

WWW search engines (2)

Nog steeds onbetrouwbaar

Het is helemaal niet de bedoeling dat WWW-zoekmachines betrouwbaar zijn! Dat is de kern van de kritiek van Hein Ragas en Onno Stegeman in het vorige nummer van Informatie Professional. Hiermee reageerden zij op het onderzoek van de Werkgroep IRT in Informatie Professional nr. 11/99. Hans van der Laan en Wouter Mettrop hebben zo hun twijfels over die kritiek, en de zoekmachines blijven inmiddels onbetwist onbetrouwbaar.

ASPIRINE GRAAG, DE echte van Bayer'. Zwijgend legt de apotheker een buisje Paracetamol voor me neer. Op mijn verbaasde blik ontdooit hij: 'Apothekers zijn alleen geïnteresseerd in de relevantie van de tien meest gevraagde geneesmiddelen. Om de relevantie van die geneesmiddelen zo hoog mogelijk te krijgen worden door apothekers allerlei heuristische gebruikt. Dat zal ik u eens even uitleggen...' Nu sta je bij de huidige politiek nergens meer van te kijken, maar bij deze apotheek zou ik niet terugkomen. Mag ik alsjeblief zelf aangeven wat ik wil hebben?

Hopelijk krijgen we van Hein Ragas en Onno Stegeman nu niet het verwijt dat wij menen dat zoekmachines en apothekers met elkaar vergeleken kunnen worden. Wij beseffen het, er zijn 'belangrijke verschillen' tussen zoekmachines en apothekers. Het voorgaande was dan ook maar een grapje. Zo was de inleiding op ons artikel 'Naast incompetent zelfs onbetrouwbaar' ook bedoeld, een aardige binnenkomer voor een overigens serieus betoog. Het was niet de bedoeling om zoekmachines en catalogi met elkaar te vergelijken en dat hebben we dan ook niet

gedaan. Helaas zijn Ragas en Stegeman die weg wel ingeslagen en ze hebben hem hardnekkig gevolgd. Verdwaald als zij zijn zouden we hen daar kunnen laten in hun conclusie dat de resultaten van ons onderzoek dus niet relevant zouden zijn. Maar zij beweren ook dat die resultaten niet geldig zouden zijn en dat misverstand moet worden rechtgezet. Misverstand, want onze opponenten hebben

een essentieel deel van de methodologie van onze waarnemingen gemist. En ja, dan gaat het onderzoek zelf er ten onrechte onbetrouwbaar uitzien. 'Dat zullen we u eens even uitleggen...'

Terzijde: Catalogi kunnen bogen op volledigheid. Laat niet het misverstand ontstaan dat de werkgroep zou hebben gekeken naar grootte of dekking van de vergeleken zoekmachines, die factor speelt geen enkele rol in het onderzoek. Er zijn grote en kleine onbetrouwbare machines.

Methodologie

Ragas en Stegeman verwijten ons onzorgvuldigheden waarvoor we ons inderdaad zouden schamen als ze waar zouden zijn. Er zijn, zo zeggen zij, twee soorten fouten gemaakt: *Fouten door aannames over de wijze van zoeken*, en *Fouten door de tests*. Als we hier nu even hun veronderstelde niet-professionele zoeker en onze apotheker buiten beschouwing laten komen die beide fouten ongeveer op hetzelfde neer: Er komen, zeggen ze, wagonladingen treffers binnen op onze vragen en mogelijk zouden we hier onze testpagina's niet in kunnen terugvinden. Hier wordt AltaVista aangehaald die maar tweehonderd treffers zou doorgeven, dus ze zitten er misschien niet eens bij. Zijn wij nou heus zo dom geweest?

Welnee, dat hebben we natuurlijk wel voorzien. Als we domweg op 'vier vijf zes' zouden zoeken of, nog erger, op 'Government' (8.326.593 treffers!), dan zouden we roekeloos tewerk gaan. Maar als we zoeken op 'Misselthwaite and IRT and Government' dan vinden we (9 december 1999) slechts vijf treffers. Al onze testvragen zijn vanzelfsprekend op deze manier zo specifiek gemaakt dat ze vrijwel uitsluitend onze eigen testpagina's vinden, of (soms) nou net niet natuurlijk, dat is juist de kern van onze bevindingen: zoekmachines zijn onbetrouwbaar. Er bestaat dus weinig kans dat de treffers niet bij de eerste tweehonderd zouden zitten (zo veel antwoorden hebben we nooit gehad), en het verschil in taal tussen 'vier vijf zes' en 'Government' speelt dus ook geen rol.

