

Society of the Query

magazine

10 ARTIKELEN
OVER ZOEKEN
OP HET WEB

Stop met zoeken, begin met vragen!

Introductie Society of the Query magazine

Miriam Rasch 1-3

Google: een database van intenties?

Kylie Jarrett 4-9

Google tussen standaardisering en individualisering

Andrea Miconi 10-13

Zoeken in een zoekmachine die steeds verandert

Dave Crusoe 14-19

Wat Google autocomplete zegt over wie we zijn

Mél Hogan & M.E. Luka 20-24

Op zoek naar de voorouders van Google

Anton Tantner 25-30

Ik ben geen zoekresultaat! Ik ben een vrij woord!

Anna Jobin & Olivier Glassey 31-36

Zoeken over de grens: China

Min Jiang 37-42

Google staat boven de wet en voor de rechter

Angela Daly 43-47

De ideologie van alternatieve zoekmachines

Astrid Mager 48-55

Pleidooi voor een onafhankelijke index van het web

Dirk Lewandowski 56-59

Slim zoeken?!

Miriam Rasch 60-64

Meer weten?

Miriam Rasch 65-66

Colofon

67



Stop met zoeken, begin met vragen!

Introductie Society of the Query magazine

Miriam Rasch



'Even googelen...' Hoe vaak komt het niet voor dat je je telefoon uit je zak haalt om snel iets op te zoeken? Zelfs als Google niet de startpagina van je browser is, zie je waarschijnlijk als je internet opstart toch als eerste een zoekbalk. Je moet er niet aan denken dat het anders zou zijn, hoe zou je in die zee van informatie ooit je weg kunnen vinden? Inmiddels weten we wat er nog meer achter zit. De onthullingen door Edward Snowden over PRISM en de NSA laten zien hoe wijdvertakt de controle is met behulp van digitale diensten als Google en Facebook en de directe banden die zij hebben met overheidsinstanties, niet alleen in de VS maar ook in Nederland. Een belangrijk issue als het gaat om privacy. Denk maar eens aan wat je allemaal opzoekt online: niet alleen feitjes (hoe heette die acteur ook alweer die een Gouden Kalf won?) of onderzoek naar studiebronnen (in welk jaar werd de eerste film met geluid vertoond?), maar ook twijfels over liefde (ben ik raar als...), symptomen van mogelijke ziektes (wat betekent dat bultje daar...) of misschien zelfs filmpjes van terroristische acties (puur uit nieuwsgierigheid natuurlijk). Al die informatie wordt opgeslagen en te gelde gemaakt. Wat betekent het dat 95% van het Nederlandse marktaandeel in online zoeken in handen is van een Amerikaanse, commerciële zoekmachine die werkt op onbekende algoritmes?

Zoeken op het web is zo gewoon geworden dat we er niet meer bij stilstaan. Dat is nu juist het problematische: we nemen de zoekmachines voor lief, terwijl het steeds minder zichtbaar is hoe ze hun werk doen. De witte, neutrale uitstraling van Google versterkt dat alleen maar. Net als de razende snelheid waarmee zoekresultaten binnen milliseconden op je scherm komen, alsof het doorzoeken van het web geen enkele moeite kost. Zoals de elektriciteit die uit het stopcontact komt, of straten die samen een wegennet vormen, lijken zoekmachines een infrastructuur. Je merkt het pas als het niet meer werkt, als de weg is opgebroken of het internet door een serverstoring niet toegankelijk is (iets wat Google zeer zelden overkomt). Zoekmachines verdienen echter een aandachtige blik, juist omdat ze de voornaamste ingang tot kennis zijn en zo een grote invloed hebben op welke kennis we tot ons nemen en daarom ook op hoe je naar de wereld kijkt.





We gebruiken zoekmachines zoals Google of Bing (of YouTube of Twitter) alsof het neutrale, objectieve tools zijn. Terwijl er steeds meer duidelijk wordt over de filterende en selecterende functie ervan, met als achterliggend doel zoveel mogelijk winst te genereren. Het verzamelen van data staat in dienst van commerciële belangen, zoveel is zeker. Maar hoe het precies werkt en welke algoritmes Google hanteert om dit voor elkaar te krijgen, blijft geheim. Wel zijn we ons steeds meer bewust van de sturende rol die Google heeft in wat we online vinden en dus lezen of bekijken. (Die rol gaat overigens steeds verder door ook te sturen wat we zoeken, via tools zoals autocomplete die je zoekopdracht aanvullen nog voor je hem zelf tot het eind hebt kunnen formuleren.) Het concept van de 'filterbubbel' zoals beschreven door Eli Pariser is inmiddels redelijk bekend: door slimme algoritmes en het verzamelen van data over onder meer je zoekgeschiedenis en locatie, krijgt de gebruiker van een zoekmachine op maat gesneden resultaten te zien. Handig toch? De andere kant van het verhaal is dat je daardoor verstoken blijft van informatie die buiten je 'filterbubbel' valt. Op den duur zou dat betekenen dat je kennis en perspectief beperkt en onvolledig blijft. Bijvoorbeeld: zoeken op 'BP' levert voor de één een pagina op met resultaten over aandelenkoersen in de oliemaatschappij, voor de ander links naar nieuws over de milieuramp die in de Golf van Mexico plaatsvond.

Zoekmachines en de manier waarop ze werken mogen haast onzichtbaar zijn, ze zijn wel alomtegenwoordig. Dat is goed te zien op de wereldkaart 'Age of Internet Empires' van Mark Graham en Stefano De Sabbata. Die toont de verdeling van de meest bezochte websites per land. Conclusie: de helft van de internetters (bij elkaar 1 miljard mensen) woont in een land waar Google de meest bezochte website is, terwijl de Chinese zoekmachine Baidu nog eens een half miljard mensen bereikt. Dat is bij elkaar opgeteld heel wat meer dan sociaal medium Facebook, dat weliswaar in vijftig landen de grootste is, maar die bij elkaar 'slechts' 280 miljoen mensen tellen. Als voor zoveel mensen een zoekmachine de ingang van het web vormt, waarom weten we er dan niet meer over? Is dat niet vreemd, zeker in een omgeving als het hoger onderwijs? Wie begint zijn zoektocht naar kennis nog ergens anders dan op Google? Wat zou je immers zonder moeten doen?





<http://geography.oii.ox.ac.uk/?page=age-of-internet-empires>

We hopen dat de volgende artikelen samen een beeld geven van wat er meespeelt als je de belangrijkste toegangspoort tot het internet gebruikt, van de problematische aspecten van online zoeken en de nuances die te vaak verborgen blijven. Natuurlijk wil niemand terug naar de tijd voor Google, het gemak waarmee zoveel informatie ontsloten wordt is ongekend en van onschatbare waarde. Maar dat neemt niet weg dat we vragen moeten stellen over zoeken – aan experts en aan elkaar.

Miriam Rasch, Instituut voor Netwerkcultuur
Amsterdam, mei 2014

Volg de volgende link voor meer informatie of voor het bestellen of downloaden van de Society of the Query Reader: Reflections on Web Search, René König en Miriam Rasch (eds), Amsterdam: Institute of Network Cultures, 2014, (gratis).

<http://www.networkcultures.org/query>

Google: een database van intenties?

Kylie Jarrett



We zoeken elke dag wel iets op op Google. Met een paar zoektermen vinden we al snel wat we zoeken, bijvoorbeeld een routebeschrijving, vliegtickets of recept. Bij dat soort vragen hebben we een specifiek doel voor ogen, ofwel een intentie bij ons handelen, namelijk een route afleggen, een vakantie boeken, of een gerecht bereiden. Je zou kunnen stellen dat Google door het verzamelen van al die zoekopdrachten, én van de bijbehorende IP-adressen, locaties en andere profielmatige data een 'database van intenties' opbouwt. Met die database kan Google onze intenties niet alleen archiveren en traceren, maar ook voorspellen en uitbuiten. Maar is de definitie 'database van intenties' wel accuraat en gepast? Wat zijn intenties precies? Hoe verwerkt Google ze in zijn systeem, en wat zijn hier de gevolgen van?



Wat zijn intenties?

Intentie heeft te maken met iets wat je wilt. Als je handelt vanuit een bepaalde intentie betekent dat je een bepaald doel voor ogen hebt. Kylie Jarrett, verbonden aan de National University of Ireland, draagt in haar artikel 'A Database of Intention?' meerdere argumenten aan om de definitie 'database van intenties' niet zomaar voor lief te nemen. Google kan onze intenties nooit geheel vastleggen in een database, omdat ze altijd verbonden zijn aan affect en emotie. Het gebruik van bepaalde zoektermen zegt wel iets over ons, maar ze weerspiegelen nooit alle achterliggende verlangens en emoties. Google kan dus wel (uiterlijk) gedrag in kaart brengen, maar niet een compleet inzicht krijgen in de (innerlijke) motivatie voor een zoekopdracht.

Google autocomplete belemmert de mogelijkheid om zelfstandig en onafhankelijk te zoeken

Ter illustratie beschrijft Jarrett hoe zij maandenlang werd geteisterd door specifieke advertenties die verschenen in haar Gmail, nadat zij op 'abortus' had gezocht. Niet met de intentie om een abortuskliniek te vinden om abortus te plegen, maar juist om Google zelf te onderzoeken (je moet weten dat op abortus in Ierland een gevangenisstraf kan staan van veertien jaar!). Dat had Google niet in de gaten. Het bedrijf koppelt intenties, doelstellingen, aan zoektermen die van te voren zijn gedefinieerd door adverteerders om zoveel mogelijk advertenties te verkopen, ook al komen deze niet overeen met de werkelijke intenties van de gebruikers.

Noch Google, noch de adverteerders, kunnen aan de ingevoerde zoektermen zien wat de inherente verlangens, emoties en motivaties van een persoon zijn. Oftewel, de algoritmes van Google kunnen emotionele of lichamelijke drijfveren niet begrijpen en zinvol opslaan. Dat besef geeft troost, vindt Jarrett. Google moet niet worden begrepen als een database van intenties, maar als een database van onnauwkeurige en incomplete intenties die Google vervolgens weer één-op-één aan ons toeschrijft.





