The l3pdfmeta package
pdf-standards and XMP-metadata

The LaTeX3 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="W5M0Mp
28  \str_gput_right:Nx \g__pdfmeta_xmppacket_str {<x:xmpmeta~xmlns:x="adobe:ns:meta/">\iow_newlin
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* **??***.)*

```
49 \hook_gput_code:nnn {begindocument} {pdf}
50   {
51     \prop_item:Nn \g_pdfmeta_standard_prop { annot_flags }
52   }
```

4.2 pdf/A

```
53 \prop_new:c { g_pdfmeta_standard_pdf/A-1b_prop }
54 \prop_set_from_keyval:cn { g_pdfmeta_standard_pdf/A-1b_prop }
55   {
56     ,name            = pdf/A-1b
57     ,type            = A
58     ,year            = 2005
59     ,pdf_version     = 1.4        %minimum
60     ,encryption      = false
61     ,external_content = false  % no F, FFilter, or FDecodeParms in stream dicts
62     ,embed_content   = false, % no EF key in filespec, no /Type/EmbeddedFiles
63     ,max_string_size = 65535
64     ,max_array_size  = 8191
65     ,max_dict_size   = 4095
66     ,max_obj_num     = 8388607
67     ,max_nest_qQ     = 28
68     ,named_actions   = {NextPage, PrevPage, FirstPage, LastPage}
69     ,annot_flags     = {
70                            \bitset_set_true:Nn  \l_pdfannot_F_bitset {Print}
71                            \bitset_set_false:Nn \l_pdfannot_F_bitset {Hidden}
72                            \bitset_set_false:Nn \l_pdfannot_F_bitset {Invisible}
73                            \bitset_set_false:Nn \l_pdfannot_F_bitset {NoView}
74                            \pdfannot_dict_put:nnn {link/URI}{F}{ \bitset_to_arabic:N \l_pdfann
75                            \pdfannot_dict_put:nnn {link/GoTo}{F}{ \bitset_to_arabic:N \l_pdfan
76                            \pdfannot_dict_put:nnn {link/GoToR}{F}{ \bitset_to_arabic:N \l_pdfa
77                            \pdfannot_dict_put:nnn {link/Launch}{F}{ \bitset_to_arabic:N \l_pdf
78                            \pdfannot_dict_put:nnn {link/Named}{F}{ \bitset_to_arabic:N \l_pdfa
79                        }
80     %booleans. Only the existence of the key matter.
81     %If the entry is added it means true (so in most cases "don't use ...")
82     %
83     %===============
84     % Rule 6.1.13-1 CosDocument, isOptionalContentPresent == false
85       ,Catalog_no_OCProperties =
86     %===============
87     % Rule 6.6.1-1: PDAction, S == "GoTo" || S == "GoToR" || S == "Thread" || S == "URI" || S
88     % means: no /S/Launch, /S/Sound, /S/Movie, /S/ResetForm, /S/ImportData, /S/JavaScript, /S
89       ,annot_action_no_A      =
90     %===============
91     % Rule 6.6.2-1: PDAnnot, Subtype != "Widget" || AA_size == 0
92     % means: no AA dictionary
93       ,annot_widget_no_AA     =
94     %===============
95     % Rule 6.9-2: PDAnnot, Subtype != "Widget" || (A_size == 0 && AA_size == 0)
96     % (looks like a tightening of the previous rule)
```

```
 97        ,annot_widget_no_A_AA    =
 98      %===============
 99      % Rule 6.9-1 PDAcroForm, NeedAppearances == null || NeedAppearances == false
100      ,form_no_NeedAppearances =
101      %===============
102      %Rule 6.9-3 PDFormField, AA_size == 0
103      ,form_no_AA              =
104      %===============
105      % to be continued https://docs.verapdf.org/validation/pdfa-part1/
106      % - Outputintent/colorprofiles requirements
107      % an outputintent should be loaded. There are more requirements
108      % but these are not tested
109      ,outputintent           =
110      ,outputintent_subtype = {GTS_PDFA1}
111      ,outputintent_profile = {sRGB.icc}
112      % - no Alternates key in image dictionaries
113      % - no OPI, Ref, Subtype2 with PS key in xobjects
114      % - Interpolate  = false in images
115      % - no TR, TR2 in ExtGstate
116    }
117
118  %A-2b ==============
119  \prop_new:c { g_pdfmeta_standard_pdf/A-2b_prop }
120  \prop_gset_eq:cc
121    { g_pdfmeta_standard_pdf/A-2b_prop }
122    { g_pdfmeta_standard_pdf/A-1b_prop }
123  \prop_gput:cnn
124    { g_pdfmeta_standard_pdf/A-2b_prop }{name}{pdf/A-2b}
125  \prop_gput:cnn
126    { g_pdfmeta_standard_pdf/A-2b_prop }{year}{2011}
127  % embedding files is allowed (with restrictions)
128  \prop_gremove:cn
129    { g_pdfmeta_standard_pdf/A-2b_prop }
130    { embed_content}
131
132  %A-3b ==============
133  \prop_new:c { g_pdfmeta_standard_pdf/A-3b_prop }
134  \prop_gset_eq:cc
135    { g_pdfmeta_standard_pdf/A-3b_prop }
136    { g_pdfmeta_standard_pdf/A-2b_prop }
137  \prop_gput:cnn
138    { g_pdfmeta_standard_pdf/A-3b_prop }{name}{pdf/A-3b}
139  \prop_gput:cnn
140    { g_pdfmeta_standard_pdf/A-2b_prop }{year}{2012}
141  % embedding files is allowed (with restrictions)
142  \prop_gremove:cn
143    { g_pdfmeta_standard_pdf/A-3b_prop }
144    { embed_content}
145
146  %    ["pdf/a-1b:2005"] = {
147  %         pdf_version           = 1.4,
148  %         format_name           = "pdf/a-1b:2005",
149  %         xmp_file              = "lpdf-pda.xml",
150  %         gts_flag              = "GTS_PDFA1",
```

```
151 %              gray_scale             = true,
152 %              cmyk_colors            = true,
153 %              rgb_colors             = true,
154 %              spot_colors            = true,
155 %              calibrated_rgb_colors  = true, -- unknown
156 %              cielab_colors          = true, -- unknown
157 %              include_intents        = true,
158 %              forms                  = true,
159 %              internal_icc_profiles  = true,
160 %              include_cidsets        = true,
161 %              include_charsets       = true,
162 %              inject_metadata        = function()
163 %                  injectxmpinfo("xml://rdf:RDF","<rdf:Description rdf:about='' xmlns:pdfaid='h
164 %              end
165 %          },
```