We laten nu maar even buiten beschouwing dat Ragas en Stegeman de gevonden *documentfouten* en *elementfouten* op één hoop gooien, en eigenlijk zelf 'bewijzen' dat ze vanzelfsprekend zouden moeten voorkomen (?), dat zou flauw zijn. In ieder geval berust hun bewering 'de conclusies [...] zijn, door de fouten in de methodologie, niet geldig' op een misverstand en dat is jammer. De hele methodologie is op de website te vinden. Sans rancune.

En nu verder

Maar van de rest van de beweringen van onze opponenten vallen – herinnert u zich nog – ons de oren van het hoofd. Vandaar dat ons grapje over die apotheker toch wel een serieuze ondertoon heeft. Laten we even ernstig blijven. Het lijkt erop dat Ragas en Stegeman niet hebben begrepen waar ons artikel over ging. In de bibliotheekwereld kennen we het onderscheid tussen onderwerps- en known-itemvragen. Over die laatste hebben wij ons artikel geschreven en de gebreken die we daarin vaststellen hebben natuurlijk ook consequenties voor de eerste (als je op 'Brontë sisters' zoekt kan het immers gebeuren dat je Cecilia Falks pagina soms wel, soms niet vindt, net zoals dat met de IRT-testpagina's zo is). De reactie van Ragas en

Stegeman gaat vrijwel uitsluitend over wat zoekmachines allemaal doen om bij onderwerpsvragen de meest 'relevante' documenten naar voren te halen. Dat was ons ook wel bekend, alleen je hoort er nooit iemand over dat het van de luimen van de zoekmachine ('zoekmachine was er even niet bij, zoekmachine was even koffiedrinken, zoekmachine had het even te druk') afhangt wat die relevante documenten zijn. Als een zoekmachine beweert een groot deel van het web en hele webdocumenten te indexeren (en dat doen de meeste zoekmachines), moeten we niet doen alsof iedereen wel weet dat dat niet zo is.

De bewering van Ragas en Stegeman dat je zoekmachines niet mag vergelijken met catalogi berust hoofdzakelijk op twee stellingen:

- zoekmachines werken nou eenmaal zo, en
- dat is voor gebruikers niet interessant.

Als zoekmachines nou eenmaal zo zouden werken – en daar hebben wij zo onze twijfels over – dan is het bijzonder interessant dat gedrag in de praktijk te onderzoeken. Al was het alleen maar voor de categorie zoekers die known-itemvragen stelt. In het door onze opponenten gestelde kader is het daarbij bijzonder boeiend dat sommige zoekmachines in dit opzicht nou juist wel (vrijwel) foutloos blijken te werken. Iets meer in detail: Voor de elementfouten geldt dat vier machines geheel foutloos werken en zes machines bijna foutloos. Drie machines zijn foutloos op het punt van documentfouten. Dat lijkt ons van belang voor alle gebruikers, hoe 'niet-professioneel' ook, en hoe slecht hun 'mentale model' ook moge zijn. Ragas en Stegeman menen dat voor de meeste gebruikers volledigheid en terugvindbaarheid niet van belang zijn. Maar simpele (en zelfs niet zo simpele) gebruikers denken vaak ten onrechte, en tot hun ongenoegen, dat 'het niet op internet staat' als ze niets of weinig vinden. En hoevelen van hen zoeken niet vertwijfeld naar die interessante pagina die ze gisteren nog hebben bekeken? Het is voor de gebruiker van een zoekmachine niet interessant wat de bedoelingen van de ontwerpers ervan zijn (zeker als hij die niet te horen krijgt), welke heuristieken zij gebruiken en hoe goed zij het met hem voorhebben. Wat voor hem telt is of ze hun slogans als 'de grootste', 'de beste', 'de compleetste' kunnen waarmaken. En vooral: hij wil resultaten, soms heel specifieke.

Onbetrouwbaar

Als samenvatting kunnen wij stellen dat aan de geldigheid van de onderzoeken van de Werkgroep IRT niet is getornd. Aanmerkingen op de methodologie berusten op een misverstand. We hebben zoekmachines niet vergeleken met catalogi, nog minder hebben we een onderzoek naar hun grootte of dekking gedaan. Het is voor de gebruiker van een zoekmachine niet interessant wat de bedoelingen van de ontwerpers ervan zijn. En dat alle ontwerpers snode lieden zouden zijn wil er bij ons niet meteen in. En overigens zijn wij onveranderd van mening dat (sommige) zoekmachines onbetwist onbetrouwbaar zijn.

Drs. H.R. van der Laan is computerraadsman te Leiderdorp en redactielid van Informatie Professional. Drs. M.W. Mettrop is medewerker bij de Bibliotheek van het Centrum voor Wetenschap en Informatica (CWI) in Amsterdam.