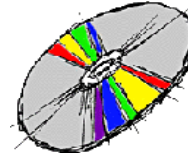


Die MagDriva-DVD

zusammengestellt von Wobo und Magnus



Dies geschieht nun z. B. mit dem Befehl:

```
pdftk MagDriva_2008.1.pdf cat 5-6
output Webperlen_2008.1.pdf
```

Vorgeschichte

Ich habe Anfang Oktober 2008 mal wieder einen Artikel gesucht und mich dafür dann durch diverse Ausdrücke und PDF-Dateien gewühlt. Alles sehr ärgerlich! Also ist bei mir die Idee gewachsen, einen Artikel-Index zu basteln. Alles eigentlich kein Problem, nur eine Fleißarbeit, genau das Richtige für dunkle Dezembertage. Nun habe ich aber noch eine weiterführende Idee. Was nützt mir die Adresse eines Artikels, wenn ich dann doch wieder den passende Artikel herausuchen muss.

Die Idee: Eine Web-Oberfläche, in der sowohl die Dokumente als auch ein Index eingehängt werden. Das alles in einem Verzeichnis, so dass man das ganze Paket auch auf eine DVD packen könnte.

Das Problem: Ich hab nur keine Ahnung, wie man das erstellt.

So habe ich erstmal das bestehende Artikelverzeichnis - hatte ich auch im ersten Eifer übersehen - auf den aktuellen Stand gebracht.

Das Programm

Als erstes musste das passende Programm her, um die im PDF-Format vorliegenden Ausgaben zu zerlegen. Nach ein wenig Sucherei bin ich auf das Konsolen-Programm „pdftk - PDF Tool Kit (Vers. 1.41)“ [1] gestoßen, das sich auch in den Mandriva-Quellen befindet. Die Installation (inkl. Abhängigkeiten) per urpmi verlief ohne Probleme.

Das Programm liefert eine Reihe von Möglichkeiten (aufteilen, mischen, Wasserzeichen Meta-Daten ändern usw.), wobei für mich nur das „fachmännische Zerlegen“ im Focus stand.

Hierbei werden aus dem Ausgangsdokument die Seiten fünf und sechs in eine neue PDF-Datei geschrieben.

Somit hatte ich das Werkzeug, aber nicht die Lust, für alle Ausgaben und Artikel den Konsolen-Hacker zu machen. Also musste eine Automatisierung her. Als Ahnungsloser habe ich dazu einen Thread [2] angestiftet um das umfangreiche Wissen der Community zu nutzen. Und dies natürlich auch mit Erfolg!

Die Script-Lösung [man-draker]

Basierend auf dem Konsolenbefehl hat man-draker die notwendigen Ergänzungen geliefert, so dass ich mehrere Befehle hintereinander in einem Script ausführen kann. Ein Beispiel sieht dann wie folgt aus (Listing1).

Zur Erläuterung: Die beiden cd-Befehle verzweigen in mein Projekt-Verzeichnis und dann ins Jahrgangsverzeichnis. Da liegen jeweils die MagDriva-Ausgaben. Darunter

liegt dann jeweils ein Artikel-Verzeichnis, in das die einzelnen Artikel kommen (entsprechender „Präfix“ für die Artikel-Dateien). Die weiteren Eingaben haben folgende Bedeutung (Listing 2).

Die GUI-Lösung [HWM]

Natürlich gibt es dafür auch eine GUI-Lösung. HWM hat eine grafische Oberfläche für pdftk als Python-Script geschaffen. Die zugehörigen Screen-Shots befinden sich im Thread [3] und das Programm [4] benötigt an Abhängigkeiten Python 2.5, Qt 4.3, PyQt 4.3 und natürlich pdftk.

Also die entpacken mit:

```
gzip -d pypdftk.tar.gz
tar xvf pypdftk.tar
```

Zum Vorschein kommen fünf Dateien. Zum Starten des Programmes liefert der Autor drei Möglichkeiten:

1. Ausführbar machen, dann auf der Konsole in das Verzeichnis der entpackten Dateien wechseln und mit ./pypdftk.py starten
2. Auf der Konsole in das Verzeichnis der entpackten Da-

3. Programm ausführbar machen, dann eine Verknüpfung auf dem Desktop anlegen und im Feld "Befehl" '>Verzeichnis>pypdftk.py' angeben

Nun hat der geneigte User eine grafische Oberfläche für pdftk zur Verfügung, mit der er seinen Zerlege-Aufgaben einfach nachgehen kann. Über den Menüpunkt „Datei“ kann eine Eingabedatei ausgewählt werden und über ein Eingabefeld werden die Seiten angegeben.

