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.The goal is to collect the requirements and to provide code to check and fulfill them.

In future is 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

```
\tl_new:N\l__pdfmeta_tmpa_tl
\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

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

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

```
\hook_gput_code:nnn {begindocument} {pdf}
  {
    \prop_item:Nn \g_pdfmeta_standard_prop { annot_flags }
  }
```

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

```
\prop_new:c { g_pdfmeta_standard_pdf/A-1b_prop }
\prop_set_from_keyval:cn { g_pdfmeta_standard_pdf/A-1b_prop }
  {
    ,name             = pdf/A-1b
    ,type             = A
    ,year             = 2005
    ,pdf_version      = 1.4        %minimum
    ,encryption       = false
    ,external_content = false  % no F, FFilter, or FDecodeParms in stream dicts
    ,embed_content    = false, % no EF key in filespec, no /Type/EmbeddedFiles
    ,max_string_size  = 65535
    ,max_array_size   = 8191
    ,max_dict_size    = 4095
    ,max_obj_num      = 8388607
    ,max_nest_qQ      = 28
    ,named_actions    = {NextPage, PrevPage, FirstPage, LastPage}
    ,annot_flags      = {
                          \bitset_set_true:Nn  \l_pdfannot_F_bitset {Print}
                          \bitset_set_false:Nn \l_pdfannot_F_bitset {Hidden}
                          \bitset_set_false:Nn \l_pdfannot_F_bitset {Invisible}
                          \bitset_set_false:Nn \l_pdfannot_F_bitset {NoView}
                          \pdfannot_dict_put:nnn {link/URI}{F}{ \bitset_to_arabic:N \l_pdfann
                          \pdfannot_dict_put:nnn {link/GoTo}{F}{ \bitset_to_arabic:N \l_pdfan
                          \pdfannot_dict_put:nnn {link/GoToR}{F}{ \bitset_to_arabic:N \l_pdfa
                          \pdfannot_dict_put:nnn {link/Launch}{F}{ \bitset_to_arabic:N \l_pdf
                          \pdfannot_dict_put:nnn {link/Named}{F}{ \bitset_to_arabic:N \l_pdfa
                        }
    %booleans. Only the existence of the key matter.
    %If the entry is added it means true (so in most cases "don't use ...")
    %
```

```
    %===============
    % Rule 6.1.13-1 CosDocument, isOptionalContentPresent == false
      ,Catalog_no_OCProperties =
    %===============
    % Rule 6.6.1-1: PDAction, S == "GoTo" || S == "GoToR" || S == "Thread" || S == "URI" || S
    % means: no /S/Launch, /S/Sound, /S/Movie, /S/ResetForm, /S/ImportData, /S/JavaScript, /S
      ,annot_action_no_A        =
    %===============
    % Rule 6.6.2-1: PDAnnot, Subtype != "Widget" || AA_size == 0
    % means: no AA dictionary
      ,annot_widget_no_AA      =
    %===============
    % Rule 6.9-2: PDAnnot, Subtype != "Widget" || (A_size == 0 && AA_size == 0)
    % (looks like a tightening of the previous rule)
      ,annot_widget_no_A_AA    =
    %===============
    % Rule 6.9-1 PDAcroForm, NeedAppearances == null || NeedAppearances == false
    ,form_no_NeedAppearances =
    %===============
    %Rule 6.9-3 PDFormField, AA_size == 0
    ,form_no_AA              =
    %===============
    % to be continued https://docs.verapdf.org/validation/pdfa-part1/
    % - Outputintent/colorprofiles requirements
    % an outputintent should be loaded. There are more requirements
    % but these are not tested
    ,outputintent         =
    ,outputintent_subtype = {GTS_PDFA1}
    ,outputintent_profile = {sRGB.icc}
    % - no Alternates key in image dictionaries
    % - no OPI, Ref, Subtype2 with PS key in xobjects
    % - Interpolate  = false in images
    % - no TR, TR2 in ExtGstate
  }

%A-2b ==============
\prop_new:c { g_pdfmeta_standard_pdf/A-2b_prop }
\prop_gset_eq:cc
  { g_pdfmeta_standard_pdf/A-2b_prop }
  { g_pdfmeta_standard_pdf/A-1b_prop }
\prop_gput:cnn
  { g_pdfmeta_standard_pdf/A-2b_prop }{name}{pdf/A-2b}
\prop_gput:cnn
  { g_pdfmeta_standard_pdf/A-2b_prop }{year}{2011}
% embedding files is allowed (with restrictions)
\prop_gremove:cn
  { g_pdfmeta_standard_pdf/A-2b_prop }
  { embed_content}

%A-3b ==============
\prop_new:c { g_pdfmeta_standard_pdf/A-3b_prop }
\prop_gset_eq:cc
  { g_pdfmeta_standard_pdf/A-3b_prop }
  { g_pdfmeta_standard_pdf/A-2b_prop }
```

```
\prop_gput:cnn
  { g_pdfmeta_standard_pdf/A-3b_prop }{name}{pdf/A-3b}
\prop_gput:cnn
  { g_pdfmeta_standard_pdf/A-2b_prop }{year}{2012}
% embedding files is allowed (with restrictions)
\prop_gremove:cn
  { g_pdfmeta_standard_pdf/A-3b_prop }
  { embed_content}

