The **pdfprivacy** package*

Laurens Sion
laurens@sion.info

December 3, 2017

Abstract

Creating pdfs with pdfLATEX populates several pdf meta-data fields such as date/time of creation/modification, information about the latex installation (e.g., pdfTEX version), and the relative paths of included pdfs. The `pdfprivacy` package provides support for emptying several of these pdf meta-data fields as well as suppressing some pdfTEX meta-data entries in the resulting pdf.

Contents

---

*This document corresponds to **pdfprivacy** v1.0, dated 2017/12/03.

1 Introduction

By default, pdfs created by pdfLaTeX contain several meta-data fields populated with information about the application used to create them, date/time information, and information about the LaTeX installation. The `pdfprivacy` package provides support to remove several of these meta-data fields from the resulting pdf. By default `pdfprivacy` prevents the following properties from being set:

- The PDF creator (`Creator` in `hyperref`)
- The PDF producer (`Producer` in `hyperref`)
- The PDF modification date (`ModDate` in `hyperref`)
- The PDF creation date (`CreationDate` in `hyperref`)
- The PTEX.Fullbanner in the advanced PDF properties
- The PTEX.FileName with the filenames of any included pdf (e.g., images)
- The PTEX.PageNumber
- The PTEX.InfoDict

In addition to that, `pdfprivacy` can also remove the document meta-data: `Author`, `Title`, `Subject`, and `Keywords`; and the `pdftrailerid` which allows you to create reproducible pdfs.

2 Requirements

`Pdfprivacy` requires a recent version of pdfTeX (version 1.40.17 or above[1]) if you want to suppress the inclusion of the pdfTeX banner.

       PTEX.Fullbanner This is pdfTeX, Version 3.14159265...

3 Usage

To use `pdfprivacy`, simple include it:

       \usepackage[⟨*options*⟩]{pdfprivacy},

Options are passed on to `pdfprivacy` as key-value pairs. By default, the options are set as follows:

       nodocdata=false,
       noeditdata=true,
       noproducerdata=true,
       noptexdata=true,
       nopdftrailerid=false

This removes/suppresses all pdf creator, producer, dates, and pdfTeX meta-data. Standard document properties such as title, author, subject, and keywords are still kept. The `pdftrailerid` is kept as well by default.

---

[1]`https://www.tug.org/texlive/Contents/live/texmf-dist/doc/pdftex/NEWS`

3.1 Options

All `pdfprivacy` options are boolean. Each option is prefixed with `no` to indicate you want the meta-data removed, you can specify the opposite as well (e.g., both `nodocdata=true` or `docdata=false` are valid options).

nodocdata=⟨*boolean*⟩ Specify whether document meta-data should be removed from the pdf. Document meta-data includes: (i) Title, (ii) Subject, (iii) Author, and (iv) Keywords.

noproducerdata=⟨*boolean*⟩ Specify whether meta-data on the pdf-producing application should be removed. This includes: (i) Creator, and (ii) Producer.

noeditdata=⟨*boolean*⟩ Specify whether meta-data on the edit and creation dates should be removed. This includes: (i) creation date, and (ii) modification date.

noptexdata=⟨*boolean*⟩ Specify whether the PTEX meta-data entries should be suppressed. This includes: (i) PTEX.Fullbanner, (ii) PTEX.FileName, (iii) PTEX.PageNumber, and (iv) PTEX.InfoDict.

nopdftrailerid=⟨*boolean*⟩ Specify whether the pdfrailerid should be omitted.

3.2 Using pdfprivacy without hyperref

When using `pdfprivacy` without `hyperref`, the `\pdfinfo` command is used to set the pdf meta-data properties (i.e., `Title`, `Subject`, `Keywords`, `Author`, `Creator`, and `Producer`). The `\pdfinfo` works by appending the provided properties to the pdf info dictionary. This means that any previously set properties will still be present in the resulting pdf, despite pdf viewers only showing the last one. To make sure these properties are not present, *do not* manually set any of these properties, or use `hyperref` to set them. `Hyperref` will only save the last property (i.e., the empty one set by `pdfprivacy`) to the pdf. To make absolutely sure that no meta-data is present in the resulting pdf, you can always manually check the pdf by opening it as a text file.[2]

3.3 Removing the pdftrailerid

The pdf trailer ID is a generated ID to uniquely identify a pdf. The generated ID is based on the time of creation. This ID can be removed to make reproducible pdfs, i.e., with the same cryptographic hash.

