The `documentmetadata-support` code[*]

Frank Mittelbach, Ulrike Fischer, LaTeX Project

May 25, 2023

1 Introduction

The kernel command `\DocumentMetadata`, which can be used as the very first declaration in a document (i.e., before `\documentclass`), defines metadata and other configuration data that applies to the document as a whole (using a key/value syntax).

While the underlying functionality is still under development (e.g., further keys will be added over time and keys marked temporary may vanish again) the code for `\DocumentMetadata` is placed in a separate bundle, so that it is easier to update it without the need to build a full LaTeX release. Over time the functionality will move fully into the kernel.

`\DocumentMetadata` also loads and activates the new PDF management code from `pdfmanagement-testphase`. As this forces the loading of the l3backend files, a backend that can't be detected automatically like `dvipdfmx`, must be set in the first `\DocumentMetadata`.

From a process perspective `\DocumentMetadata` loads the `pdfmanagement-testphase` code the first time it is called and then redefines itself to only manage key/value pairs in case it is called more than once. In particular, this means that a document without a `\DocumentMetadata` declaration has no code available for extended management of PDF output as needed for various features developed as part of the multi-year "Tagged PDF" project [1].

2 Currently supported key/values

Currently the following keys are implemented for `\DocumentMetadata`:

backend Passes the backend name to expl3. This is needed only if the needed backend can't be automatically determined or if the workflow used requires a special backend.

pdfversion Sets the PDF version explicitly, e.g., `pdfversion=1.7`

uncompress (no value) Forces an uncompressed pdf — mainly for debugging purposes.

lang Explicitly sets the Lang entry in the Catalog, e.g., `lang=de-DE`. If not given the default value used is `en-US`.

---

[*]This file has version 1.0c dated 2023-01-07, © LaTeX Project.

pdfstandard Choice key to set the pdf standard. Currently `A-1b`, `A-2a`, `A-2b`, `A-2u`, `A-3a`, `A-3b`, `A-3u` and `A-4` are accepted as values. The casing is irrelevant, `a-1b` works too. Note that using these key doesn't mean that the document actually follows the standard. LaTeX can neither ensure nor check all requirements of a standard, and not everything it can do theoretically has already been implemented. When setting an A-standard a color profile is included and the `/OutputIntent` is set and javascript action in hyperref are suppressed. The `u` variants do not enforce unicode, but they will pass the information to hyperref. The `a` variants do *not* enforce (or even test) a tagged pdf yet.

Starting with version 0.95s of `pdfmanagement-testphase` it is also possible to use the values `X-4`, `X-4p`, `X-5g`, `X-5n`, `X-5pg`, `X-6`, `X-6n`, `X-6p`, `UA-1` for a PDF/X and PDF/UA standard. These keys currently set *only* the relevant XMP-metadata.

`pdfstandard` can be used more than once to set overlapping standards, e.g:
`pdfstandard=A-2b,pdfstandard=X-4,pdfstandard=UA-1`

If XMP-metadata are added (see the following key `xmp` the needed conformance marker for the standards are set.

More information can be found in the documentation of l3pdfmeta.

xmp A boolean, if set to false no XMP metadata are added to the PDF. The initial value is true. Details are described in the documentation of l3pdfmeta.

colorprofiles This allows to load icc-colorprofiles. Details are described in the documentation of l3pdfmeta.

testphase This key is used to load testphase code. The values it accepts and their effect will change over time, when testphase packages are added or removed or when the code is moved into the kernel.

  phase-I This value loads code implementing the first phase of the project [1], i.e., it will load the tagpdf package. It will also activate tagging by issuing `\tagpdfsetup{activate,interwordspace}`. This phase is frozen.

  phase-II This is the current development phase. Currently it differs from `phase-I` only in one point: It will additionally activate tagging of paragraphs with `\tagpdfsetup{paratagging}`. In the upcoming months it will also enable automatic tagging of other basic document elements.

  new-or-1 This patches a few commands related to the output routine. The patches are needed for the tagging of paragraphs, for the tagging of header and footer and to allow the PDF management to insert code which avoids that links happening at page breaks spills into the header and footer. This code is automatically loaded if the `testphase` values `phase-I`, `phase-II` or `new-or` are used.

  new-or This loads more changes to the output routine required for the tagging. It is not compatible with every class! The code is also loaded by the `phase-II` value.