%   ["pdf/a-1b:2005"] = {
%            pdf_version             = 1.4,
%            format_name             = "pdf/a-1b:2005",
%            xmp_file                = "lpdf-pda.xml",
%            gts_flag                = "GTS_PDFA1",
%            gray_scale              = true,
%            cmyk_colors             = true,
%            rgb_colors              = true,
%            spot_colors             = true,
%            calibrated_rgb_colors   = true, -- unknown
%            cielab_colors           = true, -- unknown
%            include_intents         = true,
%            forms                   = true,
%            internal_icc_profiles   = true,
%            include_cidsets         = true,
%            include_charsets        = true,
%            inject_metadata         = function()
%                injectxmpinfo("xml://rdf:RDF","<rdf:Description rdf:about='' xmlns:pdfaid='h
%            end
%        },
```

(*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.

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

```
    ,N = 4
  }

\cs_new_protected:Npn \__pdfmeta_embed_colorprofile:n #1%#1 file name
  {
    \pdf_object_if_exist:nF {__pdfmeta_colorprofile_#1}
      {
        \pdf_object_new:nn  {__pdfmeta_colorprofile_#1}{fstream}
        \pdf_object_write:nx {__pdfmeta_colorprofile_#1}
         {
           {/N\c_space_tl
             \prop_item:cn{c__pdfmeta_colorprofile_#1}{N}
           }
           {#1}
         }
      }
  }

\cs_new_protected:Npn \__pdfmeta_write_outputintent:nn #1 #2 %#1 file name, #2 subtype
  {
    \group_begin:
     \pdfdict_put:nnx {l_pdfmeta/outputintent}{S}{/\str_convert_pdfname:n{#2}}
     \pdfdict_put:nnx {l_pdfmeta/outputintent}
       {DestOutputProfile}
       {\pdf_object_ref:n{__pdfmeta_colorprofile_#1}}
     \clist_map_inline:nn { OutputConditionIdentifier, Info, RegistryName }
       {
         \prop_get:cnNT
          { c__pdfmeta_colorprofile_#1}
          { ##1 }
          \l__pdfmeta_tmpa_tl
          {
            \pdf_string_from_unicode:nVN {utf8/string}\l__pdfmeta_tmpa_tl\l__pdfmeta_tmpa_str
            \pdfdict_put:nnx
              {l_pdfmeta/outputintent}{##1}{\l__pdfmeta_tmpa_str}
          }
       }
     \pdf_object_now:nx {dict}{\pdfdict_use:n {l_pdfmeta/outputintent} }
     \pdfmanagement_add:nnx {Catalog}{OutputIntents}{\pdf_object_last:}
    \group_end:
  }

\AddToHook{begindocument/end}
  {
     \prop_if_in:NnT\g_pdfmeta_standard_prop {outputintent}
      {
        \exp_args:Nx
        \__pdfmeta_embed_colorprofile:n
          {\prop_item:Nn \g_pdfmeta_standard_prop {outputintent_profile}}
        \exp_args:Nxx
        \__pdfmeta_write_outputintent:nn
          {\prop_item:Nn \g_pdfmeta_standard_prop {outputintent_profile}}
          {\prop_item:Nn \g_pdfmeta_standard_prop {outputintent_subtype}}
      }
```

```
  }
```

4 XMP-metadata (work in progress)

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

```
%
\str_const:Nx\c__pdfmeta_bom_str
 {
  \bool_if:nTF
    {
      \sys_if_engine_luatex_p: || \sys_if_engine_xetex_p:
    }
    {
     \char_generate:nn {65279}{12}
    }
    {
      \char_generate:nn {239}{12}
      \char_generate:nn {187}{12}
      \char_generate:nn {191}{12}
    }
 }
% writing the stream:
\str_new:N \g__pdfmeta_xmppacket_str

\str_gset:Nx      \g__pdfmeta_xmppacket_str
 {<?xpacket~begin="\c__pdfmeta_bom_str"~id="W5M0MpCehiHzreSzNTczkc9d"?>\iow_newline:}
\str_gput_right:Nx \g__pdfmeta_xmppacket_str
 {<x:xmpmeta~xmlns:x="adobe:ns:meta/">\iow_newline:}

%here more contents ...
\str_gput_right:Nx \g__pdfmeta_xmppacket_str { </x:xmpmeta> \iow_newline: }
%here padding??
\str_gput_right:Nn \g__pdfmeta_xmppacket_str { <?xpacket~end="w"?> }
%
```

4.1 The stream object

This code is currently not activated.

```

%\hook_gput_code:nnn {shipout/lastpage}{pdf}
%  {
%    \bool_if:NT \g_pdfmeta_xmp_bool
%      {
%        \pdf_object_new:nn {__pdf_Metadata} {stream}
%        \pdf_object_write:nx {__pdf_Metadata}
%         {
%           { /Type /Metadata /Subtype /XML }
%           {\g__pdfmeta_xmppacket_str}
%         }
%      }
%  }
```

⟨/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_string_from_unicode: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

\str_convert_pdfname:n 163
\str_gput_right:Nn 218, 222, 224
\str_gset:Nn 216
\str_new:N 7, 214
sys commands:
\sys_if_engine_luatex_p: 202
\sys_if_engine_xetex_p: 202

T

tl commands:
\c_space_tl 152
\tl_new:N 6