

Zoekinstructies van de door RSScockpit gebruikte zoekmachine Lucene

Booleaanse operatoren (AND OR NOT zie [wikipedia](#) voor uitleg)

De volgende operatoren worden gebruikt:

AND and	Met de AND operator worden documenten gevonden die beide termen bevatten. Als er niets staat tussen twee woorden wordt de AND gebruikt.
&&	&& kan ook worden gebruikt.
OR or +	Met de OR operator worden documenten gevonden die 1 van de 2 woorden bevat. Met de + voor een woord, wordt aangegeven dat dat woord in ieder geval aanwezig moet zijn.
NOT not	De NOT operator zorgt ervoor dat het woord dat volgt op NOT niet in het document aanwezig is. Dit moeten altijd 2 woorden zijn, met 1 woord werkt het niet.
!	! kan ook worden gebruikt.
-	Zie vorige.
()	Door haakjes kun je groeperen.
\	Lucene kent een aantal speciale karakters (+ - && ! () { } [] ^ " ~ * ? : \) zoals hierboven is weergegeven. Als ze toch in een zoekwoord zitten, kun je ze neutraliseren door het teken \

Op woorden zoeken

Dat kan met lossen woorden.

Een vaste volgorde van woorden kan worden aangegeven met dubbele aanhalingstekens "zoekmachine optimalisatie"

Een zin (phrase) is een groep van woorden omgeven door dubbele aanhalingstekens

Trunceren/wildcards

Truncatie is een teken dat (alle) andere tekens kan vervangen.
Dit werkt niet binnen een zin met dubbele aanhalingstekens.

Het truncatie teken voor 1 letter is een ?

Om meerdere letters te trunceren wordt een * gebruikt.

Voorbeelden:

Te?t zoekt text en test

Test* zoekt test, tester, tests

De truncatie tekens kunnen ook binnen een woord gebruikt worden, maar niet aan het begin van een woord als eerste karakter.

Fuzzy zoeken (Wikipedia over [Fuzzy Logic](#))

Dit is om met onzekerheden die in [natuurlijke taal](#) voorkomen te kunnen omgaan met formele middelen.

Lucene ondersteunt fuzzy zoekt op basis van de Levenshtein Distance, of Bewerken Distance algoritme. Om van fuzzy zoeken gebruik maken voeg een tilde "~" toe aan het einde van een woord.

Bij Lucene 1.9 kun je een getal tussen de 0 en 1 aangeven om de kans dat een woord wordt meegenomen in het resultaat aan te geven.

Voorbeeld: prijs~0.8

De default die wordt meegegeven als er geen getal wordt aangegeven is 0.5

Proximity zoeken (~)

Dit is zoeken op woorden binnen een bepaalde afstand van elkaar.

Met een getal geef je aan binnen hoeveel woorden de woorden van elkaar af moeten staan

"zoekmachine optimalisatie" ~10

De woorden zoekmachine en optimalisatie moeten max. 10 woorden van elkaar af staan.