

GEEL EN GROEN

HET BEDROG VAN DE ONLINE PUBLIEKSCATALOGUS*

De OPC, online publiekscatalogus, is het instrument waarmee de samenwerkende Nederlandse serieuze bibliotheken hun bezoekers toegang geven tot hun collecties. Die bezoekers hebben in de meeste gevallen geen toegang tot de magazijnen waar de collecties staan opgesteld en zij zijn daardoor afhankelijk van wat de catalogus te bieden heeft. Alleen door de informatie die de eindgebruiker aan de OPC weet te ontfutselen kan hij weten of boeken of tijdschriften die hij zoekt aanwezig, beschikbaar en uitleenbaar zijn. Geeft de OPC geen resultaat, of, erger nog, doet het apparaat het niet, dan is dat een regelrechte ramp. De bibliotheek is dan in feite dicht.

De meeste UB's hebben een aantal terminals staan, van waaruit de OPC wordt aangesproken. Via de PC's van het Publieksnetwerk, via de terminals en PC's in het netwerk en via inbellen met een modem is het mogelijk de geautomatiseerde catalogus van de collecties van een groot aantal bibliotheken te bekijken. Ook is het mogelijk de OPC's van nog eens achttien grotere, en vooral kleinere, collecties in den lande te raadplegen. Het ziet er allemaal indrukwekkend uit. Maar het mooiste is dat, vergeleken met de vroegere Leidse boekjes of een alfabetische kaartcatalogus, met een veel groter aantal zoek sleutels in de bibliografische bestanden kan worden gezocht.

Laten we gewoon eens een boek opzoeken, bijv. via het scherm dat de bezoeker/inlogger krijgt aangeboden in de UB Leiden. Met de zoek sleutel AUT, achternaam van de auteur, zal weinig mis kunnen gaan. DAT HAD U MAAR GEDACHT: de meest eenvoudige zaken gaan mis, en echt niet omdat die domme bezoeker (eindgebruiker in het jargon) er weer niets van weet te brouwen. En zelfs al zou dat zo zijn, een systeem dat van zijn gebruikers domme mensen maakt is altijd een slecht systeem.

Op auteur

Dus op zoek naar een boek. Kijken we naar het voorbeeldje in het AUT-menu. Behalve de fiscalist 'a.j. van den bos' wordt hier Nederlands grootste schrijver, Willem Frederik Hermans, als voorbeeld gebruikt voor de syntaxis van het zoekcommando. Dat er in het menu geen hoofdletters worden gebruikt stoort, maar is niet wezenlijk van belang. Wij zoeken eerst onder 'hermans, willem frederik'. Resultaat is 147 treffers. Niet gek. Dan gebruiken we voor de aardigheid, of omdat we te lui zijn om die twee voornamen voluit te tikken, ook de tweede syntax voor het commando (hermans,w f), maar o wee, hadden we dat maar niet gedaan. Alle 147 treffers zijn nu verdwenen, en er is nog maar één titel overgebleven, toepasselijkerwijs een een-acter, uit 1953. Luiheid en vrolijkheid worden dus onmiddellijk afgestraft. Gierigheid ook trouwens, want de boeken van Hermans zoekt men niet in de bibliotheek, maar in de boekhandel. Wie voor de zekerheid zoekt op 'hermans,w' (wat niet als suggestie door de OPC wordt aangeboden) krijgt weer wél alle Hermansen met initiaal W., ook anderen dan Willem Frederik.

Een landelijke rondgang op 26 maart 1995 door de aangesloten OPC's op zoek naar Hermans en zijn alter ego levert het volgende resultaat. Vooraf dient vermeld te worden dat het menu van de overige catalogi die via OPC zijn te raadplegen aangeven of verbinding mogelijk is ('vrij') of niet ('onbereikbaar'). Bij herhaald raadplegen van dit menu blijkt echter dat sommige instellingen zich spontaan bij het menu voegen, en dat andere er onverklaard van worden verwijderd. Dat kan niet anders dan de intelligente bezoeker voor problemen stellen.