Vervreemding en onteigening

Dat is niet onschuldig. Volgens Jarrett is het belangrijk om je te realiseren dat bij zoekopdrachten op Google onze intenties worden vervreemd en onteigend. Wat betekent dat? Intenties komen in eerste instantie voort uit een 'geleefde' en 'belichaamde' wens. Met andere woorden, ze ontstaan in ons lichaam, tezamen met een bepaald gevoel of een meer of minder bewuste gedachte, in relatie tot persoonlijke handelingen of ervaringen. Wanneer we een zoekterm intypen, verandert die geleefde, affectieve intentie in een duidelijk gedefinieerde en door Google geïnterpreteerde intentie. Ze wordt in feite van je afgepakt, waardoor vervreemding en onteigening optreden. Je hebt niet meer de controle over je eigen intenties – noch inhoudelijk, noch in bezitsmatige zin. De intenties worden omgezet in 'dode' advertentiedata die 'gecommodificeerd', oftewel te gelde gemaakt kan worden. Levende intenties veranderen in data op de servers van Google, verhandeld en omgezet in geld uit de verkoop van advertenties. Jarrett ziet dit als een vorm van uitbuiting, omdat onze energie, gedachten en activiteiten worden getransformeerd tot iets waar alleen Google financieel op vooruit gaat. Deze manier van werken, waarbij commerciële bedrijven data van internetgebruikers vervreemden en onteigenen om vervolgens uit te buiten en te controleren, is volgens Jarrett symptomatisch voor het hedendaags kapitalisme.

Autocomplete: het voorspellen van intenties

Met de Autocomplete-functie doet Google suggesties voor zoektermen, op basis van ingetypte letters en/of woorden. Jarrett merkt op dat de suggesties die uit een database komen invloed kunnen hebben op de manier waarop we onze intenties verwoorden bij online zoekopdrachten. Denk alleen al aan de aanvullingen als je zoekt naar een café in een jou onbekende stad. De kans is groot dat je doorklikt naar een van de zoeksuggesties die verschijnen. In het artikel 'Wat Google autocomplete zegt over wie we zijn' van M.E. Luka en Mél Hogan zie je wat minder onschuldige voorbeelden. Soms laten we zelfs Google onze eigen wensen verwoorden, als we klikken op een suggestie. Lekker makkelijk? Al typende laten we ons leiden door een commercieel geproduceerd algoritme en eerder ingevoerde zoekopdrachten van de meerderheid. Volgens Jarrett belemmert dit de creativiteit en de mogelijkheid om vrij, zelfstandig en onafhankelijk na te denken bij een zoekopdracht.





Iedereen heeft een data-dubbelganger

Google leest niet alleen intenties af aan ons zoekgedrag om die technisch te verwerken in een algoritme, maar koppelt ze ook aan personen. Alle data die de algoritmes van Google over iemand verzamelen, vormen samen een 'algoritmische identiteit'. Dit is als het ware je 'data-dubbelganger', ontstaan en gevormd door gebruik van de producten van Google. Terwijl Google nooit met zekerheid de identiteit van een gebruiker kan achterhalen, is het systeem gebouwd om intenties aan je echte identiteit toe te schrijven – waarmee vervolgens nog gerichter advertenties kunnen worden verkocht en geld verdiend.

Op basis van al die informatie doet Google bepaalde aannames over wie we zijn, waar we zijn, en wat we willen. Maar kloppen die aannames wel? Intenties kunnen verkeerd worden geïnterpreteerd. Dit gebeurde bijvoorbeeld in Amerika, nadat een onschuldig koppel op Google had gezocht naar rugzakken en snelkookpannen. Zij werden onprettig verrast door FBI-agenten aan de deur, die hen ervan verdachten een bom te willen bouwen. Je kan als Google-gebruiker moeilijk afstand kan doen van je data-dubbelganger, en dat kan wel eens vervelend uitpakken, zeker nu gebleken is dat geheime diensten als de NSA gebruik maken van Googles databases en algoritmes. Jarrett roept op tot verzet en vraagt zich af of er alternatieve systemen zijn die beter kunnen voorzien in onze behoeften. Is het mogelijk om een database aan te leggen van werkelijke intenties? Zouden we dat wel willen? Zijn er technische, sociale en economische modellen te bedenken die de gebruiker van een zoekmachine níet reduceren tot data en die de gebruiker controle en zeggenschap geeft over wat er met zijn data gebeurt?

Lees het hele artikel van Kylie Jarrett in de *Society of the Query Reader*:

http://networkcultures.org/query/wp-content/uploads/sites/4/2014/06/1.Kylie_Jarrett.pdf

A Database of Intention? (pdf).





Discussiepunten:

1: Zoeken met Google levert het bedrijf via advertentieverkopen geld op. Kylie Jarrett is geen voorstander van de toenemende economische en kapitalistische macht van Google. Maar de zoekmachine zou je natuurlijk ook vanuit een heel ander perspectief kunnen beschouwen: het is een goede, nuttige dienst die ons helpt in onze zoektocht naar informatie. Hoe erg vind je de toenemende macht van Google, afgezet tegen de frequentie van gebruik van deze handige dienst?

2: Google investeert in nieuwe producten en diensten. Onlangs werd bekend dat Google een bedrijf heeft overgenomen dat militaire robots bouwt. Aangezien zoekopdrachten inkomsten voor Google opleveren, zou je kunnen stellen dat je door te zoeken indirect het bouwen van dodelijke militaire robots steunt. Wat vind je hiervan?

<http://www.nbcnews.com/technology/google-buys-major-military-robot-maker-why-does-search-giant-2D11744237>

3: Opdracht: Maak jij vaak gebruik van de suggesties die Google met de Autocomplete-functie doet? Sta eens stil bij het feit dat zoeksuggesties onze gedachten beïnvloeden. Via onderstaande link kan je zoeken op Google zonder suggesties. Zie je daar de voordelen van in?

<http://www.bbc.com/future/story/20131106-is-google-autocomplete-evil>
<https://www.google.com/webhp?complete=0>

4: DuckDuckGo en Ixquick zijn voorbeelden van alternatieve privacy-vriendelijke zoekmachines die je zoekopdrachten niet opslaan en verwerken tot advertentieverkopen. Kende je deze alternatieve zoekmachines? Waarom zou je ze wel/niet willen gebruiken?

<https://duckduckgo.com/>
<https://www.ixquick.com/>





Verder lezen en kijken:

De lezing die Kylie Jarrett gaf op Society of the Query #2 is op het blog beschreven en terug te zien onder de volgende links:

<http://networkcultures.org/wpmu/query/2013/11/11/kylie-jarrett-search-for-the-google-god-metaphysics-and-the-social-imaginary-of-search/>

<http://vimeo.com/79880130>

Kylie Jarrett – Search for the Google God: Metaphysics and the Social Imaginary of Search', Institute of Network Cultures, 11 november 2013.

+ Stephan Fellingner, 'Google is god', 1 juni 2012.

<http://fellingner.nl/2012/06/01/google-is-god/>

+ Frederiek Pennink, 'Book Review: Google and the Culture of Search (2013)', Institute of Network Cultures, 28 mei 2013.

<http://networkcultures.org/wpmu/query/2013/05/28/book-review-google-and-the-culture-of-search-2013/>

+ Joep Hutschemakers, 'De rijkheid van intentiedata: scoren in de laatste minuut', 16 juli 2013.

<http://www.marketingfacts.nl/berichten/de-rijkheid-van-intentiedata-scoren-in-de-laatste-minuut>

+ Anna Jobin, 'Google's Autocompletion: Algorithms, Stereotypes and Accountability', 22 oktober 2013.

<http://sociostrategy.com/2013/googles-autocompletion-algorithms-stereotypes-accountability/>

+ Douwe Schmidt en Maurits Martijn, 'Op zoek naar je digitale alter ego'. De Correspondent, 28 oktober 2013.

<https://decorrespondent.nl/231/op-zoek-naar-je-digitale-alter-ego/7104636-30f9c8e1>



Google tussen standaardisering en individualisering

Andrea Miconi



Afbeelding 1: Googling for Obama (Bron: <http://thefilterbubble.com>)

Kun je je nog een leven voorstellen zonder Google? Waarschijnlijk niet. Ik googel, jij googelt, haast iedereen googelt. Het is zo vanzelfsprekend om iets op Google op te zoeken dat de zoekmachine niet meer weg te denken is uit onze hedendaagse cultuur. Sterker nog, je zou Google kunnen zien als een technologie die onze cultuur ordent en zo onze dagelijkse activiteiten beïnvloedt. De zoekresultaten die de algoritmes leveren, hebben invloed op wat wij wel en niet lezen, tot welke informatie wij toegang hebben, welke bronnen we raadplegen. Hoewel Google al meer dan tien jaar deze omvangrijke functie heeft, weten we nog steeds niet hoe die algoritmes precies werken. Is dat niet vreemd?

Andrea Miconi, verbonden aan de Universiteit van Milaan, maakt zich zorgen over de manier waarop Google onze cultuur vormt. Hij beschrijft in het artikel 'Dialectic of Google' de groeiende macht van Google als een groot probleem dat maar weinig mensen onderkennen. De meeste gebruikers ervaren Google als betrouwbare bron voor objectieve informatie, een handige, neutrale, gebruiksvriendelijke technologie die bovendien gratis is. Het vertrouwen in de zoekmachine is groot, zoals ook het marktaandeel laat zien: in Nederland ligt dat boven de 95%. De witte, simpele interface van Google oogt neutraal, maar in de praktijk zijn de resultaten die Google levert juist verre van neutraal. De strategieën die de zoekmachine hanteert leiden eerder tot een beperkte en vooringenomen representatie van de werkelijkheid. Miconi noemt dit een 'bias': een neiging om een beperkt perspectief op een zaak te hebben én dat beperkte perspectief te willen houden (zie <http://en.wikipedia.org/wiki/Bias>).



Google als gatekeeper

De zoekresultaten die Google levert, sturen ons beeld van de werkelijkheid, terwijl de manier waarop de zoekresultaten tot stand komen ontransparant blijft. Dat is een verschil met traditionele media, die je welbewust opzoekt met het besef dat ze niet neutraal zijn. Google wordt echter als objectieve instantie en niet als 'gatekeeper' gezien, terwijl het bedrijf volgens Miconi juist wel die rol heeft. 'Gatekeeping' wordt eerder geassocieerd met massamedia zoals televisie en kranten, en heeft te maken met het actief selecteren en filteren van informatie. Terwijl we bij televisieprogramma's en nieuwsberichten in de krant beseffen dat dit gebeurt, accepteren we zonder argwaan de zoekresultaten van Google. We vragen ons niet af hoe de selectie en rangschikking van zoekresultaten tot stand komt, terwijl er bij de zoekmachine juist erg veel achter de schermen gebeurt.

De pretentie neutraal te zijn valt volgens Miconi onder Googles strategie. Hij beschrijft ook twee technologische instrumenten van Google: PageRank en cookies. PageRank, een van de vele algoritmes van Google, rangschikt zoekresultaten op basis van het aantal links dat naar die specifieke website voert. Hoe meer links naar een website, des te hoger die komt te staan in de lijst met zoekresultaten. PageRank is een standaard algoritme dat bij alle zoekresultaten wordt gebruikt.

