

wordif_doc

R.J. Toscani

HOE HET PROGRAMMA ‘WORDIF.SH’ TE GEBRUIKEN:

1. Voorbereiding:

Maak van te voren twee (sub-)mappen aan in de projectmap, en noem die bijvoorbeeld ‘oud’ en ‘nieuw’ (ook andere namen zijn toegestaan.)

Leid nu platte-tekstbestanden af van de te vergelijken bronteksten (in b.v. InDesign), bij voorkeur één bestand per hoofdstuk of artikel.

B.v. tekst van een InDesign-artikel selecteren met ‘Ctrl-A’, kopiëren met ‘Ctrl-C’ en in een PLATTE editor plakken met ‘Ctrl-V’.

Sla deze PLATTE-tekstbestanden (tekscodering Unicode [UTF-8]) zodanig op in ‘oud’ en ‘nieuw’ dat voor elk bestand in ‘oud’ er een daarmee te vergelijken bestand bestaat in ‘nieuw’.

Om twee bestanden in verschillende mappen aan elkaar te kunnen relateren moeten hun bestandsnamen beginnen met hetzelfde NUMMER gevuld door een UNDERSCORE, eventueel gevuld door een vrij te kiezen gedeelte (dat onderling verschillend mag zijn).

Het programma zal tekstbestanden in die in ‘oud’ zijn opgeslagen vergelijken met die in ‘nieuw’ waarvan de naam met hetzelfde NUMMER begint, bijvoorbeeld:

Tabel 1		
oud/	-	nieuw/
1_voorwoord.txt	wordt vergeleken met:	1_inleiding.txt
2_highlights.txt	wordt vergeleken met:	2_highlights.txt
3_tips.txt	wordt vergeleken met:	3_tips&tricks.txt

De laatste deel van de namen kan dus verschillend zijn.

Let op:

- Het aantal genummerde bestanden in ‘oud’ moet gelijk zijn aan dat in ‘nieuw’.
- Per map mag elke waarde van NUMMER maar één keer voorkomen.
- In bestandsnamen mogen na de ‘_’ spaties voorkomen, maar vermijd het bij voorkeur.

2. Werking van het programma:

Om het programma te starten, typ in een terminal het volgende commando in, en druk daarna op <ENTER>:

```
wordif.sh -d oud nieuw
```

De optie ‘-d’ impliceert dat er – zoals hierboven – twee mappen worden opgegeven (niet twee tekst-bestanden).

Het programma werkt nu één-voor-één alle paren van tekstbestanden in genoemde mappen af en genereert hieruit verschil-bestanden.

Deze worden elk als apart HTML-bestand in de map ‘diff’ geplaatst. De bestandsnaam is als volgt opgebouwd:

DATUM_TIJD_‘diff’_NUMMER’.html’

Het NUMMER in de namen van de verschil-bestanden komt overeen met dat van de vergeleken tekstbestanden.

Indien het programma doublures en/of in één map ontbrekende waarden van NUMMER tegenkomt, dan geeft het daarvan waarschuwingen in de terminal, maar vervolgt wel zo goed mogelijk zijn taak.

Zodra het programma klaar is, geeft het hiervan een melding in dezelfde terminal.

In elk van de resulterende .html-bestanden in de map ‘diff’ is:

- Zwarte tekst: WEL aanwezig op die plaats in BEIDE vergeleken bestanden in ‘oud’ en ‘nieuw’;
- Rode tekst: WEL aanwezig in het bestand in ‘oud’; NIET op die plaats in het vergeleken bestand in ‘nieuw’;
- Groene tekst: WEL in het vergeleken bestand in ‘nieuw’; NIET op die plaats aanwezig in het bestand in ‘oud’.

Vanuit elk van deze html-bestanden kan b.v. geprint worden naar PDF. Verder zijn er nog de volgende mogelijkheden:

- Indien tussen ‘wordif.sh’ en ‘oud’ ook nog de optie ‘-p’ wordt opgegeven, dan worden de verschilbestanden direct naar PDF in plaats van naar HTML weggeschreven.
- Indien optie ‘-d’ wordt weggelaten, dan moeten er twee tekstbestanden worden opgegeven in plaats van twee mappen. Er volgt dan slechts één vergelijking in plaats van een reeks vergelijkingen, en één verschilbestand.