Naam bibliotheek	A	B	C	D
1. Leiden UB	147	1	158	1715620
2. Centrale catalogus Ned.	360	22	449	7 milj.
3. OCLC. World Cat.	0	0	0	29 milj.
4. KB, Den Haag	167	0	188	2042347
5. Groningen, UB	150	5	165	1748835
6. Tilburg, KUB	59	0	80	790860
7. Maastricht, UL	5	1	10	488124
8. Enschede, TH	2	1	6	356404
9. Nijmegen, KU	95	0	124	1520494
10. R'dam, UB	52	0	58	734135
11. R'dam, medische bibl.	0	1	5	122919
12. Provincie Overijssel	0	0	0	26482
13. A'dam, KNAW	2	0	6	228074
14. Kampen, Theol. Hogesch.	4	1	7	145802
15. Den Haag, Vredespaleis	0	0	0	90368
16. Tilburg, Kath. Leerg.	19	0	19	51915
17. Breda, KMA	0	0	2	114402
18. Zoetermeer, Min. OCW	0	0	0	59937
19. Den Haag, NBLC	111	1	114	385443
20. Leeuwarden, Prov. Bibl.	76	0	80	465259
21. Deventer, SAB	73	20	99	234536
22. Leiden, KITLV	6	2	6	201598
23. Kampen, Theol. Hogesch. (B)	0	0	2	129015

Diagram 1. Zoekacties Hermans via de voorbeelden op AUT.

Toelichting:

- kolom A = zoekactie met hermans,willem frederik
- kolom B = zoekactie met hermans,w f
- kolom C = zoekactie met hermans,w
- kolom D = totaal aantal titels in de betreffende OPC

Opmerkingen.

1. Bijna alle OPC's gebruiken het voorbeeld met 'van den bos' en 'hermans'. Alleen Maastricht (7), KITLV (22) en het Vredespaleis (15) gebruiken geheel verschillende namen. De KNAW (13) gebruikt 'johannes herder' en 'rudi van den bos', overigens geen auteurs die in de bibliotheek van de KNAW te vinden zijn, als ze al bestaan. Het Ministerie (18) heeft van de voornamen van 'hermans' anton frederik gemaakt, welke fictieve persoon evenmin met werk in de ministeriële collectie is vertegenwoordigd. Bij de Katholieke Leergangen in Tilburg (16) heeft men de andere grote Nederlandse schrijver als voorbeeld genomen, de bekende katholieke reve, maar dan

wel met de voornamen 'gerard karel' — als Siamese tweeling verenigd met zijn geleerde broer.

2. De vergelijking tussen kolom A en B geeft de verschillen in resultaat aan van de twee in de OPC als gelijkwaardig alternatief gepresenteerde zoekmethoden. Kolom C geeft altijd meer hits dan A en B tesamen, omdat andere auteurs, bijvoorbeeld Ward H., Wilhelmina H., daarin mede begrepen zijn. Het droeve vermoeden vat post dat de verschillende OPC's Willem Frederik Hermans en W.F. Hermans als twee verschillende auteurs beschouwen.

3. Kolom D is alleen gegeven om aan hermansianen de mogelijkheid te geven de hermans-factor van de verschillende verzamelingen te berekenen. Dit valt verder buiten ons onderwerp. Het is wel aardig dat het Zoetermeerse ministerie geen titels van Hermans verzamelt, maar het geeft te denken dat OCLC met zijn ruim 29-miljoen titels al evenmin een Hermans bevat. Is hier misschien sprake van een foutje in de zoek-software?

4. De Stads- en Atheneum Bibliotheek te Deventer heeft als enige een grote score op de zoekactie hermans,w f. Wie die set van 20 titels bekijkt valt van de ene verbazing in de anderen. Hier moet een volstrekt dyslexe catalogisant bezig zijn geweest, met het zelf maken van titels die alle ontleend hadden kunnen worden aan de GGC. Het lijkt erop dat de 22 titels van deze categorie via Deventer in de CCN terecht zijn gekomen.

5. Tenslotte moet nog opgemerkt worden dat een aantal OPC's in den lande (niet die van Leiden) reageerden op de zoekactie 'hermans,willem frederik' alsof 'frederik' de achternaam van een tweede auteur was.

6. Voor KITLV (22) en Kampen-B (23) werd op 30 maart 1995 gezocht.

Op titel

Een van de grote voordelen van een geautomatiseerde catalogus is dat men ook kan zoeken op woorden uit de titel (zoektermen in het jargon). Dat was met de traditionele auteurscatalogus, de Leidse boekjes, nauwelijks mogelijk. Alleen de anonieme werken waren daar onder hun titel opgenomen.