De zoekmachinetechnologie berust op twee tegengestelde principes: individualisering en standaardisering

De tweede strategische stap, het aanpassen van de zoekresultaten door cookies te gebruiken, ziet Miconi daarentegen als een vorm van individualisering. Je zoekgeschiedenis, interesses en voorkeuren worden opgeslagen door middel van cookies, die vervolgens worden ingezet om gepersonaliseerde zoekresultaten te leveren. Dit heet ook wel de 'filterbubbel' – niet iedereen krijgt exact dezelfde resultaten, want die zijn óók toegespitst op jouw persoonlijke profiel. Gebruikers van Google zijn zich over het algemeen niet bewust dat Google niet alle resultaten toont, oftewel informatie achterhoudt. Kun je aan de 'bubbel' ontkomen? Miconi vreest van niet. Hij beschrijft de bubbel als een onzichtbare kooi van informatie, die op ondoorgroendelijke wijze bepaalt welke informatie we wel en niet te zien krijgen.





Twée tendensen, één technologie

Dialectiek (<http://www.marxists.org/nederlands/woordenboek/d/dialectiek.htm>) is een filosofische methode die de werkelijkheid niet statisch maar dynamisch opvat, ook wel gekenmerkt als een geheel van tegenstellingen. 'Dialectiek van Google' zoals Miconi dat bedoelt, verwijst naar de dynamische manier waarop Google werkelijkheid en wereldbeeld beïnvloedt. Die zijn onderhevig aan de sturende werking van innovaties van Google, zoals de lancering van nieuwe producten en diensten als Google Now, en aan het aanpassen van algoritmes. Dialectiek verwijst ook naar de hierboven genoemde tegengestelde tendensen die de zoektechnologie kenmerken. Enerzijds is er sprake van standaardisering door PageRank, anderzijds van individualisering door het aanpassen van zoekresultaten op individueel niveau.

De zoekmachinetechnologie berust dus op twee tegengestelde principes: individualisering en standaardisering. Er is volgens hem sprake van een technologische versie van wat hij beschrijft als het meest traditionele sociologische probleem: de spanning tussen individu en maatschappelijke structuur en de opkomst van een nieuwe orde die de relatie tussen beide reguleert. Vanuit dit perspectief is Google de nieuwe orde die overal een vinger in de pap heeft en met haar technologie zowel de wereldwijde informatiemaatschappij als het individu beïnvloedt. Google voorziet iedere individuele gebruiker van zijn eigen filterbubbel en past door middel van cookies ieders zoekresultaten aan op eerder zoekgedrag. Tegelijkertijd is het PageRank algoritme de standaard geworden waarmee informatiebronnen in zoekresultaten wereldwijd worden gerangschikt. Individualisering en standaardisering werken elkaar hierbij niet tegen; ze worden gecombineerd in Google's zoektechnologie.

Lees het hele artikel van Andrea Miconi in de *Society of the Query Reader*:

http://networkcultures.org/query/wp-content/uploads/sites/4/2014/06/2.Andrea_Miconi.pdf

Dialectic of Google (pdf).





Discussiepunten:

1: Miconi stelt dat Google een soort toets is voor wat echt is of wat bestaat. Herken je dat? Toets jij vaak feiten of gebeurtenissen aan de hand van de resultaten die Google geeft?

2: Beschouw je de macht van Google als een maatschappelijk probleem?

3: Opdracht: gebruik met meerdere mensen Google om te zoeken naar manieren om de filterbubbel van Google te omzeilen. Gebruik dezelfde zoektermen en vergelijk de zoekresultaten. In hoeverre verschillen de zoekresultaten?

<https://vimeo.com/79880130>

Kylie Jarrett, Search for the Google God: Metaphysics and the Social Imaginary of Search

4: Kylie Jarrett had het in haar haar lezing bij de tweede Society of the Query-conferentie ook over de invloed van Google op onze realiteit. Wat zijn haar stellingen hierover en komen deze overeen met die van Miconi?

Verder lezen en kijken:

- + Consumentenbond, 'Wat Google ons niet laat zien', Digitaalgids januari/februari 2014.
http://www.consumentenbond.nl/morello-bestanden/gidsarchief/digitaalgids/2013/dg201301p28_Focus_het_Google-filter.pdf

- + Andrea Miconi, 'The Big G and the Never-Ending Story of Knowledge Monopolies' (pdf), Mediascapes Journal, n. 1 (2013), 158-160.
<http://ojs.uniroma1.it/index.php/mediascapes/article/view/10265/10169>

- + Eli Pariser, 'The Filter Bubble. What the Internet Is Hiding From You', New York: Penguin. 2011,
http://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles/





Wat Google autocomplete zegt over wie we zijn

Mél Hogan & M.E. Luka 

Zodra je een zoekterm intypt in Google komen er zoeksuggesties in beeld. Deze functie staat ook wel bekend als 'Autocomplete' of 'Google Suggest' (zie ook het artikel van Kylie Jarrett). De suggesties zijn gebaseerd op miljarden andere zoekopdrachten. Maar dat is het niet het enige. Verschillende gebruikers krijgen namelijk verschillende suggesties, omdat onder meer factoren als locatie en taal worden meegenomen (zie voor meer daarover het artikel van Jobin en Glassey). Bovendien worden sommige zoekopdrachten helemaal niet getoond in de zoeksuggesties. Volgens de servicevoorwaarden van Google worden zoektermen die te maken hebben met nationaliteit, haat of geweld, persoonlijk identificeerbare informatie, pornografie of verwijderde content niet opgenomen in de suggesties. Maar is dat wel zo?

Mél Hogan, postdoctoraal onderzoeker Digital Curation aan de Universiteit van Colorado, laat zien dat dit niet helemaal klopt. Google doet volgens Hogan namelijk wel degelijk suggesties die raken aan geweld. In de tabel hieronder zie je de suggesties van Google bij het zoeken naar 'woman is', waaruit blijkt dat Google onder andere voorstelde om te zoeken naar 'vrouw is onthoofd'.

2013
woman is: the future of man that meant to insult me a woman the earth the reflection of her man a danger cat too pretty to work the weaker vessel a devil beheaded

Tabel 1: Woman is 2013



De zoeksuggesties zijn soms hilarisch, bijvoorbeeld 'Dude is a hair on an elephant' maar Hogan legt de nadruk op de bedroevende, zielige en vulgaire zoeksuggesties die wijzen op de duistere kant van de mens en waar we naar zoeken als we ons onbespied wanen. Het schetst een beeld van dat waar we ons zorgen over maken, zoals bijvoorbeeld dochters met overgewicht (http://www.slate.com/blogs/xx_factor/2014/01/21/parents_ask_google_is_my_s_on_gifted_and_is_my_daughter_overweight.html).

Hogans bevinding is dat vooral de woorden die een relatie aangeven, zoals daughter of boyfriend (zie Tabel 2), leiden tot de verschrikkelijkste zoeksuggesties.

2010	2011	2012	2013
daughter is: overweight depressed moving away pregnant a bully mean tired all the time calling ringtones losing hair a prostitute	daughter is: mean pregnant always cold constipated a bully a loner cutting herself tired all the time overweight calling ringtone	daughter is: pregnant depressed cutting herself <u>spanish</u> pregnant with dads child constipated mean out of control disrespectful to mother a brat	daughter is: pregnant calling ringtone <u>spanish</u> an atheist <u>sexting</u> getting married depressed a brat taller than me a tomboy
boyfriend is: depressed distant ignoring me selfish insecure a virgin moving away <u>a jerk</u> <u>too big</u>	boyfriend is: distant depressed selfish sick ignoring me cheap too big gay boring <u>mad at me</u>	boyfriend is: distant selfish a virgin boring depressed a jerk sick controlling <u>too clingy</u> <u>insecure</u>	boyfriend is: depressed immature a virgin <u>a douchebag</u> boring ignoring me moving away <u>selfish</u> <u>distant</u> gay

Tabel 2: Daughter and Boyfriend is 2010-2013

Zoektermen zoals 'love' of 'death' verwijzen daarentegen vaak naar songteksten, boeken of films. Het verzamelen van de zoeksuggesties is in die zin ook een manier om populaire cultuur te documenteren.



Vervuild en voorspelbaar

De screenshots mogen door de jaren heen iets zeggen over onze gedeelde emoties, zorgen en populaire cultuur, toch waarschuwen Hogan en Luka tegen het trekken van al te sterke conclusies. De zoeksuggesties zijn volgens hen namelijk afkomstig uit een 'vervuild en voorspelbaar' archief. Zij verwijzen hiermee naar een citaat van Geert Lovink:

Vergeet nooit dat Google informatie verzamelt met een commercieel doel. Het is geen openbaar archief. Daarnaast wordt Google's zoekmachine steeds meer 'vervuild', wat leidt tot nutteloze en voorspelbare zoekresultaten.
<http://networkcultures.org/wpmu/geert/2010/06/26/back-from-gent-notes-on-memories-of-the-future>

De toenemende vervuiling van Google houdt in dat er steeds meer onbruikbare links tussen de zoekresultaten zitten. Hoewel Google haar algoritmes geheim houdt, waardoor we niet weten precies weten hoe de zoekresultaten en suggesties precies tot stand komen, is het wel duidelijk dat commerciële drijfveren de resultaten beïnvloeden en dat niet iedereen dezelfde resultaten en suggesties krijgt. De screenshots die Hogan verzamelt weerspiegelen daarom niet één-op-één het collectieve zoekgedrag van Google-gebruikers. Eerder geven zij een impressie.

De zoeksuggesties veranderen door de jaren heen; de oude worden vervangen door nieuwe. De screenshots van Hogan zijn vanuit dit vergankelijke perspectief waardevolle momentopnames. Stille beelden van algoritmes, databases en emoties die altijd in beweging zijn.

Lees het hele artikel van Mél Hogan en M.E. Luka in de *Society of the Query Reader*

http://networkcultures.org/query/wp-content/uploads/sites/4/2014/06/18.Hogan_Luka.pdf

Polluted and Predictive, in 133 Words (pdf).