(*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.

```
166 \pdfdict_new:n   {l_pdfmeta/outputintent}
167 \pdfdict_put:nnn {l_pdfmeta/outputintent}
168   {Type}{/OutputIntent}
169 \prop_const_from_keyval:cn { c__pdfmeta_colorprofile_sRGB.icc}
170   {
171     ,OutputConditionIdentifier=IEC~sRGB
172     ,Info=IEC~61966-2.1~Default~RGB~colour~space~-~sRGB
173     ,RegistryName=http://www.iec.ch
174     ,N = 3
175   }
176 \prop_const_from_keyval:cn { c__pdfmeta_colorprofile_FOGRA39L_coated.icc}
177   {
178     ,OutputConditionIdentifier=FOGRA39L~Coated
179     ,Info={Offset~printing,~according~to~ISO~12647-2:2004/Amd~1,~OFCOM,~ %
180   paper~type~1~or~2~=~coated~art,~115~g/m2,~tone~value~increase~curves~A~(CMY)~and~B~(K)}
181     ,RegistryName=http://www.fogra.org
182     ,N = 4
183   }
184
185 \cs_new_protected:Npn \__pdfmeta_embed_colorprofile:n #1%#1 file name
186   {
187     \pdf_object_if_exist:nF {__pdfmeta_colorprofile_#1}
188       {
189         \pdf_object_new:nn  {__pdfmeta_colorprofile_#1}{fstream}
190         \pdf_object_write:nx {__pdfmeta_colorprofile_#1}
191           {
192             {/N\c_space_tl
193               \prop_item:cn{c__pdfmeta_colorprofile_#1}{N}
194             }
195             {#1}
```

```
196            }
197          }
198      }
199
200   \cs_new_protected:Npn \__pdfmeta_write_outputintent:nn #1 #2 %#1 file name, #2 subtype
201      {
202        \group_begin:
203        \pdfdict_put:nnx {l_pdfmeta/outputintent}{S}{/\str_convert_pdfname:n{#2}}
204        \pdfdict_put:nnx {l_pdfmeta/outputintent}
205          {DestOutputProfile}
206          {\pdf_object_ref:n{__pdfmeta_colorprofile_#1}}
207        \clist_map_inline:nn { OutputConditionIdentifier, Info, RegistryName }
208          {
209            \prop_get:cnNT
210            { c__pdfmeta_colorprofile_#1}
211            { ##1 }
212            \l__pdfmeta_tmpa_tl
213            {
214              \pdf_text_convert:nVN {utf8/string-print}\l__pdfmeta_tmpa_tl\l__pdfmeta_tmpa_str
215              \pdfdict_put:nnx
216                {l_pdfmeta/outputintent}{##1}{\l__pdfmeta_tmpa_str}
217            }
218          }
219        \pdf_object_now:nx {dict}{\pdfdict_use:n {l_pdfmeta/outputintent} }
220        \pdfmanagement_add:nnx {Catalog}{OutputIntents}{\pdf_object_last:}
221        \group_end:
222      }
223
224   \AddToHook{begindocument/end}
225      {
226        \prop_if_in:NnT\g_pdfmeta_standard_prop {outputintent}
227          {
228            \exp_args:Nx
229            \__pdfmeta_embed_colorprofile:n
230              {\prop_item:Nn \g_pdfmeta_standard_prop {outputintent_profile}}
231            \exp_args:Nxx
232            \__pdfmeta_write_outputintent:nn
233              {\prop_item:Nn \g_pdfmeta_standard_prop {outputintent_profile}}
234              {\prop_item:Nn \g_pdfmeta_standard_prop {outputintent_subtype}}
235          }
236      }
237   ⟨/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.