De zoektocht met elementen uit de titel verloopt zo mogelijk nog curieuzer dan het zoeken op auteur. We nemen nu eens niet het voorbeeld dat de catalogus ons aanbiedt, maar zoeken nu naar titels die we echt willen inzien. De TST zoek sleutel wordt als tweede in het lijstje van zoek sleutels gegeven. Op het scherm wordt deze sleutel omschreven als 'titel'. We zoeken op het woord Mekka. Score is 3 treffers. Nu zoeken we Mekka met de zoek sleutel TTL. Deze sleutel staat omschreven op het scherm als 'woorden uit de titel'. Score is nu 24 treffers, waaronder de 3 die we ook met TST hadden gevonden. Met het scheermes van Ockham kunnen we de TST zoek sleutel elimineren. Weg met die onvolledig werkende zoek sleutel TST waarmee met maar 3 van de 24 mogelijke treffers op het scherm krijgt! Waarom staat die misleidende zoek sleutel op zo'n prominente plaats? Een van de 24 treffers is, natuurlijk, Snouck Hurgronje's meesterwerk Mekka uit 1888. De OPC noemt het (zowel met TST en TTL), maar alleen het exemplaar in Instituut Kern, in de Nonnensteeg te Leiden. De verschillende sets van dit prachtboek (2 dln en platenatlas) in de Leidse UB moet men dan in de Leidse boekjes opzoeken (waar de auteur die wereldwijd

bekend staat als Snouck etc. overigens opgezocht moet worden onder Hurgronje).
Wie zou die moeite nemen? Goed voor het boek, slecht voor het bezoek.

Als we nu vervolgens Mekka opzoeken met de zoek sleutel GTR (het trefwoord van de nieuwe onderwerpsontsluiting) krijgen we te horen dat er 0 treffers zijn. Hoe nu? Gelukkig wordt de index van trefwoorden getoond, en daar zien we dat het juiste trefwoord niet Mekka is, maar: Mekka, bedevaart naar. Op dat trefwoord zijn 23 treffers, en dit zijn allemaal andere titels dan wat een zoekactie via de sleutel TTL op Mekka oplevert. Via zoeken met ALL vinden we tenslotte 49 treffers (48 titels en een thesaurusverwijzing).

Resumerend:

zoekterm	zoeksleutel	treffers
Mekka	TST	3
Mekka	TTL	24
Mekka	GTR	23
Mekka	ALL	49

Diagram 3. Mekka-treffers volgens vier verschillende zoeksleutels.

Hetzelfde gaan we nu doen met het woord Islam, dat zich leent om onder nog wat meer zoeksleutels te laten opzoeken.

zoekterm	zoeksleutel	treffers
Islam	AUT	47 (waarvan 16 info)
Islam	TST	66
Islam	TTL	> 1000 (wordt niet getoond)
Islam	ALL	> 1000 (wordt niet getoond)
Islam	COR	140
Islam	RKS	178
Islam	CON	19
Islam	IMP	124
Islam	PAO	20
Islam	SYS*	'te groot om te tonen'
Islam	GTR	336
Islam?	GTR	608
Islam	BCL**	16

Diagram 4. Islam-treffers volgens verschillende zoeksleutels.

* SYS met aanduiding Fw.

** BCL gecumuleerde opgave van treffers met de subrubrieken 11.80, 11.81, 11.82, 11.83, 11.84.

Opmerkingen.

1. De cijfers hoeven niet met elkaar in tegenspraak te zijn. De 16 info-regels bij AUT komen vrijwel overeen met de 20 treffers van PAO. Opvallend is opnieuw het enorme verschil tussen de treffers op TST en TTL.

2. Driemaal bleek een set te groot om te tonen te zijn: bij gebruik van zoek sleutels TTL, ALL, SYS. De sleutel TTL is in Leiden voor dit algemene trefwoord dus niet te gebruiken, terwijl de sleutel TST weer veel en veel te weinig vertoont.

