

Quoi de neuf ?

Changing PDF Titles With pdftk

👤 sejh 📁 Computer, English, Linux, Open Source ⌚ 2014/11/26 2014/11/26 ⌵ 2 Minutes

Have you ever noticed that many PDF converters, or scanning programs, create PDF titles that are anything but meaningful? I'm using XSane for scanning, and all the PDF files get this title:

XSane scanned image

That's totally meaningless. Others create PDF from Microsoft Word, and many of those PDF files have *Microsoft Word* in the title — in most cases, the title even begins with Microsoft Word, which makes it hard to identify the document you're looking for in the window bar when you have several of them open.

With [pdftk \(PDF Toolkit\)](http://en.wikipedia.org/wiki/Pdftk) (<http://en.wikipedia.org/wiki/Pdftk>), you can fix this easily. I've used it only on Linux, but apparently it's available for other major platforms, too. Be warned that this is a command-line program.

So here's what I do to change a PDF title.

Here's a typical „Microsoft Word“ PDF file:

```
Atlas~/private/stefan> l neugier_handout.pdf
-rw-r--r-- 1 stefan users 531831 25. Okt 2012  neugier_handout.pdf
```

1. First step is to dump the PDF metadata to a file which I call report.txt:

```
Atlas~/private/stefan> pdftk neugier_handout.pdf dump_data output report.t
```

Here's what's in the PDF metadata:

```
Atlas~/private/stefan> cat report.txt
InfoBegin
InfoKey: ModDate
InfoValue: D:20081229161229+01'00'
InfoBegin
InfoKey: CreationDate
InfoValue: D:20081229161229+01'00'
InfoBegin
InfoKey: Author
InfoValue: Charakterstärke
InfoBegin
InfoKey: Title
InfoValue: Microsoft Word - Neugier_Handout.doc
InfoBegin
InfoKey: Creator
InfoValue: Word
InfoBegin
InfoKey: Producer
InfoValue: Mac OS X 10.4.11 Quartz PDFContext
PdfID0: 911d0c6f06613f3690fa270fad39d33b
PdfID1: 911d0c6f06613f3690fa270fad39d33b
NumberOfPages: 4
```

2. Second step is to edit the metadata file:

```
Atlas~/private/stefan> vi report.txt
```

Here's what I've changed. Note that I've used ASCII-7, because pdftk doesn't seem to be able to properly handle UTF-8 and friends:

```
Atlas~/private/stefan> grep Neugier report.txt
InfoValue: Neugier - Staerkentraining
```

3. Third step is to update the metadata in the PDF file. Note that the output must be written to another file — pdftk refuses to overwrite the original file:

```
/private/stefan> pdftk neugier_handout.pdf update_info report.txt output neug:
```

4. Last step is to make the copied PDF file the original PDF file:

```
Atlas~/private/stefan> mv neugier_handout.pdf.copy neugier_handout.pdf
```

And done. Verify that the title meets your expectations in the PDF viewer of your choice:

```
Atlas~/private/stefan> okular neugier_handout.pdf
```

The steps are easily scriptable if you're so inclined.

Verschlagnwortet mit:

PDF,
pdftk

Veröffentlicht von sejh



See About. Siehe About. [Alle Beiträge von sejh anzeigen](#)

[Bloggen auf WordPress.com.](#)