

Automatisierte Dokumentenerstellung

Dr.-Ing. Herbert Voß
Freie Universität Berlin

5. September 2015

- 1 Einführung
 - Aufgabenstellung

- 2 Beispiel
 - Daten holen
 - Konvertierung

Aufgabenstellung

Selten hat man im naturwissenschaftlichen und technischen Bereich ein reines Textdokument. Sehr oft sind externe Abbildungen oder Daten einzubinden, beziehungsweise entsprechend darzustellen.

Aufgabenstellung

Selten hat man im naturwissenschaftlichen und technischen Bereich ein reines Textdokument. Sehr oft sind externe Abbildungen oder Daten einzubinden, beziehungsweise entsprechend darzustellen. Handelt es sich um statistische Auswertungen, die fortlaufend erfolgen sollen, so kann das automatisiert werden.

Aufgabenstellung

Selten hat man im naturwissenschaftlichen und technischen Bereich ein reines Textdokument. Sehr oft sind externe Abbildungen oder Daten einzubinden, beziehungsweise entsprechend darzustellen. Handelt es sich um statistische Auswertungen, die fortlaufend erfolgen sollen, so kann das automatisiert werden.

\TeX ist ein Programm (Compiler oder Interpreter), eignet sich daher für eine Stapelverarbeitung.

Vorgehensweise

Automatisiert heißt, dass der Anwender sein Skript/Programm/etc. startet und der Rest automatisch abläuft.

Vorgehensweise

Automatisiert heißt, dass der Anwender sein Skript/Programm/etc. startet und der Rest automatisch abläuft.

Dies kann erfolgen durch

- Ein Shellskript (Linux, MAC) oder Batchdatei (Windows)

Vorgehensweise

Automatisiert heißt, dass der Anwender sein Skript/Programm/etc. startet und der Rest automatisch abläuft.

Dies kann erfolgen durch

- Ein Shellskript (Linux, MAC) oder Batchdatei (Windows)
- Eine Skriptsprache (Perl, Python, Java, etc.)

Vorgehensweise

Automatisiert heißt, dass der Anwender sein Skript/Programm/etc. startet und der Rest automatisch abläuft.

Dies kann erfolgen durch

- Ein Shellskript (Linux, MAC) oder Batchdatei (Windows)
- Eine Skriptsprache (Perl, Python, Java, etc.)
- Eine T_EX-Datei

Beispiel

Aufgabenstellung: Es soll der Verlauf von Windgeschwindigkeit, Temperatur und Luftdruck grafisch dargestellt werden. Das erstellte PDF-Dokument soll über das Internet abrufbar sein.

Beispiel

Aufgabenstellung: Es soll der Verlauf von Windgeschwindigkeit, Temperatur und Luftdruck grafisch dargestellt werden. Das erstellte PDF-Dokument soll über das Internet abrufbar sein.

Die Daten bekommen wir aus dem Netz:

Beispiel

Aufgabenstellung: Es soll der Verlauf von Windgeschwindigkeit, Temperatur und Luftdruck grafisch dargestellt werden. Das erstellte PDF-Dokument soll über das Internet abrufbar sein.

Die Daten bekommen wir aus dem Netz:

`http://geo.fu-berlin.de/met/wetter/index.html`

Beispiel

Aufgabenstellung: Es soll der Verlauf von Windgeschwindigkeit, Temperatur und Luftdruck grafisch dargestellt werden. Das erstellte PDF-Dokument soll über das Internet abrufbar sein.

Die Daten bekommen wir aus dem Netz:

`http://geo.fu-berlin.de/met/wetter/index.html`

`http://www.met.fu-berlin.de/de/wetter/iframe_wetter.php`

Beispiel

Aufgabenstellung: Es soll der Verlauf von Windgeschwindigkeit, Temperatur und Luftdruck grafisch dargestellt werden. Das erstellte PDF-Dokument soll über das Internet abrufbar sein.