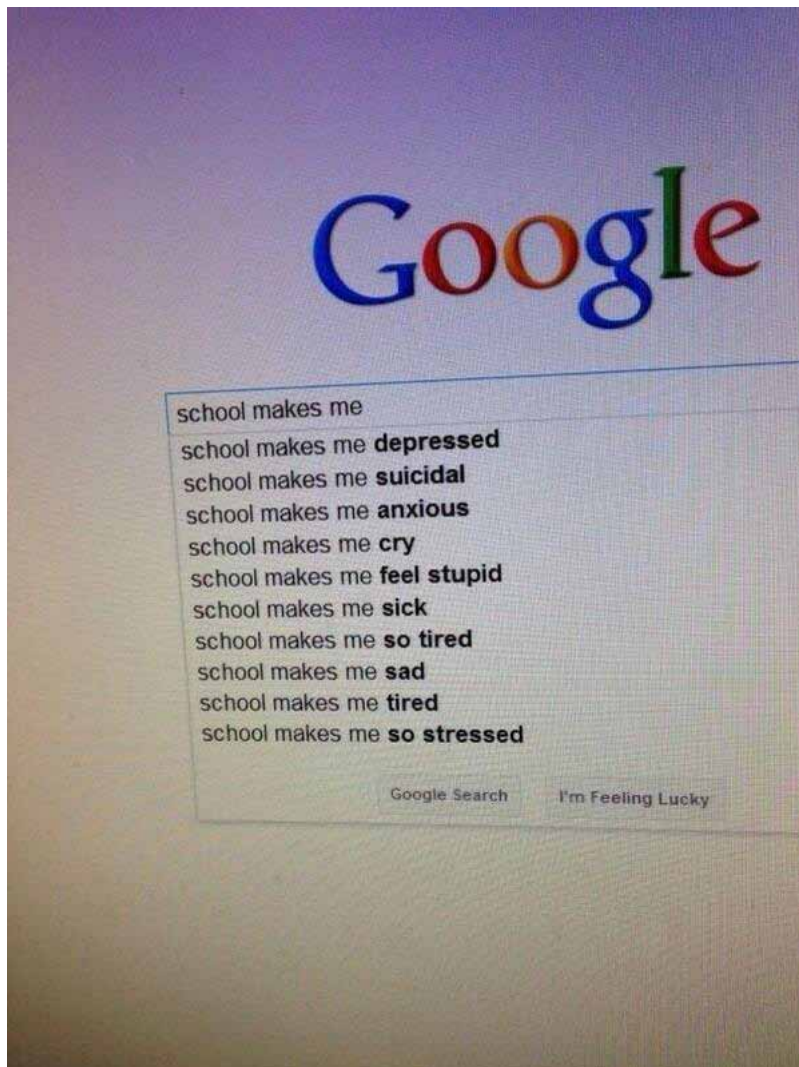
Discussiepunten:

1: Opdracht: vertaal een aantal zoektermen die Hogan heeft gebruikt, bijvoorbeeld 'dochter' en voer deze in bij Google.nl, aangevuld met 'is' en een spatie. Bespreek met elkaar de zoeksuggesties die Google levert. Welke vind je het meest opmerkelijk en waarom?

2: Opdracht: bekijk de website Google Poëzie. Volg de instructies, maak een gedicht en maak er een screenshot van.

<http://nederlands.googlepoetics.com/>

<http://nederlands.googlepoetics.com/post/45261569093/hoer>



Afbeelding 1: #sadgooglesearches

(<http://memetrip.tumblr.com/post/58300727010/sadgooglesearches>)





3: Zoek op Twitter of Google naar #sadgooglesearches. Wat vind je? Wat is #sadgooglesearches en hoe verhoudt dit zich tot het kunstproject van Hogan?

Verder lezen:

- +** Claire Cullen, 'Hilarious Hashtag "Sad Google Searches" Trends on Twitter', 15 augustus 2013.
<http://www.independent.ie/blog/hilarious-hashtag-sad-google-searches-trends-on-twitter-29501757.html>

- +** Nicholas Diakopoulos, 'Sex, Violence, and Autocomplete Algorithms: What Words do Bing and Google Censor from Their Suggestions?' 2 augustus 2013.
http://www.slate.com/articles/technology/future_tense/2013/08/words_banned_from_bing_and_google_s_autocomplete_algorithms.html

- +** Jalees Rehman, 'Google Zeitgeist: Annoying Philosophers, Weird Germans and White Pakistanis', 9 december 2013.
<http://www.3quarksdaily.com/3quarksdaily/2013/12/google-zeitgeist-annoying-philosophers-weird-germans-and-white-pakistanis.html>



Zoeken in een zoekmachine die steeds verandert

Dave Crusoe



Afbeelding 1: The + operator has been replaced

(http://www.wired.com/images_blogs/business/2011/10/google-beta02-f.jpg)

Hoe heb jij leren zoeken op het web? Op school, of heb je het jezelf aangeleerd? Het is goed mogelijk dat je ooit hebt geleerd dat je bij zoekopdrachten met Google een plusteken kunt gebruiken om zoektermen aan elkaar te koppelen en om zo webpagina's te vinden waarop de exacte frase voorkomt. Maar sinds oktober 2011 heeft het plusteken deze functie niet meer. Het plaatsen van een plus voor aanvullende zoektermen is door Google vervangen door het gebruik van dubbele aanhalingstekens (zie Afbeelding 1), wat sommigen als nadelig ervaren(<http://www.wired.com/business/2011/10/google-kills-its-other-plus-and-how-to-bring-it-back/>). Dubbele aanhalingstekens kosten meer moeite; één extra handeling per ingevoerde aanvullende zoekterm. De functie van het plusteken bestond al 12 jaar en was plotseling verdwenen.

Google brengt vaak veranderingen aan in de gebruikte zoekalgoritmes (zie http://seopalbg.com/wp-content/uploads/2012/02/Google_Algorithm_Changes_02-011.png). Die veranderingen zijn haast onzichtbaar, waardoor je niet snel stil staat bij de veranderlijkheid van het grootste webportaal dat we kennen. Welke gevolgen heeft dat voor gebruikers en de manier waarop we leren om met zoekmachines om te gaan, bijvoorbeeld in het onderwijs?



Boolean operators

In het essay 'Educating for Search' geeft David Crusoe, medeoprichter van het Public Learning Media Laboratory (<http://www.plml.org/>), richtlijnen die beter onderwijs in het gebruik van zoekmachines moeten bevorderen. Hij observeerde een aantal lessen in online zoeken op Amerikaanse middelbare scholen en constateerde dat het onderwijs is blijven hangen in de tijd waarin zoekmachines vooral opereerden op basis van de zogeheten booleaanse logica (http://nl.wikipedia.org/wiki/Booleaanse_algebra). Bij de booleaanse logica, die met name wordt toegepast op doorzoekbare databases van bibliotheken, kun je de operators AND, NOT en OR gebruiken of de tekens '+', '-' en '"', om specifieke zoektermen te combineren of juist uit te sluiten.

Hoewel Google de functie van het plusteken in 2011 verwijderde, zijn de booleaanse operators toch gedeeltelijk nog van toepassing. Een minteken sluit nog steeds zoektermen bij een zoekopdracht uit en met OR kun je zoeken naar webpagina's met een van beide zoektermen, zoals te lezen is op de 'Zoekoperators' pagina (<https://support.google.com/websearch/answer/136861?hl=n>) van Google.

Google kan de zoekmachine aanpassen zonder dat veel gebruikers dit doorhebben

Toch is kennis van de booleaanse logica in steeds mindere mate van belang om te leren omgaan met zoekmachines, aldus Crusoe. Belangrijker zijn kennis van de werkwijze van Google en algemene vaardigheden op gebied van informatiewijsheid, in combinatie met de vele functies die Google biedt. Crusoe geeft vier praktische tips ten behoeve van het gebruik van zoekmachines:

- 1: Leer hoe zoekmachines werken; besef dat zoekmachines maar een deel van het web indexeren.
- 2: Gebruik woorden en zinnen waarvan je mag verwachten dat ze veel voorkomen op de websites die je zoekt.
- 3: Wees geduldig; de eerste zoekopdracht of resultatenpagina leidt niet per se tot het beste resultaat.
- 4: Leer goede, bruikbare en geschikte bronnen te onderscheiden van slechte bronnen.





Veranderlijkheid, signalen en intenties

Google is geen stabiele technologie die altijd hetzelfde blijft, maar ontwikkelt zich razendsnel door. Het bedrijf kan de zoekmachine aanpassen, zonder dat veel gebruikers dit door hebben, zoals gebeurde bij de functie van het plusteken. Bovendien worden Googles zoekresultaten constant gefilterd en gerangschikt, waardoor geen enkele zoekopdracht precies dezelfde resultaten zal opleveren. Google heeft de afgelopen jaren de algoritmes aangepast met het doel zoveel mogelijk spam te voorkomen en zo snel mogelijk 'verse' of actuele resultaten te bieden. Met die veranderlijkheid van Google en de resultaten zou in het onderwijs meer rekening gehouden moeten worden.

Iedereen zou er veel aan hebben om te leren dat een 'query', een zoekopdracht, veel meer omvat dan slechts een woord of een combinatie van woorden, ingevoerd in een zoekbalk. De resultaten zijn namelijk onderhevig aan alle 'signalen' die Google ontvangt over de zoeker, nu en vanuit het verleden. Denk hierbij bijvoorbeeld aan locatie, tijdstip en zoekgeschiedenis. Crusoe legt uit dat Google in toenemende mate signalen van haar gebruikers interpreteert om zo te achterhalen wat de gebruikers echt willen weten, wat hun intentie is. De signalen vormen samen een portret van de gebruiker en zijn van grote invloed op de zoekresultaten. Een query kunnen we het beste beschouwen als het totaal van alle signalen die de gebruiker naar de zoekmachine stuurt, aldus Crusoe. Maar zijn onze intenties wel goed af te lezen van de signalen die we naar Google sturen? Meer hierover kan je lezen in het artikel van Kylie Jarrett: Database van intenties.

Lees het hele artikel van Dave Crusoe in de *Society of the Query Reader*

http://networkcultures.org/query/wp-content/uploads/sites/4/2014/06/16.Dave_Crusoe.pdf

Educating for Search: Understanding the Past and Present Search Technology to Teach for Future Resilience (pdf).





Discussiepunten:

1: Je kunt personalisatie op basis van je zoekgeschiedenis uitschakelen via onderstaande link. Hiermee voorkom je dat je voormalige zoekgeschiedenis een signaal is dat wordt verwerkt in je zoekopdracht. Wist je dat dit mogelijk was? Welke voor- en/of nadelen zie je bij het personaliseren van zoekresultaten op basis van je zoekgeschiedenis?

<https://support.google.com/accounts/answer/54048?hl=en%20>

2: Opdracht: Ga naar de Zoekoperators pagina van Google. Hoeveel van de genoemde functies op die pagina kende je al? Experimenteer met een aantal functies en bespreek vervolgens met elkaar welke je nuttig vindt en welke niet.

<https://support.google.com/websearch/answer/136861?hl=nl>

3: Volgens Crusoe is het de verantwoordelijkheid van het onderwijs om studenten in aanraking te brengen met zoekmachinefuncties en praktische vaardigheden. Ben je dit met hem eens?

4: Bekijk de tips van Crusoe en probeer voor jezelf de volgende vragen te beantwoorden. Hoe komt het dat Google niet het hele web indexeert? Welke veel voorkomende frasen en zinnen vind je handig om te gebruiken als je met Google zoekt? Hoe geduldig ben jij als je online zoekt? Hoe onderscheid je goede van slechte bronnen?

5: Opdracht: Het Public Learning Media Laboratory, mede opgericht door Crusoe, heeft lesmateriaal voor groepen studenten ontwikkeld over zoekmachines en het formuleren van zoekopdrachten door middel van booleaanse operators: Boolify Lesson: Introduction to Web Search. Ga hiermee aan de slag:


https://docs.google.com/document/d/1ebthehQ0ETWzmonLDNe-ul6WioTvMJJb_bB87CNI5Q4g/edit


<http://www.plml.org> (Lesson last updated: April 22, 2013).







