

Beknopte instructie Windows Grep

Dit is een korte introductie van Windows Grep (Wingrep). Gebruik de (uitstekende) on-line helpfunctie van het programma om meer uitgebreide informatie op te vragen.

Wingrep is een programma waarmee je kunt zoeken naar tekenreeksen (*strings*) in een opgegeven verzameling bestanden. Wingrep werkt regelgericht, hetgeen betekent dat uitsluitend binnen tekstregels wordt gezocht (dat zijn regels die eindigen met een <Enter>). Daarnaast kun je bij het zoeken zgn. reguliere expressies gebruiken. Dit zijn zoekopdrachten die speciale symbolen bevatten.

Zodra je Wingrep hebt gestart (in de *Beginner mode*), wordt de *Search assistant* gestart. Deze begeleidt je door het proces van het opgeven van een zoekopdracht voor een groep bestanden. Eerst kun je opgeven welke tekenreeks je zoekt. Je kunt hierbij een aantal opties instellen. De optie "Normal (regular expressions)" biedt de meeste zoekmogelijkheden. Als deze optie is ingesteld, wordt je zoekstring niet als woord beschouwd – als je b.v. opgeeft "boek", vind je alle regels waarin de tekenreeks "boek" voorkomt, je vindt dus ook regels die woorden als "boekbinden", "zwartboek", en "geboekt" bevatten. Zie hieronder voor een aantal voorbeelden van reguliere expressies.

Als je op de knop [Next >] klikt, kun je opgeven in welke map (*directory*) de bestanden zich bevinden die je wilt doorzoeken. Linksonder selecteer je een station, linksboven selecteer je één of meer mappen op dat station. Met behulp van de knoppen [>] en [<] kun geselecteerde mappen respectievelijk toevoegen aan en verwijderen uit het *include*-venster. Als je vervolgens op de knop [Next >] klikt, kun je opgeven welke bestandstypen je wilt doorzoeken. Voor corpusonderzoek zullen dit vaak .txt bestanden zijn. In het geval van CHILDES-bestanden, voer je bij "Custom file specification" -bestanden de specificatie *. in. Ook hier gebruik je de knoppen [>] en [<] om bestandstypen toe te voegen aan of te verwijderen uit het *include*-venster.

In principe geeft Wingrep in de uitvoer uitsluitend de regels die de opgegeven zoekstring bevatten. Als je in de uitvoer klikt met de rechter muisknop, kun je opgeven dat er tot maximaal 5 regels voor en na de gevonden regels wordt getoond als context. Je kunt echter ook via de optie View -> Selected file het complete bestand bekijken. De regelnummers van de regels met gevonden zoekstrings zijn in het blauw weergegeven.

Reguliere expressies:

^	begin van de regel
\$	eind van de regel
.	(punt) willekeurig karakter
*	(asterisk) nul of meer maal het voorgaande karakter of de voorgaande expressie
.*	een willekeurige reeks (willekeurige) karakters
+	één of meer maal het voorgaande karakter of de voorgaande expressie
?	nul of één maal het voorgaande karakter of de voorgaande expressie
\	zorgt dat het volgende symbool <i>niet</i> als speciaal symbool wordt beschouwd
()	markeert een expressie
[]	groep karakters

Verder is het nuttig om te weten dat spaties worden beschouwd als onderdeel van de zoekstring. Voorbeelden:

je	Zoekt naar de string "je".
je je	Zoekt naar de string "je je".
[a-z] je [!\?\.]	Zoekt naar de string "je", voorafgegaan door een willekeurige letter en gevolgd door een spatie, uitroepteken, vraagteken of punt, en daarmee dus aan het einde van een woord, maar niet het woord "je" zelf.
CHI.*poes	Zoekt alle regels die de string "CHI" bevatten en (na een willekeurige reeks karakters) de string "poes". Met andere woorden: alle regels waarop de persoon gecodeerd als CHI het woord <i>poes</i> zegt