4. GTR maakt een aanzienlijk verschil tussen zoekterm Islam en zoekterm Islam?. Dat is wellicht te verklaren door aan te nemen dat alle treffers van GTR-Islam ook begrepen zijn in GTR-Islam?. Ik heb verder niet nagegaan of deze overlap inderdaad tussen beide verkregen sets bestaat. Opvallend is het grote verschil in treffers tussen het trefwoorden-systeem van de Gemeenschappelijke Onderwerpsontsluiting (GOO) en de basis-classificatie: 608 tegen 16. Ik heb er geen goede verklaring voor.

3. De beperking dat een set van meer dan duizend titels niet getoond wordt getuigt van het paternalisme van de makers van het systeem. Er zijn bezoekers die wel degelijk de moeite nemen om duizend of meer titels te bekijken. Want wat is duizend boeken nou helemaal? Twee of drie kasten vol? Mag daar niet in alle rust naar gekeken worden? Kennelijk niet, menen de makers van het systeem. Maar de catalogus is wel de enige toegang tot de bibliotheekcollectie, de bezoeker kan immers niet zelf voor de kast gaan staan en boek voor boek ter hand nemen. De beperking van duizend is een klap in het gezicht van de wetenschappelijk onderzoeker. Het is hoog tijd dat er wordt terug geslagen.

Diep treurig

Resumerend kan vastgesteld worden dat het zo eenvoudige zoeken op auteur of titel nog verre van simpel is. Gaat men op zoek naar fouten in de wijze van presentatie van meerdelige werken of boeken die in een reeks zijn verschenen, dan openen zich wijds perspectieven voor vrolijke uren — als het niet zo diep treurig was. We laten dan nog buiten beschouwing de vaak onjuiste mededelingen over de uitleenbaarheid van boeken (met name die op studiezalen), en de talrijke fouten bij de locatie van tijdschriftseries. Met een beetje doorzoeken blijkt het systeem aan alle kanten te rammelen. De presentatie van een typografisch ‘moeilijke’ taal, het Turks (een Europese taal met het Latijnse alfabet), laat het al helemaal afweten. De presentatie van uit andere alfabetten getranscribeerde titels is ronduit angstaanjagend, en komt niet meer in aanmerking voor citeren vanaf het scherm. Dit eurocentrisme, of is het gewoon gemakzucht?, mag vreemd genoemd worden voor een universiteit die zich zozeer met zijn niet-westerse studiën wil profileren.

Wie een geautomatiseerde catalogus aanspreekt en om informatie vraagt heeft er recht op om een adequate reactie te krijgen en niet bij toepassing van verschillende zoek sleutels zeer verschillende resultaten te krijgen. Wie de cijfers in bovenstaande diagrammen bekijkt en wie de beschreven ervaringen kent en herkent kan niet anders dan alle vertrouwen in de OPC verliezen. Onder de bezoekers wordt behoorlijk geklaagd over de OPC, en niet zelden pakt de bezoeker een andere terminal als hij bij zijn zoektochten is vastgelopen. En of de bezoeker echt alles te zien krijgt wat de lokale catalogus in zich bergt blijft voor hem steeds de vraag.

De OPC is een arrogant, klant-onvriendelijk en onbetrouwbaar systeem. Het vertoont maar een deel van de titels die de bezoeker wil zien. De rest houdt het onzichtbaar vast, tenzij men even zotte zoekacties gaat ondernemen als waarmee ik mij, blijkens bovenstaande diagrammen, heb bezig gehouden, en dan nog weet je het nooit zeker. De OPC maakt de boeken kwijt, en dat is in een bibliotheek het ergst

mogelijke. De OPC ondermijnt daardoor de kernactiviteit van de bibliotheek, die is: wetenschappelijke informatievoorziening en de bijbehorende dienstverlening.

Een ramp

De OPC is dus een ramp, maar daarmee nog geen natuurramp. Door menselijk handelen zou van de OPC een bruikbaar instrument gemaakt kunnen worden. Ik ga daartoe hier geen aanbevelingen doen. Ik heb er een hekel aan om het werk van anderen te doen. Ik heb hiervoor alleen aangetoond, vanuit het perspectief van de bezoeker van een bibliotheek — zoals ik mij primair zelf ook altijd ben blijven beschouwen —, dat de OPC nu niet naar behoren functioneert. Ik ben er overigens helemaal geen voorstander van dat de oude catalogus, de Leidse boekjes, tot in lengte van jaren in gebruik blijft. Er moet iets gebeuren, dat is duidelijk.