Verder lezen:

-  Andry Baio, 'Google Kills Its Other Plus, and How to Bring It Back', Wired, 25 oktober 2011.
<http://www.wired.com/business/2011/10/google-kills-its-other-plus-and-how-to-bring-it-back/>

-  Patrick Mackaaij, 'Zoeken leren (googelen) met Power Searching with Google', 13 september 2013.
<http://www.eenmanierom.nl/power-searching-with-google/>

-  Melanie Pinola, 'Google Tips And Tricks Every Student Should Know', Life-hacker, 25 januari 2014.
<http://www.lifehacker.com.au/2014/01/google-tips-and-tricks-every-student-should-know/>

-  Maarten Sprenger, *Slim zoeken op internet. Voor thuis, op school en waar je ook bent*, De wereld van bovenaf, september 2013.





Op zoek naar de voorouders van Google

Anton Tantner



Een sollicitant stuurt een mail naar een potentiële werkgever. De mail komt binnen, en de naam wordt onmiddellijk gegoogeld. De sollicitant is online al helemaal uitgepluisd voordat het cv is bekeken. Je gaat op een date en nog voordat jullie goed gesprek hebben gevoerd, zijn jullie via Google al veel over elkaar te weten gekomen. Hedendaags scenario's, of niet alleen maar? Waren zulke verwickelingen vroeger ook denkbaar, vóór de opkomst van digitale zoekmachines?

Zoekmachines zijn eeuwen geleden ontstaan, ruim voor het digitale tijdperk, stelt Anton Tantner, hoogleraar geschiedenis aan de Universiteit van Wenen. Hij onderzoekt hoe men in de periode van de 15e tot en met laat 20e eeuw naar informatie zocht en hoe deze werd geïndexeerd. Daarnaast vergelijkt hij problemen van toen met die van nu. Wat blijkt? Privacy-issues, gebrekkig toezicht op de geheimhouding van geregistreerde informatie en het gebruik van zoekmachines door de overheid om persoonlijke informatie te achterhalen, zijn niet alleen zorgen in het digitale tijdperk. Ook in het analoge tijdperk, lang voor het bestaan van Google en inlichtingen- en veiligheidsdiensten als de NSA, speelden deze problemen al.

In het artikel 'Before Google. A Pre-History of Search Engines in Analogue Times', noemt Tantner diverse voorbeelden van zoekmachines die werden geraadpleegd in het analoge tijdperk. Hij onderscheidt daarbij drie categorieën: boeken, individuen en instellingen.





Boeken

Encyclopedieën en bijbels zijn de bekendste voorbeelden van boeken waarin door middel van registers en indexen een grote hoeveelheid kennis is geïndexeerd en doorzoekbaar is gemaakt. Minder bekend is het bibliometrisch systeem dat sinds 1963 wordt gebruikt om verwijzingen binnen de wetenschappelijke literatuur vast te leggen. Door te registreren welke artikelen naar welke andere artikelen verwijzen ontstaat een zogeheten 'citatie-index': een lijst van gerangschikte bronnen op basis van het aantal onderlinge verwijzingen. Hoe vaker er naar de bron verwezen wordt, des te hoger de bron in de index komt te staan. Het analyseren van citaten is belangrijk in wetenschappelijke kring, maar speelt tevens een cruciale rol in digitale zoekmachines. Het PageRank principe, een van de algoritmes waarmee Google websites rangschikt in zoekresultaten, werkt namelijk op een vergelijkbare manier: hoe vaker er wordt verwezen naar een website, hoe hoger de website in de zoekresultaten terechtkomt.

Individueen

Individueen die in het begin van de vijftiende eeuw in de grote Europese steden informatie verzamelden en verhandelden, beschouwt Tantner als 'menselijke zoekmachines'. Zij heten ook wel 'contactmakelaars'. Zoals bemiddelaars tussen werkzoekende bediendes en huishoudens die er een nodig hadden. Deze bemiddelaars werden destijds veel geraadpleegd, maar waren niet geliefd bij iedereen. Ze vroegen commissie en beschikten over veel persoonlijke informatie. Tussenpersonen op het gebied van handel, in de late middeleeuwen en de vroegmoderne periode, hadden een vergelijkbare functie. Als er een handelaar de stad binnenkwam hielpen zij, ook tegen een commissie, hem met het vinden van resellers, die hun goederen kochten om deze vervolgens door te verkopen.

Bedienden zelf kunnen volgens Tantner bij uitstek ook als menselijke zoekmachines worden beschouwd. Rond 1800 was het heel gebruikelijk voor de adel en bourgeoisie om voor langere tijd een bediende in huis te hebben. Zij kenden vaak veel geheimen over de familie, bijvoorbeeld op financieel gebied. Er was sprake van een zekere spanning tussen de bediende als hulp en het beeld van de bediende als informant of spion die geheimen doorspeelde. Zogeheten 'leenlakeien' waren ook populair in die tijd, maar hun hulp was in tegenstelling tot bedienden van korte duur. Zij verstrekten reizigers informatie over de stad, bijvoorbeeld de beste herbergen om te bezoeken.





Een ander type menselijke zoekmachine in grote steden was de conciërge, die alle huurders van grote wooncomplexen door en door kende. Niet de eigenaren, maar conciërges woonden in de huurhuizen, samen met de huurders. Zij inden de huur, maar moesten ook zoveel mogelijk te weten komen over de gebeurtenissen in het huis en dit rapporteren aan de eigenaren. In Wenen, waar het beroep volgens Tantner is ontstaan, verstrekten de conciërges zelfs informatie aan de politie en vice versa, zodat verdachte huurders werden gerapporteerd en er geen verdachte personen in huurhuizen belandden. Dit soort samenwerkingen brengt ons bij de derde categorie: instellingen.

Individen die in het begin van de vijftiende eeuw in de grote Europese steden informatie verzamelden en verhandelden, beschouwt Tantner als 'menselijke zoekmachines'

Instellingen

In de voorgeschiedenis van digitale zoekmachines is de institutionalisering van de uitwisseling van alledaagse informatie een belangrijke ontwikkeling. Aan het einde van de zestiende eeuw werd er tussen personen onderling steeds minder informatie uitgewisseld. Tussen de zeventiende tot vroeg negentiende eeuw werd die rol overgenomen door formele instellingen die dienden als een centrale informatiebank (zie Afbeelding 1). Zogenaemde inlichtingenkantoren, ofwel registratiekantoren, verzorgende de officiële registratie en uitwisseling van informatie over individuen (uit de stad), werkgelegenheid, appartementen, handelswaren en geld. Tegen betaling kon iedereen bij een inlichtingenkantoor informatie inwinnen. De kantoren hadden als doelstelling om kennis in de samenleving te organiseren en om op iedere vraag een antwoord te hebben. Wat al snel doet denken aan Google en haar missie: '(...) alle informatie ter wereld te organiseren en universeel toegankelijk en bruikbaar te maken'.



Aangezien inlichtingenkantoren informatie over mensen (identiteit, adres, zaken, doel, arbeidsverleden, criminele antecedenten, sociaal kapitaal en geschiedenis) verzamelden en verkochten, werden ze al snel beschouwd als politie-achtige, controlerende instanties die persoonlijke informatie uitbuiten door er in te handelen. Deze kritiek leidde uiteindelijk tot het staken van hun activiteiten.

Examples for European Intelligence Offices



- Bureau d'adresse, Paris 1630 – 1643 (Théophraste Renaudot)
- Office of Address (project), London/Oxford 1647 (Samuel Hartlib)
- Bureau d'adresse, Paris 1676 (François Colletet)
- Fragamt (question office), Vienna 1707 – ca. 1810 (Ghelen family)
- Universal Register Office, London 1750 – 1760ies (Fielding brothers)
- Fragamt (question office), Lemberg/Lviv 1782 (Friedrich Wilhelm Schütz)

Afbeelding 1: Voorbeelden van Europese inlichtingenkantoren
(slide uit <https://vimeo.com/79882082>)

Volgens Tantner is de voorgeschiedenis van zoekmachines en de bijbehorende problematiek nog niet compleet. Wat Tantner nu al wel concludeert, is dat zowel de zoekmachines van toen als die van nu ambivalent zijn en vergelijkbare problemen veroorzaken. Wij kunnen zoekmachines beschouwen als onze 'frenemy': vriend én vijand. Ze zijn vriendelijk en nuttig, maar tegelijkertijd niet onschuldig. Denk bijvoorbeeld aan de relatie tussen Google en intelligentiediensten als de NSA. Zoekmachines zijn erg machtig, omdat zij beschikken over een grote hoeveelheid persoonlijke en gedetailleerde informatie. Dat geldt zowel voor het analoge als het digitale tijdperk.

Lees het hele artikel van Anton Tantner in de *Society of the Query Reader*:
http://networkcultures.org/query/wp-content/uploads/sites/4/2014/06/9.Anton_Tantner.pdf

Before Google: A Pre-History of Search Engines in Analogue Times (pdf).



Discussiepunten:

1. In het analoge tijdperk vroegen zogenaamde menselijke zoekmachines om een betaling voor hun informatie. Hoe verhoudt dit zich tot het digitale tijdperk? Is zoeken op Google echt 'gratis'?

2. De zoekmachine Ask.com heette eerst AskJeeves. Wat is de connectie tussen de oorspronkelijke naam van deze digitale zoekmachine en zoekmachines uit het analoge tijdperk?

<http://nl.ask.com/?o=312&l=dir>

3. Zie jij Google als je vriend, vijand of beide? Vind je het een voor- en/of nadelig dat de zoekmachine wordt gebruikt om persoonsgebonden informatie op te zoeken? Wat vind je ervan dat er mogelijk uitvoerig wordt gegoogeld in sollicitatieprocedures?