  The `testphase` key can only be used in the first `\DocumentMetadata`.

debug This key activates some debug options. It takes a list of key-values as value. Currently the following keys are known:

para with the default and only value `show`. It will activate the `paratagging-show` option of `tagpdf`,

log with the values as described in the documentation `tagpdf`,

uncompress which does the same as `uncompress` as main key

pdfmanagement a boolean which allows to deactivate the pdfmanagement.

firstaidoff This accepts a comma lists of keywords and disables the patches related to them. More information can be found in the documentation of pdfmanagement-firstaid.

References

[1] Frank Mittelbach and Chris Rowley: *LATEX Tagged PDF — A blueprint for a large project.* https://latex-project.org/publications/indexbyyear/2020/

3 The Implementation

```
1  ⟨@@=pdfmanagement⟩
2  ⟨∗code⟩

3
4  \RequirePackage{pdfmanagement-testphase}

5  \ExplSyntaxOn\makeatletter
```

\DocumentMetadata \DocumentMetadata should not be used after \documentclass so we error in this case. It can be used more than once but follow-up calls should not do the initialization code.

```
6  \cs_set_protected:Npn \DocumentMetadata #1
7    {
8      \cs_if_eq:NNTF \documentclass \@twoclasseserror
9        { \msg_error:nn { meta } { after-class } }
10       {
```

The wanted backend must be detected first, we read the init key and then force the loading of the backend. The backend can contain management commands, so the boolean should be set to true first.

```
11         \bool_gset_true:N \g__pdfmanagement_active_bool
12         \keys_set_groups:nnn { document / metadata} {init}{ #1 }
13         %if no backend has been loaded force it now:
14         \str_if_exist:NF \c_sys_backend_str
15           {
16             \sys_load_backend:n {}
17           }
```

Now we load the extra backend code:

```
18         \ExplSyntaxOn\makeatletter
19           \file_input:n {l3backend-testphase-\c_sys_backend_str.def}
20         \ExplSyntaxOff\makeatother
```

Set the default language (this requires that the backend has been loaded), process the rest of the keys, and setup the generic driver.

```
21          \keys_set_filter:nnn  { document / metadata } { init } { lang=en-US, #1 }
22          \bool_if:NT \g__pdfmanagement_active_bool
23            {
24                \PassOptionsToPackage{customdriver=hgeneric-testphase}{hyperref}
25            }
```

`\pdfmanagement_add:nnn` has collected values in this hook.

```
26          \hook_use_once:n {pdfmanagement/add}
```

Now we redefine `\DocumentMetadata` so that it only process the keys on any further calls.
    We need to update the hyperref option if the active status changes.

```
27          \cs_set_protected:Npn \DocumentMetadata ##1
28            {
29              \keys_set_filter:nnn  { document / metadata } { init } { ##1 }
30              \str_remove_all:cn {opt@hyperref.sty}{customdriver=hgeneric-testphase}
31              \bool_if:NT \g__pdfmanagement_active_bool
32               {
33                  \PassOptionsToPackage{customdriver=hgeneric-testphase}{hyperref}
34               }
35            }
```

Load more modules, the testphase code and the firstaid code. The code is only loaded in the first `\DocumentMetadata` call!

```
36          \g__pdfmanagement_testphase_tl
37          \RequirePackage{pdfmanagement-firstaid}
38        }
39    }
```

(*End definition for* `\DocumentMetadata`. *This function is documented on page* **??**.)

```
40 %FMi defined elsewhere
41 %FMi
42 %FMi \clist_new:N \g__pdfmanagement_firstaidoff_clist
43 %FMi \tl_new:N \g__pdfmanagement_testphase_tl
44 % UFi should the definition move to here?
45 \keys_define:nn { document / metadata }
46   {
47     backend .choices:nn =
48       { dvipdfmx , dvips , dvisvgm , luatex , pdftex , pdfmode , xdvipdfmx , xetex }
49       {
50         \sys_load_backend:n {#1}
51       },
52     backend .groups:n = { init } ,
53   }
54
55 \keys_define:nn { document / metadata }
56   {
57     ,pdfversion .code:n =
58       {
59         \pdf_version_gset:n { #1 }
60         \AddToDocumentProperties[document]{pdfversion}{#1}
61       }
62     ,uncompress .code:n =
63       {
```