Het is duidelijk dat de bezoeker van de bibliotheek bij PICA's OPC aan het kortste eind trekt. Pica heeft geen boodschap aan de moderne analfabeet die zijn boeken niet eens op het beeldscherm kan vinden. Pica heeft het ook zo georganiseerd dat het nooit met zulke domoren wordt geconfronteerd. Wie de ronkende zelfvoldaanheid van het Pica-establishment leest op de pagina's van bijvoorbeeld het jubileumnummer van Pica-mededelingen (ISSN 0165-7984), jrg. 17, nummer 4 (december 1994) en wie leest hoe deze heren en dames elkaar en zichzelf feliciteren met wat Pica heeft bereikt (succes, mirakel. Gideonsbende, creatief (p. 2), prachtig hulpmiddel, verrijking, voorbeeldig, enorme uitdaging (p. 4), betrouwbaarheid, hoogwaardig, degelijk, innovatief (p. 5), ziekte (p. 11), Pica, c'est moi!, restaurant, beneveld, mooie jaren (p. 16), begrijpt dat bij al dat feestgedruis de bezoeker van een bibliotheek, die verhinderd wordt een boek te vinden in de OPC en zijn beklag daarover doet, door de politie verwijderd moet worden als verstoorder van rust en orde.

In het Pica-jargon is de gebruiker niet die wanhopige bezoeker die van OPC-terminal naar OPC-terminal loopt om een boek te zoeken dat hij echt nodig heeft, en waarvan in heel Nederland misschien maar één exemplaar aanwezig is. Die ongelukkige heet de eindgebruiker en is rechteloos. Maar om die eindgebruiker gaat het nu juist. De adequate bibliografische dienstverlening aan die eindgebruiker is het enige bestaansrecht van de bibliotheken, en daardoor ook van superstructuren, of moet ik zeggen uitwassen?, als de Stichting Pica. De bezoeker heeft alleen maar de OPC, dus de kwaliteit daarvan is het enige criterium waarop de doelmatigheid van al het overige werk van Pica moet worden beoordeeld. Als die eindgebruiker naar een bibliotheek komt in de verwachting de boeken, die hij nodig heeft voor onderwijs of onderzoek en die ook in die bibliotheek aanwezig zijn -, en als die eindgebruiker dan door een onhanteerbaar en onbetrouwbaar instrument als de OPC in de steek wordt gelaten, dan komt die bezoeker inderdaad bedrogen uit. En dat is wat ik bedoel met de titel van deze voordracht: het bedrog van de OPC.

* Ingekorte versie van een voordracht, gehouden in de Universiteitsbibliotheek te Leiden op 31 maart 1995

J.J. WITKAM

VAKGENOOTSCHAP

Stuudiemiddag <i>Legenda aurea</i>	49
Digitaal tekstcorpus	49
Betaalgemak IVN	50
Vierde Bert van Selm-lezing	50
Het beeld van de Middeleeuwen in de negentiende eeuw	51
Van Ginneken-dag	51
Bavianen en slijkgeuzen op muziek	51
WNT op CD-ROM	52
De derde grote taaltest van de V.A.N.	52

GEEL EN GROEN

Het bedrog van de online publiekscatalogus	53
--	----

GEPROMOVEERD

C.F.M. Streng, G. Warnar, J. Ytsma, J.B. Oosterman, P. Ackema en I. Bejczy	59
---	----

VAN DE ADMINISTRATIE

64

PAS VERSCHENEN

Catalogus bibliotheek Vestdijk	65
Communicatie	66
Intra- en extramurale adreslijsten IVN	67
Een Beytie Hollansche	68
Handelingen IVN congres	68
Handboek Literatuuronderwijs Nederlands 1995	69
Over de geschiedenis van de taalkunde in Nederland	69
Nieuwe AD&L-uitgaven	70
Prachtbijbel blijkt prachthandel	71
Groot Tesourus van Afrikaans	71
SNL-reeks	72
Javaanse reyse	72
'Dutch Literature in the Age of Rembrandt' in het Nederlands vertaald	73
Rotwijven en meidepiemels uit Den Bosch	74
Cahiers Zuidnederlandse letterkunde	75

RECENSIE

Een onvoldragen Girafje (W. HENDRIKS)	76
---------------------------------------	----