<http://blog.iusmentis.com/2013/05/01/mag-je-sollicitanten-googelen/>





Verder lezen en kijken:

- +** De lezing die Anton Tantner gaf op Society of the Query #2 is op het blog beschreven en terug te zien op Vimeo: 'Anton Tantner – Human Search Engines and Intelligence Offices of the Analogue Age', Institute of Network Cultures, 11 november 2013.
<http://networkcultures.org/wpmu/query/2013/11/11/anton-tantner-human-search-engines-and-intelligence-offices-of-the-analogue-age/>
<http://vimeo.com/79882082>

- +** Karina Ahles-Frijters, 'Google legt zich neer bij dubieuze rol overheid op het internet', Dutch Cowboys, 16 september 2013.
<http://www.dutchcowboys.nl/google/29764>

- +** Hans Cools, Marika Kebluses en Badeloch Noldus (ed.), Your Humble Servant. Agents in Early Modern Europe. Verloren, recensie Annemarieke Blanckesteijn, historischhuis.nl.
<http://historischhuis.nl/recensiebank/review/show/454>

- +** Barton Gellman en Ashkan Soltani, 'NSA Infiltrates Links to Yahoo, Google Data Centers Worldwide, Snowden Documents say', The Washington Post, 30 oktober 2013.
http://www.washingtonpost.com/world/national-security/nsa-infiltrates-links-to-yahoo-google-data-centers-worldwide-snowden-documents-say/2013/10/30/e51d661e-4166-11e3-8b74-d89d714ca4dd_story.html

- +** Maarten Hell, 'Persoonsregistratie en identificatie tijdens het ancien regime', Ons Amsterdam 10 (oktober 2013).
<http://www.onsamsterdam.nl/tijdschrift/jaargang-2013/33-tijdschrift/tijdschrift-jaargang-2013/2639-nummer-10-oktober-2013?showall=&start=3>

- +** Janneke Sløetjes, 'Telt jouw privacy nog voor Google?' Bits of Freedom, 3 februari 2012.
<https://www.bof.nl/2012/02/03/telt-jouw-privacy-nog-voor-google/>





Ik ben geen zoekresultaat! Ik ben een vrij woord!

Anna Jobin & Olivier Glassey

In de Britse cult spy-fi serie *The Prisoner*, probeert een ontvoerde ex-spion te ontsnappen uit een vakantiedorp, waarin een geheime organisatie hem gevangen houdt. De organisatie probeert door middel van grootschalige surveillance informatie los te krijgen over alle bewoners in het dorp, inclusief de ex-spion. Gedrag, interesses, gewoontes en taalgebruik: alles wordt gemonitord vanuit een geheime controlepost. Maar hoeveel informatie er precies wordt verzameld en hoe die wordt verwerkt, weten de bewoners niet. De geheime organisatie bestuurt letterlijk van achter de schermen het dorp. Bovendien weet ze alle vragen van de bewoners te beantwoorden, zonder dat duidelijk is hoe de antwoorden tot stand komen. De hoofdpersoon vertrouwt het zaakje niet en komt in verzet tegen het systeem.

Anna Jobin en Olivier Glassey, beiden verbonden aan de Universiteit van Lausanne, gebruiken het scenario van *The Prisoner* als metafoor voor de relatie tussen Google en woorden. Woorden zijn immers de kern van elke zoekopdracht. We voeren woord na woord in op Google, die ze vervolgens verwerkt om bepaalde zoekresultaten te leveren. Maar hoe Googles algoritmes precies werken is onbekend. Woorden zijn als het ware gevangen in een ontransparant systeem zonder mogelijkheid te ontsnappen aan de door commercie gedreven algoritmes. Woorden zijn overgeleverd aan 'semantisch determinisme', wat wil zeggen dat Google ervan uitgaat dat alles in de wereld in woorden is uit te drukken, en in het verlengde daarvan kan worden geïndexeerd door Google.

Commodificatie en categorisatie

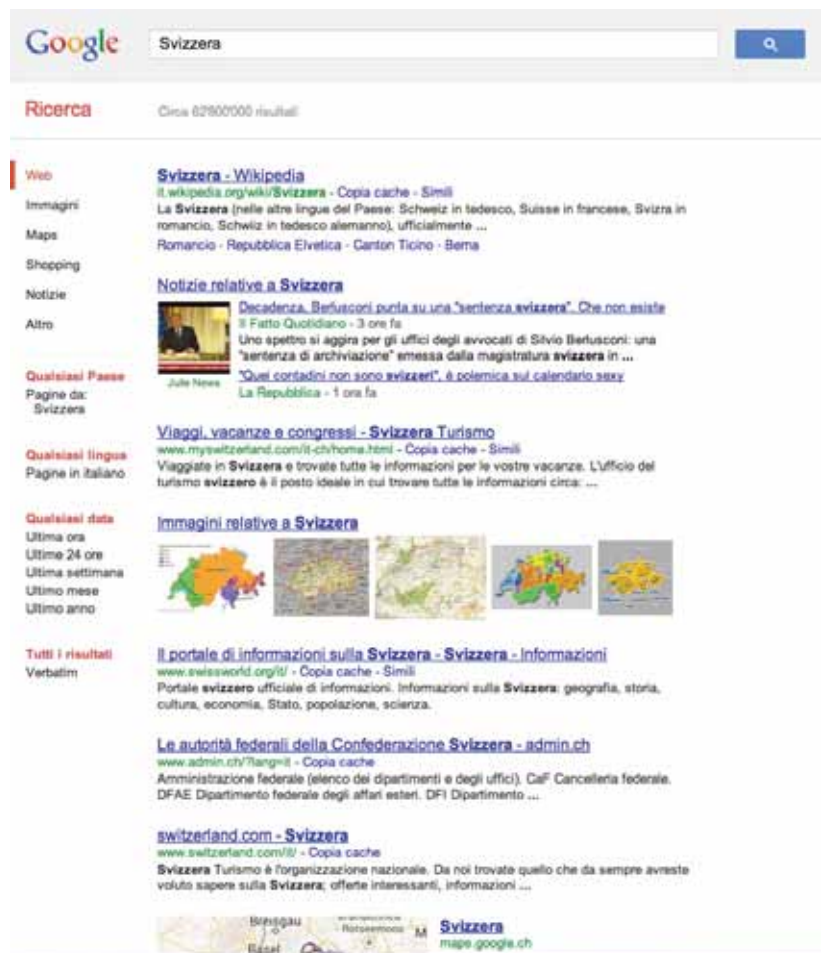
Jobin en Glassey stellen dat de gevangenschap van de woorden komt doordat Google woorden commodificeert en categoriseert. Commodificatie wil zeggen dat de woorden op een procesmatige wijze te gelde worden gemaakt via advertentieverkopen. Dat gebeurt aan de hand van de zogeheten Keyword planner, waarmee bepaalde combinaties van zoektermen tegen betaling aan advertenties worden gekoppeld.



Google plaatst de keywords (combinaties van zoektermen) in categorieën van het Google Adwords advertentiesysteem. Jobin en Glassey onderzoeken hoe Google omgaat met het woord Zwitserland of afgeleiden daarvan. Zij kozen voor Zwitserland als casestudy omdat het land maar liefst vier officiële talen kent – elk met zijn eigen woord voor Zwitserland. Het blijkt dat bijvoorbeeld 'Wohnungen Schweiz' (appartementen Zwitserland) in de categorie 'Schweiz Sehenswürdigkeiten' (Zwitserland bezienswaardigheden) belandt. Jobin en Glassey menen dat het plaatsen van keywords in categorieën vaker arbitrair en ontransparant gebeurt. Bovendien is de categorisering soms onlogisch. Waarom is 'appartementen Zwitserland' niet gewoon ondergebracht in de categorie 'Zwitserland vastgoed'?

Google stelt woorden door deze categorisering gelijk aan betekenissen, zeggen Jobin en Glassey, en dat staat de vrije interpretatie van woorden in de weg. Anders gezegd, we kunnen niet zelf meer bepalen wat een woord 'betekent', want dat is al vastgelegd door Google in databases, bovendien geleid door commerciële drijfveren. De verkoop van woorden verandert zodoende in de handel in betekenissen, aldus Jobin en Glassey.

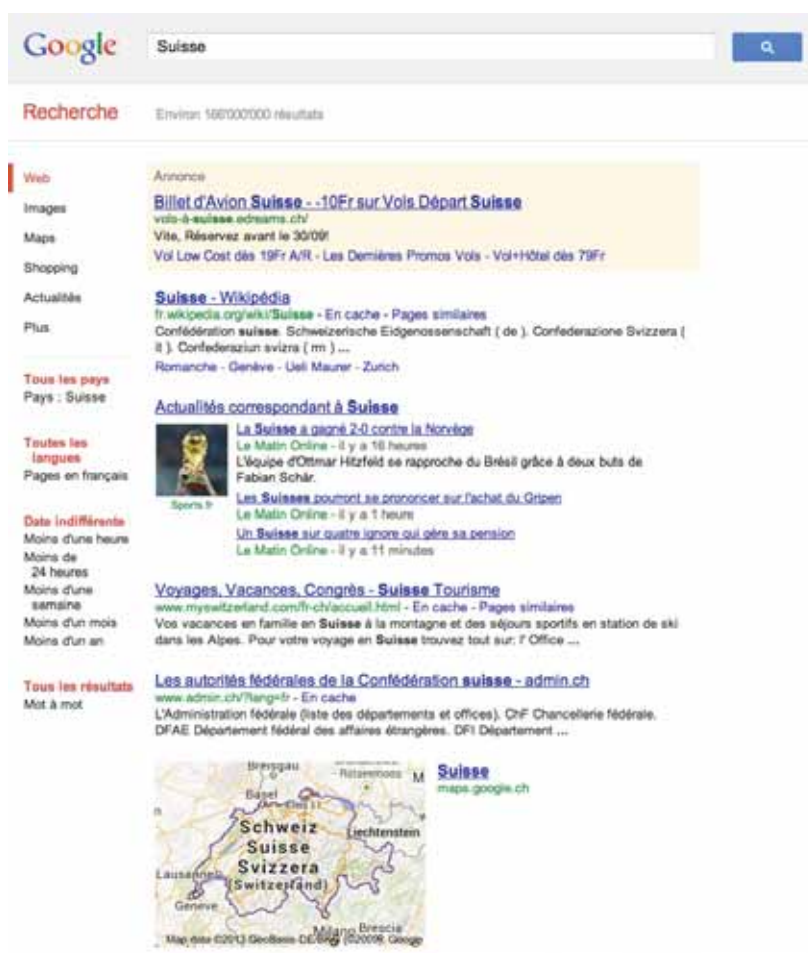
Het proces van commodificatie en categorisatie is van invloed op de samenstelling van zoekresultaten, waaronder advertenties. Maar welke taalgerelateerde factoren spelen daar nog meer een rol in?



Afbeelding 1: 'Svizzera' ingevoerd op Google.ch met Italiaanse taalinstelling

Filterbubbel of taalbubbel

Jobin en Glassey tonen aan de hand van een casestudy aan dat Google zoekresultaten levert die zijn gefilterd op taalinstellingen en locatie. In een experiment zochten zij meerdere malen met de zoekterm 'Switzerland' op de algemene Google-pagina van Zwitserland, <http://google.ch>. Hierbij gebruikten zij verschillende taalinstellingen, wat realistisch is vanwege de meertaligheid in Zwitserland: Duits, Frans, Italiaans en Reto-Romaans zijn alle officiële talen. Google bleek bij verschillende taalinstellingen uitermate verschillende zoekresultaten te leveren, zowel wat betreft de samenstelling van de bronnen als de volgorde ervan. 'News for Switzerland' werd bijvoorbeeld alleen getoond bij het Engels als taalinstelling. Toen zij 'Zwitserland' googelden in dezelfde taal als waarmee de zoekmachine was ingesteld – bijvoorbeeld 'Svizzera' op Google Svizzera (met Italiaanse taalinstelling) – waren er onderling zelfs nog grotere verschillen. Zo leverde Google Svizzera bijvoorbeeld alleen links naar Italiaanse nieuwsmedia en helemaal niet naar de Zwitserse media (zie Afbeelding 1), in tegenstelling tot Google Schweiz (Duitstalig) en Google Suisse (Fransstalig) die dat wel deden (zie Afbeelding 2).