4

```
64            \pdf_uncompress:
65          }
66      ,uncompress .value_forbidden:n = true
67      ,lang .code:n =
68          {
69            \pdfmanagement_add:nnn {Catalog} {Lang}{(#1)}
70            \AddToDocumentProperties[document]{lang}{#1}
71          }
72    %,xmpmeta .bool_gset:N = \g_pdfmeta_xmp_bool %see pdfmeta unused and undefined for now!
73    % this uses internal command from pdfmeta, it should probably move there ...
74      ,pdfstandard .code:n =
75          {
76            \exp_args:Nnx
77            \keys_set:nn {document / metadata} {_pdfstandard=\str_uppercase:n{#1}}
78          }
79      ,_pdfstandard .choices:nn =
80        {A-1B,A-2A,A-2B,A-2U,A-3A,A-3B,A-3U,A-4}
81          {
82            \prop_if_exist:cT { g__pdfmeta_standard_pdf/#1_prop }
83              {
84                \prop_gset_eq:Nc \g__pdfmeta_standard_prop { g__pdfmeta_standard_pdf/#1 _prop }
85              }
86            \AddToDocumentProperties [document]{pdfstandard}{#1}
87          }
88      ,_pdfstandard / unknown .code:n =
89          {
90            \msg_warning:nnn{pdf}{unknown-standard}{#1}
91          }
92      ,testphase .multichoice:
93      ,testphase / tagpdf .code:n =
94          {
95            \tl_gput_right:Nn\g__pdfmanagement_testphase_tl
96              {
97                \file_if_exist_input:nF {tagpdf-latex-lab-testphase.ltx}
98                  {
99                    \RequirePackage{tagpdf}
100                   \AddToDocumentProperties [document]{testphase/tagpdf}{loaded}
101                   \tagpdfsetup{activate,paratagging,interwordspace}
102                   \AddToDocumentProperties [document]{tagging}{active}
103                   \AddToDocumentProperties [document]{tagging/para}{active}
104                   \AddToDocumentProperties [document]{tagging/interwordspace}{active}
105                 }
106             }
107         }
108     ,testphase / unknown .code:n =
109         {
110           \tl_gput_right:Nn\g__pdfmanagement_testphase_tl
111             {
112               \AddToDocumentProperties [document]{testphase/#1}{loaded}
113               \file_if_exist_input:nF {#1-latex-lab-testphase.ltx}
114                 {
115                   \msg_warning:nnn{meta}{latex-lab-pkg-missing}{#1}
116                   \AddToDocumentProperties [document]{testphase/#1}{missing}
117                 }
```

```
118                   }
119               }
120        ,activate .multichoice:
121        ,activate / tagging .code:n =
122           {
123             \PackageWarning{pdfmanagement-testphase}
124              {The~activate~key~is~deprecated.\MessageBreak
125              Tagging~is~activated~with~'testphase=tagpdf'~directly}{}
126           }
127        ,debug .code:n =
128           {
129             \keys_set:nn { document / metadata / debug } {#1}
130           }
131        ,debug / para .code:n =
132           {
133             \AddToHook
134               {
135                  package/tagpdf/after
136               }
137               {
138                  \tagpdfsetup{paratagging-show}
139               }
140           }
141        ,debug / log .code:n =
142           {
143             \AddToHook
144               {
145                  package/tagpdf/after
146               }
147               {
148                  \tagpdfsetup{log=#1}
149               }
150           }
151        ,debug / uncompress .code:n =
152           {
153             \pdf_uncompress:
154           }
155        ,debug / pdfmanagement .bool_gset:N = \g__pdfmanagement_active_bool
156        ,debug / firstaidoff .clist_gset:N = \g__pdfmanagement_firstaidoff_clist
157     }
158
159
```

3.1 Messages

```
160  %UFi is meta the right module name here?
161  \prop_gput:Nnn \g_msg_module_type_prop { meta } { LaTeX }
162  \prop_gput:Nnn \g_msg_module_name_prop { meta } { DocumentMetadata }
163
164  \msg_new:nnn { meta } { after-class }
165          {
166              \token_to_str:N \DocumentMetadata \c_space_tl
167              should~be~used~only~before~\token_to_str:N\documentclass
168          }
```

```
169 \msg_new:nnn { meta } { latex-lab-pkg-missing }
170         {
171            LaTeX-lab~package~'#1'~not~found.
172         }
173 \ExplSyntaxOff\makeatother

174 ⟨/code⟩
```

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.