Important: The pdf trailer id is kept by default because it is optional but strongly recommended by the pdf standard.[3] Not including this entry could break some workflows that rely on the trailer ID to uniquely identify files. If this is not a concern, the `nopdftrailerid=true` option can be used to remove it as well.

---

[2]Note: privacy-sensitive content may be present in the pdf in an encoded form. When manually verifying, make sure to look for the pdf property (e.g., `/Author` (⟨*content*⟩) as well to verify it is empty.

[3]`https://www.adobe.com/content/dam/acom/en/devnet/acrobat/pdfs/pdf_reference_1-7.pdf#G8.1639448`

4 Implementation

4.1 Loading packages

Load `ifthen` for ifthenelse and booleans for managing user preferences.

```
1 \RequirePackage{ifthen}
```

Load `kvoptions` for laoding key-value options.

```
2 \RequirePackage{kvoptions}
```

4.2 Storage for options and settings

Some booleans for keeping track if hyperref is loaded and if all meta-data should be removed.

```
3 \newboolean{pdfprivacy@hyperrefloaded}
4 \newboolean{pdfprivacy@removeall}
5 \setboolean{pdfprivacy@removeall}{false}
```

4.3 Declaring package options

Declare all the package options and their complementary ones.

```
 6 \DeclareBoolOption[false]{nodocdata}
 7 \DeclareComplementaryOption{docdata}{nodocdata}
 8 \DeclareBoolOption[true]{noproducerdata}
 9 \DeclareComplementaryOption{producerdata}{noproducerdata}
10 \DeclareBoolOption[true]{noeditdata}
11 \DeclareComplementaryOption{editdata}{noeditdata}
12 \DeclareBoolOption[true]{noptexdata}
13 \DeclareComplementaryOption{ptexdata}{noptexdata}
14 \DeclareBoolOption[false]{nopdftrailerid}
15 \DeclareComplementaryOption{pdftrailerid}{nopdftrailerid}
```

Additional `all` option to just remove everything.

```
16 \DeclareVoidOption{all}{\setboolean{pdfprivacy@removeall}{true}}
```

Process the options.

```
17 \ProcessKeyvalOptions*
```

4.4 Process the user preferences

Check if we need to remove `all` and override the other preferences.

```
18 \ifthenelse{\boolean{pdfprivacy@removeall}}{%
19 \setboolean{pdfprivacy@nodocdata}{true}
20 \setboolean{pdfprivacy@noproducerdata}{true}
```

```
21 \setboolean{pdfprivacy@noeditdata}{true}
22 \setboolean{pdfprivacy@noptexdata}{true}
23 \setboolean{pdfprivacy@nopdftrailerid}{true}
24 }{%
25 }
```

Check if hyperref is loaded as some meta-data needs to be set using hyperref.

```
26 \AtBeginDocument{
27 \@ifpackageloaded{hyperref}{\setboolean{pdfprivacy@hyperrefloaded}{true}}{\setboolean{pdfpriva
28 }
```

4.5   Remove or suppress all the necessary meta-data

Process docdata preference.

```
29 \AtBeginDocument{
30 \ifthenelse{\boolean{pdfprivacy@nodocdata}}{%
31 \ifthenelse{\boolean{pdfprivacy@hyperrefloaded}}{%
32 \hypersetup{pdfinfo={ Author={}, Subject={}, Title={}, Keywords={} }}%
33 }{%
34 \pdfinfo{/Author () /Title () /Subject () /Keywords () }%
35 }
36 }{%
37 }
38 }
```

Process producerdata preference.

```
39 \AtBeginDocument{
40 \ifthenelse{\boolean{pdfprivacy@noproducerdata}}{%
41 \ifthenelse{\boolean{pdfprivacy@hyperrefloaded}}{%
42 \hypersetup{pdfinfo={ Creator={}, Producer={} }}%
43 }{%
44 \pdfinfo{/Creator () /Producer ()}%
45 }
46 }{%
47 }
48 }
```

Process editdata preference.

```
49 \ifthenelse{\boolean{pdfprivacy@noeditdata}}{%
50 \pdfinfoomitdate1
51 }{}
```

Process the PTEX data preference.

```
52 \ifthenelse{\boolean{pdfprivacy@noptexdata}}{
53 \pdfsuppressptexinfo-1
54 }{}
55
```

Process the pdftrailerid preference.

```
56 \ifthenelse{\boolean{pdfprivacy@nopdftrailerid}}{
57 \pdftrailerid{}
58 }{}
59
```

Index

Numbers written in italic refer to the page where the corresponding entry is described; numbers underlined refer to the code line of the definition; numbers in roman refer to the code lines where the entry is used.

Change History