Afbeelding 2: 'Suisse' ingevoerd op Google.ch met Franse taalinstelling



Google maakt op basis van je IP-adres en locatie bepaalde aannames over je taalvoorkeur en dat leidt tot verschillende zoekresultaten. Problematisch, want informatie over locatie of de aangenomen voorkeurstaal hoeft niets te zeggen over de daadwerkelijke voorkeur die iemand heeft.

Google maakt op basis van je IP-adres en locatie bepaalde aannames over je taalvoorkeur en dat leidt tot verschillende zoekresultaten



Jobin en Glassey vinden het vreemd is dat de inwoners van eenzelfde land verschillende zoekresultaten krijgen. Taalgebruik en taalinstellingen zijn op Google bepalend voor welke resultaten gebruikers wel en niet te zien krijgen. De optie om je taalinstelling op Google te wijzigen is misleidend, omdat het erop lijkt alsof dat slechts betrekking heeft op de taal van de interface. Dat je door zo'n wijziging heel andere zoekresultaten kunt krijgen blijft onvermeld.

Jobin en Glassey bouwen met hun onderzoek voort op het eerdergenoemde concept van de 'filterbubbel' (zie Introductie) en laten zien dat de filterbubbel verder reikt dan alleen het individuele niveau. Niet alleen persoonlijk zoekgedrag is van invloed op het filteren van zoekresultaten. Ook taal, locatie en taalinstellingen spelen mee als het erom gaat of iemand wel of geen toegang heeft tot bepaalde bronnen en informatie.

Net als Kylie Jarrett beschrijft in haar artikel 'A Database of Intention?', wijzen Jobin en Glassey op de commerciële en ontransparante algoritmes die onze denkprocessen en verwoordingen aansturen, bijvoorbeeld door de autocomplete-functie, gerelateerde zoekopdrachten en voorgestelde gecorrigeerde zoektermen. De rol van taal en taalinstellingen hierbij moet worden onderkend, vinden zij. Wees je bewust van wat de zoekmachine met taal doet en wat dit betekent. Denk jij wel eens na over de zoekresultaten die je niet krijgt, als gevolg van jouw taalgebruik, locatie of taalinstellingen?

Lees het hele artikel van Anna Jobin en Olivier Glassey in de *Society of the Query Reader*:

http://networkcultures.org/query/wp-content/uploads/sites/4/2014/06/11.Jobin_Glassey.pdf

'I Am not a Web Search Result! I Am a Free Word': The Categorization and Commodification of 'Switzerland' by Google (pdf).





Discussiepunten:

1: Opdracht: Ga naar het Belgische Google en zoek op 'België' met verschillende taalinstellingen: Nederland, Duits, Frans en Engels. Bekijk de eerste 10 links, inclusief de advertenties. In hoeverre verschillen de zoekresultaten met elkaar?

<https://www.google.be/>

2: Opdracht: Doe nu hetzelfde maar zoek op 'België' in dezelfde taal als die is ingesteld. Bijvoorbeeld 'Belgien' als je op de Duitse versie van <https://www.google.be> zoekt. Doe dit weer met 4 verschillende taalinstellingen. Wat valt je op als je de zoekresultaten met elkaar vergelijkt?

<https://vimeo.com/79884090>

Society of the Query #2 - Pascal Jürgens: Measuring Personalization

3: Pascal Jürgens stelde bij de tweede Society of the Query conferentie dat er sinds de publicatie van Eli Pariser, *The Filter Bubble: What The Internet Is Hiding From You* (2012), geen sprake meer is van een filterbubbel die leidt tot gepersonaliseerde zoekresultaten. Hij maakt wel net als Jobin en Glassey aannames over de invloed van lokalisering op zoekresultaten. Wat zijn deze aannames en sluiten deze goed aan op de stellingen van Jobin en Glassey?





Verder lezen en kijken:

- +** James Grimmelman, 'The Google Dilemma', New York Law School Law Review Januari 2009: 939.
http://works.bepress.com/james_grimmelman/19

- +** Boris Groys, 'Google: Words beyond Grammar', dOCUMENTA (13): 100 Notizen - 100 Gedanken.
<http://www.hatjecantz.de/boris-groys-2785-1.html>

- +** Zoe Kleinman, 'How the Internet is Changing Language' BBC News, 16 augustus 2010.
<http://www.bbc.co.uk/news/technology-10971949>

- +** Steven Levy, 'Did Google Just Pop the Filter Bubble?' 10 januari 2012.
<http://www.stevenlevy.com/index.php/01/10/did-google-just-pop-the-filter-bubble>

- +** Chris Smith, 'How is Web Personalisation Affecting the News?' Guardian Professional, 4 september 2012.
<http://www.theguardian.com/media-network/media-network-blog/2012/sep/04/news-personalisation-content-publisher-media>





Zoeken over de grens: China

Min Jiang



We denken al snel dat het internet werkelijk een world wide web is, en dat zoekmachines internationaal, zelfs globaal opereren. Maar het is een illusie dat zoekmachines grenzeloze wereldwijde technologieën zijn, aldus Min Jiang, van de University of North Carolina in Charlotte. Bij het ontstaan van het world wide web zo'n twintig jaar geleden heerste er een kosmopolitische visie op cyberspace (<http://nl.wikipedia.org/wiki/Cyberspace>): iedereen zou op een haast utopische manier met elkaar in verbinding staan via het internet. Maar dat is door de jaren heen ingrijpend veranderd. Nationale overheden oefenen tegenwoordig een grondige controle uit op het web, dat daardoor in toenemende mate is genationaliseerd en gelokaliseerd, stelt Jiang. Er mag sprake zijn van een globale technische infrastructuur, ervaringen op internet verschillen op nationaal en lokaal niveau, vooral wat betreft de beschikbaarheid van content en informatie. Je vindt niet overal dezelfde informatie en niet overal is alle informatie beschikbaar.

Bij grenzen denk je al snel aan douaneposten en territoria. Maar de vier grenzen die Jiang met betrekking tot zoekmachines waarneemt zijn onzichtbaar. Die 'herterritorialiseren' het internet in haar woorden. Taalkundige, politieke, technische en economische grenzen zijn volgens haar bepalend voor de informatie die wel of niet via zoekmachines beschikbaar wordt gemaakt. Zo is dat wat je vindt op het Chinese internet heel erg geconditioneerd door je locatie, wie je bent, hoe je zoekt, wat je weet en welke taal je spreekt.

Taalkundige grenzen

Taalkundige grenzen hebben te maken met de verschillende zoekresultaten die je krijgt op basis van je locatie en de taal die je gebruikt. In China is er een grote diversiteit aan talen en dialecten. Die brengen geo-taalkundige grenzen met zich mee tussen het Chinese vasteland, Hong Kong, Taiwan en andere regio's. Zo is op het vasteland vereenvoudigd Mandarijn de dominante gesproken en geschreven taal, terwijl dit in Taiwan Kantonees is, dat gebruik maakt van traditionele karakters. Taal is voor Chinese internetters op die manier gebonden aan hun locatie. Dit is van invloed op de zoekmachine(s) die zij (kunnen) gebruiken en de zoekresultaten die zij krijgen.





Aangezien de meeste websites in China in het vereenvoudigd Chinees zijn, zullen mensen die alleen de traditionele taal beheersen een minder groot en divers aanbod aan informatie(bronnen) tot zich kunnen nemen via zoekmachines. Bovendien indexeert Baidu, de grootste zoekmachine in China, alleen websites met vereenvoudigde Chinese karakters. Op andere zoekmachines zoals Google en Bing zijn zoekopdrachten in zowel traditoneel als vereenvoudigd Chinees mogelijk, maar door censuur en door de hoeveelheid websites die in de verschillende talen wordt geïndexeerd is het verschil aan zoekresultaten voor dezelfde zoektermen enorm.

Begin dit jaar zou de zoekmachine Bing zoekopdrachten in het vereenvoudigd Chinees hebben gecensureerd (<http://www.theguardian.com/technology/2014/feb/11/bing-censors-chinese-language-search-results>). Het merkwaardige hierbij was dat de zoekresultaten in het traditioneel Chinees niet werden gecensureerd. Daarnaast gebeurde dit in Amerika. Hieruit blijkt dat zoekresultaten kunnen worden geconditioneerd door je taal, ongeacht je locatie. Tegelijkertijd is het in China juist de locatie die bepalend is voor de taal die men spreekt, schrijft en waar men zoekopdrachten in formuleert.

De conditionering van zoekresultaten op basis van je taal en locatie is in de ogen van Jiang problematisch omdat men hierdoor belangrijke informatie kan mislopen, bijvoorbeeld nieuws over de Arabische Lente. Veel Chinese inwoners hebben het nieuws over de toedracht van de Arabische Lente helemaal niet goed kunnen volgen, vanwege censuur en een gebrek aan informatie in hun taal. Op de Society of the Query conferentie noemde Jiang Yeeyan.org als hulpmiddel om taalkundige grenzen mee over te steken: een vrijwilligersproject en website met de missie om internationaal nieuws, door middel van vertaling, zo toegankelijk mogelijk te maken voor de Chinese bevolking.

Zowel de Chinese staat als online diensten filteren informatie ook op lokaal niveau aan de hand van het lokale politieke klimaat en de wetgeving





Politieke grenzen

Het is bekend dat de Chinese staat Amerikaanse websites blokkeert en zoekresultaten van zoekmachines filtert met de beruchte 'Great Firewall of China' (https://en.wikipedia.org/wiki/Great_Firewall_of_China). Daarnaast heeft Google sinds 2006, toen het bedrijf zich in China vestigde, zoekresultaten gefilterd op het Chinese Google-portaal Google.cn. Het bedrijf mocht in China opereren onder voorwaarde dat het zelf zoekresultaten zou censureren. Google stopte hiermee in 2010, nadat een groep Chinese hackers Googles servers zou hebben aangevallen

(<http://www.nytimes.com/2010/01/13/world/asia/13beijing.html?pagewanted=all&r=0>) om informatie over Chinese mensenrechtenactivisten te verkrijgen.