Anschließend wird im Meldungsfenster die erfolgreiche Rückmeldung

Listing 1

```
#!/bin/bash
cd /daten/Projekte/Archiv_MagDriva
cd 2008
pdftk MagDriva_2008.1.pdf cat 5-6 output Artikel/Webperlen_2008.1.pdf &>fehler.lst; if [ $? -gt 0 ]; then exit ; fi
pdftk MagDriva_2008.1.pdf cat 7-7 output Artikel/Uservorstellung_jkowalzik_2008.1.pdf &>fehler.lst; if [ $? -gt 0 ]; then exit ; fi
```

Der Link [4] liefert eine Archiv-Datei, die alle benötigten Dateien enthält.

teien wechseln und mit python pypdftk.py starten

(oder auch nicht) gegeben. Für große Aufgaben kann auch eine Eingabeda-

tei genutzt werden, welche (jeweils mit vollem Pfad) die Eingabedatei, die Ausgabedatei und die Seitenzahl benötigt, jeweils durch ein Leerzeichen getrennt. Das ganze Programm ist somit einfach und effektiv zu nutzen.

Meine DVD

Mit den oben beschriebenen Hilfen hatte ich nun alles, um einen Index für die MagDriva zu basteln. In einem OO-Calc-Blatt habe ich dann den Index mit Magdriva-Ausgabe, Artikel und Seitenzahlen erfasst. Aus diese Basistabelle habe ich dann über ein paar Funktionen in eigenen Calc-Tabellen die benötigten Scripte bzw. Eingabedateien für die GUI gebastelt.

Diese Tabellen wurden dann jeweils als Textdatei/Script (csv-Format mit Leerzeichen als Trennzeichen) gespeichert. Damit war dann das endgültige Zerlegen - nach einer Menge Tests - relativ schnell vom Tisch.

Den gesamten Index habe ich dann nochmal im OO-Writer erfasst, jeweils mit einem Link auf die Artikel-Dateien und als HTM-Datei gespei-

chert. Nun diese Seite und die Gesamtausgaben und Einzelartikel auf eine DVD gebrannt. Fertig war meine (Test-)DVD.

Nachdem nun bei mir alles geklappt hatte, habe ich die Scripte an wobo geschickt, damit er nun den richtigen Index auf unserer Seite einrichten und auch die DVD für alle bereitstellen konnte. Aber das kann er natürlich besser selber erzählen.

Das ist schnell erzählt. Die meiste Vorarbeit haben ja Magnus und mandraker bereits vollbracht. Ich habe lediglich die Pfade im Script etwas verändert und auf unserem Server die nötige Struktur erstellt. Dann ließ ich dem Ganzen seinen Lauf.

Die Erstellung des Indexes war eine Schreibaarbeit im Wiki[5], dann war die Online-Version fertig. Für die Downloadversion schrieb ich noch eine HTML-Seite als Index, packte diese zusammen mit der ganzen Struktur in einen Tarball - fertig [6].

Listing 2:

```
>fehler.lst : hängt alle Terminal-
Ausgaben von pdftk an die Datei
fehler.lst an
```

```
; : Kommandotrennzeichen
```

```
if : Anfang der Bedingung
```

```
[ : Test Anfang (Leerzeichen da-
vor und dahinter!)
```

```
$? : Das Ergebnis der letzten
Kommandoausführung (hier des
vorangehenden pdftk-Befehls)
```

```
-gt 0 : Test ob größer null, dann
Fehler aufgetreten
```

```
] : Ende Test (Leerzeichen davor
und dahinter!)
```

```
; : Kommandotrennzeichen
```

```
then : Wenn Bedingung wahr
```

```
exit : brich das Skript ab
```

```
; : Kommandotrennzeichen
```

```
fi : Ende der Bedingung
```

Links zu MagDriva DVD

- [1] Projektseite pdftk : <http://www.accesspdf.com/pdftk/index.html>
- [2] Thread „Script zum Programm-Automatisierung“: <http://www.mandrivauser.de/viewtopic.php?id=23861>
- [3] Thread „MagdDriva-DVD“: <http://www.mandrivauser.de/viewtopic.php?id=23841>
- [4] Python-Programm: <http://www.mynetcologne.de/~nc-mayha2/pypdfk.tar.gz>
- [5] MagDriva Online: <http://www.mandrivauser.de/doku/doku.php?id=allgemein:magdriva:indexseite>
- [6] MagDrivaDVD-Download: ftp://ftp.mandrivauser.de/magazin/MagDriva_DVD.tar.bz2

Kürzlich in der Redaktion unseres Newsportals. Es ging um eine Meldung, die zwei Redakteure gleichzeitig geschrieben haben:

Thorsten: *Ich habe jetzt mal deinen Kopf genommen und meinen Körper darunter gesetzt ... ich stell mir das gerade vor, das sähe irgendwie komisch aus*

wobo: *Alles mit Olivers Kopf sieht ... (upps!)*

Oliver: *Wir reden da demnächst mal drüber ...*

wobo: *Der Thorsten hat aber angefangen!*

Thorsten: *Der lügt doch, glaub dem kein Wort ... (wegrenn)*

wobo: *Irgenwann krieg ich Dich!*

Kriegt wobo den Thorsten? Glaubt Oliver dem wobo oder dem Thorsten? Wir dürfen gespannt bleiben....