Die Daten bekommen wir aus dem Netz:

<http://geo.fu-berlin.de/met/wetter/index.html>

http://www.met.fu-berlin.de/de/wetter/iframe_wetter.php

2015-09-01 13:00 MESZ

Windrichtung: 284 °

Windgeschw.: 2.5 m/s

Windspitze: 4.8 m/s

red. Luftdruck: 1006.9 hpa

Luftfeuchtigkeit: 63 %

Lufttemperatur: 26.3 °C

wget

Holen des HTML-Codes der Seite mit wget:

```
wget -q http://www.met.fu-berlin.de/de/wetter/iframe_wetter.php
```

wget

Holen des HTML-Codes der Seite mit wget:

```
wget -q http://www.met.fu-berlin.de/de/wetter/iframe_wetter.php
```

Aus diesem Dokument heraus mit:

```
\immediate\write18{rm -f iframe_wetter.php}
```

```
\immediate\write18{wget -q http://www.met.fu-berlin.de/de/wetter/ifra
```

wget

Holen des HTML-Codes der Seite mit wget:

```
wget -q http://www.met.fu-berlin.de/de/wetter/iframe_wetter.php
```

Aus diesem Dokument heraus mit:

```
\immediate\write18{rm -f iframe_wetter.php}
```

```
\immediate\write18{wget -q http://www.met.fu-berlin.de/de/wetter/ifra
```

wget

Holen des HTML-Codes der Seite mit wget:

```
wget -q http://www.met.fu-berlin.de/de/wetter/iframe_wetter.php
```

Aus diesem Dokument heraus mit:

```
\immediate\write18{rm -f iframe_wetter.php}
```

```
\immediate\write18{wget -q http://www.met.fu-berlin.de/de/wetter/iframe_wetter.php}
```

Als Ausgabe erhalten wir die Datei `iframe_wetter.php`:

HTML-Datei

```
1 <html>
2 <head>
3 <title>Wetterinformationen</title>
4 <meta http-equiv="Content-Type" content="text/html;_charset=iso-8859-1"
  >
5 </head>
6 <body>
7 <p style='font-family : Verdana, Arial, Helvetica, sans-serif; line-
  height:140%; font-size:75%'>2015-09-10 11:00 MESZ<br />
  Windrichtung: <b>82 &deg;</b><br />Windgeschw.: <b>3.9 m/s</b><br
  />Windspitze: <b>6.5 m/s</b><br />red. Luftdruck: <b>1026.1 hpa</b
  ><br />Luftfeuchtigkeit: <b>64 %</b><br />Lufttemperatur: <b>16.4
  &deg;C</b></p></body>
8 </html>
```

Konvertierung

Eine einfachere Weiterverarbeitung ergibt sich nach Umwandlung der HTML-Datei in eine reine Textdatei (ASCII):

```
html2text iframe_wetter.php
```

Konvertierung

Eine einfachere Weiterverarbeitung ergibt sich nach Umwandlung der HTML-Datei in eine reine Textdatei (ASCII):

```
html2text iframe_wetter.php
```

Durchführung in einem Schritt durch »Pipen« von der Ausgabe von `wget`:

```
1 wget -q -O- http://www.met.fu-berlin.de/de/wetter/iframe_wetter  
   .php | html2text > wetter.txt
```

Konvertierung

Eine einfachere Weiterverarbeitung ergibt sich nach Umwandlung der HTML-Datei in eine reine Textdatei (ASCII):

```
html2text iframe_wetter.php
```

Durchführung in einem Schritt durch »Pipen« von der Ausgabe von `wget`:

```
1 wget -q -O- http://www.met.fu-berlin.de/de/wetter/iframe_wetter
   .php | html2text > wetter.txt
```

```
1 2015-09-10 11:00 MESZ
2 Windrichtung: 82 °
3 Windgeschw.: 3.9 m/s
4 Windspitze: 6.5 m/s
5 red. Luftdruck: 1026.1 hpa
6 Luftfeuchtigkeit: 64 %
7 Lufttemperatur: 16.4 °C
```

html2text

1	10.40	15.6	1026.1	3.0
2	10.50	15.8	1026.1	3.0
3	11.00	15.8	1026.1	3.1
4	11.10	16.4	1026.1	3.1

Die Ausgabe

