nn</code>		
..... 144, 158		
<code>\prop_gset_eq:NN</code>		
..... 136, 150		
<code>\prop_if_in:NnTF</code>		
..... 242		
<code>\prop_item:Nn</code>		
..... 67, 209, 246, 249, 250		
<code>\prop_new:N</code>		
..... 48, 69, 135, 149		
	T	
	tl commands:	
	<code>\c_space_tl</code>	208
	<code>\tl_if_in:NnTF</code>	53
	<code>\tl_new:N</code>	6