

DAT ZOEKEN WE OP IMDb

De Internet Movie Database is al ruim 25 jaar een gezaghebbende bron voor alles wat met films te maken heeft. Loont het in een tijd van universele zoekmachines nog steeds de moeite in zo'n gespecialiseerde database te gaan zoeken?

Door: Eric Sieverts

De IMDb blijkt al sinds 1998 eigendom van Amazon. De achtergrond van Amazons belangstelling laat zich makkelijk raden. Het interface van IMDb is weinig modieus; Google heeft daar weinig invloed op gehad. Qua structuur is IMDb overduidelijk een database.

Weinig te combineren

Voor een database biedt IMDb's zoekfunctie wonderlijk genoeg weinig combinatiemogelijkheden. Zo kun je bijvoorbeeld niet gestructureerd zoeken naar films waarin een bepaalde acteur onder een bepaalde regisseur speelde. En wie (een deel van) de naam van een acteur of regisseur combineert met een woord uit de titel van een film, vindt helemaal niets. Wanneer je IMDb als een gewone zoekmachine beschouwt, zijn zulke (onterechte) nul-resultaten natuurlijk heel ongewoon.

Wat kan er dan wel? Je kunt in allerlei velden zoeken, maar ook dan valt vaak tegen wat je vindt. Of je nu in alle gegevens, in filmplots of in keywords zoekt naar de combinatie 'iceberg sea', je vindt niet 'Titanic', hoewel beide zoekwoorden in de beschrijving van die film voorkomen en ook als keywords zijn toegekend. En bij 'iceberg shipwreck' idem.

Als een AND-relatie al niet goed werkt, hoe zit het dan met OR en NOT? Als je die operatoren tussen twee woorden zet, krijg je juist de AND-combinatie. Begrijpt u het nog? Combinatiezoeken is dus niet waar de IMDb het van moet hebben.

Fuzzy zoeken

Als je wel wat vindt, weet je zelden hoeveel dat is, want IMDb vermeldt meestal geen aantallen. Resultaten van verschillende zoekacties zijn dus moeilijk te vergelijken. Pas door wat redeneren zie je dat IMDb eventueel toegevoegde truncatietekens gewoon negeert, maar dat standaard wel zeer 'fuzzy' gezocht wordt. Met *woody* vind je ook *wood*, *woodley*, *sherwood*, *eastwood* enzovoort. *Misdemeanors* levert ook *'myth demeanors'*, *'witch demeanors'* en *'mist demeanor'*. (Nooit geweten dat er ook (tv-)films met die titels waren...) Exacte match is desgewenst wel mogelijk, maar alleen bij persoonsnamen.

Geavanceerd zoeken

Naast het onbevredigende zoekmachine-achtige zoeken, zijn er wel een paar interessante speciale mogelijkheden. Zo is er Advanced Title Search, waarbij je een titelwoord kunt combineren (of filteren) met een hele rits formele kenmerken, waaronder user rating, filmlocatie en genre. Verrassend is dat vrijwel geen van die inperkingen met keuzelijstjes werkt, zodat je maar wat moet uitproberen. Advanced Name Search biedt minder combinatieopties en ook minder nuttige, tenzij je geïnteresseerd bent in welke vrouwelijke acteurs met sterrenbeeld Stier een best actress-nominatie gehad hebben. Door geen naam in te vullen krijg je wel alle personen die aan een gekozen

combinatie van criteria voldoen. Bij de titels geldt dat uiteraard ook.

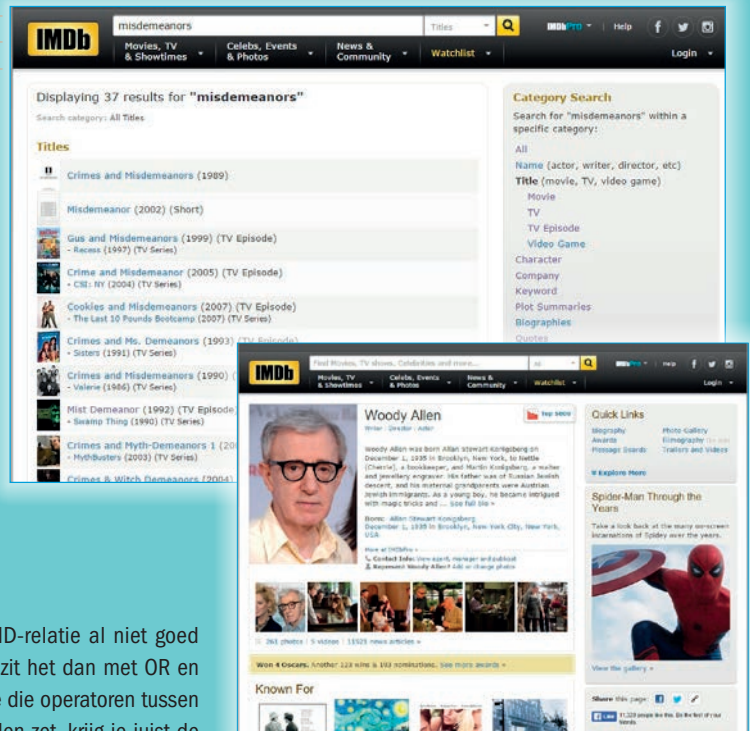
Daarnaast zijn er twee 'overlap'-zoekfuncties: welke personen samen in twee opgegeven films zaten en in welke films twee opgegeven personen samen zaten. Allemaal best leuk, maar het blijft een gemiste kans dat je van alle andere combinatiemogelijkheden die je op basis van de rijke inhoud van deze database zou kunnen bedenken, niets zelf kunt uitproberen.

Voor echt zoeken naar Google

Voor zoekers uit de Google-generatie voelt zoeken in de IMDb vaak als een regelrechte ramp. Er gebeurt zelden dat je verwacht, en er wordt niet met je meegedacht. Als je niet een van de speciale IMDb-opties nodig hebt, kun je beter naar Google uitwijken. Alle interessante inhoud van de IMDb is daar wel vindbaar. En daar kun je wel alles combineren wat je wilt. Met *iceberg shipwreck site:imdb.com* zitten bij de eerste resultaten (van de 1630) wel meteen diverse IMDb-pagina's over 'die ene film'. En op de combinatie *elaine seinfeld site:imdb.com* vind je wel ruim 7000 resultaten in plaats van de 'no results' van IMDb zelf. En verder doorlinken tussen al die elementen uit de database gaat vervolgens wel goed.

Dat IMDb toch een beetje met zijn tijd meegaat, blijkt uit het feit dat zijn inhoud wel als bron voor de Linked Open Data-cloud fungeert.

Eric Sieverts, redacteur van IP en freelance docent en adviseur



URL	www.imdb.com
Booleaans combineren	heel beperkt
Truncatie	nee (zoekt wel fuzzy)
Speciale zoekvelden	ja
Advanced zoekscherm	ja
Parametrische filters	nee
Ook in Google	ja
Semantische markup	ja