Google verhuisde vervolgens haar servers van het Chinese vasteland naar Hong Kong, omdat in dat voormalig Brits protectoraat nog wel sprake is van vrijheid van meningsuiting. Google biedt sindsdien in Hong Kong ongecensureerde resultaten. In Hong Kong heerst een ander politiek klimaat en gelden andere regels; Google is met de verhuizing dus een geopolitieke grens over gestapt.

Dat de Chinese staat censuur uitoefent met de 'Great Firewall of China' is volgens Jiang niet het grootste probleem. Het is volgens haar erger dat het filteren van informatie grootschalig wordt uitbesteed aan online diensten, waaronder de in China dominantste zoekmachine Baidu. Deze kunnen op veel subtielere wijze informatie filteren dan de staat dit doet, zoals je ook kunt zien in de video hierboven. Baidu filtert bijvoorbeeld zowel de zoekresultaten als de websites die gebruikers kunnen bezoeken. Daarnaast is er sprake van 'search engine bias', het bevoordelen van bepaalde resultaten, aangezien Baidu veel meer naar eigen diensten als Baidu Baike (een site die lijkt op Wikipedia, http://nl.wikipedia.org/wiki/Baidu_Baike) en overheidssites verwijst dan naar vergelijkbare bronnen. Zowel de Chinese staat als online diensten filteren informatie ook op lokaal niveau aan de hand van het lokale politieke klimaat en de wetgeving, waardoor de zoekresultaten van zowel Baidu als Google per Chinese regio zeer verschillend kunnen zijn.





Technische en economische grenzen

Zoektechnologie weet in toenemende mate de gebruiker ervan te lokaliseren, aldus Jiang. Zoekresultaten worden aangepast aan de locatie van het IP-adres. Lokalisering is een groeiende trend, met name omdat commerciële bedrijven met lokale advertenties hun diensten en producten aan de man willen brengen. Er gaat tegenwoordig erg veel geld om in deze business, waardoor geo-technische grenzen vaak gepaard gaan met geo-economische grenzen.

Lokalisatietechnologieën en lokale bedrijven hebben aldus een steeds grotere invloed op de toegankelijkheid van informatie. Dat heeft nadelige gevolgen voor de diversiteit van beschikbare informatie en kennis die met zoekmachines tot je beschikking staat, aldus Jiang.

Grenzen oversteken?

Relatief weinig mensen in China maken gebruik van een proxyserver (<http://nl.wikipedia.org/wiki/Proxyserver>) of VPN (http://nl.wikipedia.org/wiki/Virtueel_Particulier_Netwerk) om de internetcensuur te omzeilen. Zij die dat wel doen werken bijvoorbeeld bij buitenlandse bedrijven die zich in China hebben gevestigd. De meeste Chinezen weten echter niet eens dat dit kan, zijn niet tech-savy genoeg, of misschien wel helemaal niet geïnteresseerd in het buitenlandse online aanbod. Maar weinig mensen in China lezen en spreken Engels: volgens schatting van Jiang ongeveer 100.000 van de 1,3 miljard inwoners. De meerderheid van de Chinezen is zodoende aangewezen op de informatie die de staat beschikbaar maakt; een alternatief bestaat niet voor hen.

Welke mogelijkheden zijn er om de onzichtbare grenzen over te steken? Waar hechten we als internetgebruikers waarde aan, lokale of globale content? Jiang vindt dat we op zoek moeten naar een balans tussen lokalisering en globalisering op het internet. Zij pleit voor opties om onze zoekervaringen zelf te personaliseren, in plaats van de ondoorzichtige wijze waarop zoekmachines dit voor ons doen. We zouden in staat gesteld moeten worden om zelf te bepalen welke bronnen we willen raadplegen als we naar informatie zoeken.

Lees het hele interview met Min Jiang in de *Society of the Query Reader*:

http://networkcultures.org/query/wp-content/uploads/sites/4/2014/06/10.Min_Jiang.pdf

Search Control in China (pdf).







Discussiepunten:

- 1: Anna Jobin en Olivier Glassey stellen in het artikel "I Am not a Web Search Result! I Am a Free Word!" The Categorization and Commodification of 'Switzerland' net als Jiang dat taal de beschikbare informatie van een zoekmachine beïnvloedt. Vergelijk de verschillende perspectieven en stellingen uit de twee artikelen met elkaar. Sluiten deze goed op elkaar aan, of niet?
- 2: De grenzen die Jiang beschrijft zijn volgens haar van toepassing op het Chinese internet. Zou je een voorbeeld kunnen bedenken waarbij informatiestromen op het internet in Nederland onderhevig zijn aan deze grenzen?
- 3: Jiang pleit voor meer opties voor personalisatie en keuzevrijheid in relatie tot zoekmachines en informatiebronnen die zij tot zoekresultaten verwerken. Vind je dit een reëel voorstel?

Verder lezen en kijken:

-  De lezing die Min Jiang gaf op Society of the Query #2 is op het blog beschreven en terug te zien: 'Min Jiang – Search Without Borders', Institute of Network Cultures, 11 november 2013.
<http://networkcultures.org/wpmu/query/2013/11/11/min-jiang-search-without-borders/>
<http://vimeo.com/79287215>
-  Eugene Volokh, 'Chinese Search Engine, Operating in the U.S., has First Amendment Right to Skew Search Results for Political Reasons', The Washington Post, 28 maart 2014.
<http://www.washingtonpost.com/news/volokh-conspiracy/wp/2014/03/28/chinese-search-engine-operating-in-the-u-s-has-first-amendment-right-to-skew-search-results-for-political-reasons/>





- +** Eileen Yu, 'China's New Search Engine Unlikely to do much Work', ZDnet, 24 maart 2014.
<http://www.zdnet.com/cn/chinas-new-search-engine-unlikely-to-do-much-work-7000027611/>
- +** Dominic Rushe, 'Microsoft Blames "system error" but Denies Censoring Chinese Search Results', The Guardian, 12 februari 2014.
<http://www.theguardian.com/technology/2014/feb/12/microsoft-bing-censor-chinese-search-results-system-error>
- +** XiaChu, Bing Bests Baidu Censorship, Greatfire.org.
<https://zh.greatfire.org/blog/2014/mar/bing-bests-baidu-censorship>
- +** Ö. Canel, 'De Chinese zoekmachine Baidu', VPRO Tegenlicht, 12 april 2011.
<http://tegenlicht.vpro.nl/nieuws/2011/april/baidu.html>
- +** Liz Tung, 'What Makes Censorship in China So Effective?' 10 september 2013.
<http://www.theatlantic.com/china/archive/2013/09/what-makes-censorship-in-china-so-effective/279467/>
- +** Han-Teng Liao, "'Wikipedias" (or its Copycat) Dominate "Chinese" Search Engine Result Pages (SERPs)', 6 mei 2013.
<http://people.oii.ox.ac.uk/hanteng/2013/05/06/wikipedias-or-its-copycat-dominate-chinese-search-engine-result-pages-serps/>





Google staat boven de wet en voor de rechter

Angela Daly



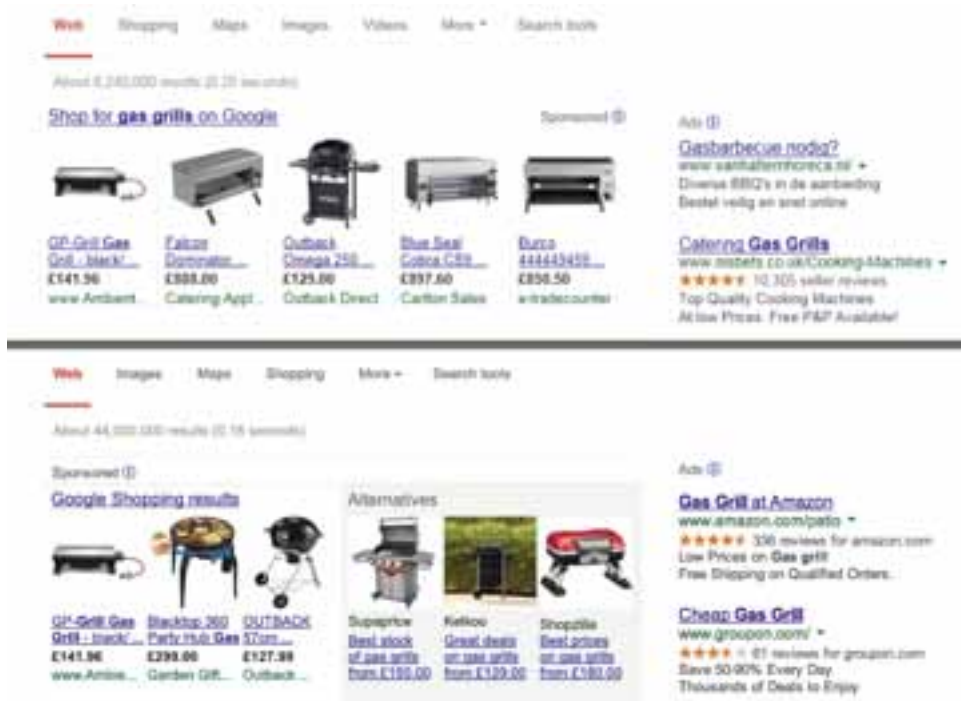
Hoe Google's Europese zoekadvertenties kunnen veranderen

<https://www.youtube.com/watch?v=NIT-QirW3uM>

In deze video wordt een mogelijke doorbraak aangekondigd in een langlopend onderzoek naar misbruik van de dominante positie op de markt van Google. Enkele bedrijven waaronder Microsoft, spanden in 2010 een mededingingszaak aan tegen Google omdat de zoekmachine de eigen diensten zou voortrekken in resultaten en advertenties, ten nadele van concurrerende diensten. Google zou links naar eigen diensten prominenter plaatsen dan die naar andere gespecialiseerde zoekmachines. Dit zijn websites die een specifiek type online content doorzoekbaar maken, bijvoorbeeld webshops of prijsvergelijkingsites. Daarnaast zou Google de zogeheten 'Quality Score' van gesponsorde links naar concurrerende diensten hebben verlaagd. De advertenties van concurrenten zouden hierdoor minder vaak in de zoekresultaten zijn verschenen omdat de Quality Score daarin een bepalende factor is.

Google heeft sinds 2010 al drie schikkingsvoorstellen gedaan. Twee keer werden de voorstellen te licht bevonden, maar nu lijkt het erop dat de Europese Commissie akkoord zal gaan. Google garandeert hierin de links naar producten en diensten van concurrerende bedrijven duidelijk zichtbaar en afgekaderd in de zoekmachineresultaten te zullen verwerken (zie Afbeelding 1).





Afbeelding 1: Google's huidige (boven) en voorgestelde nieuwe weergave (onder) van gesponsorde links. (Bron: Europese Commissie)

In het geval dat zowel de Europese Commissie als de aanklagers akkoord gaan met de schikking, ontloopt Google een grootschalig onderzoek naar de precieze werking van de zoekmachine en bovendien een grote boete van miljarden euro's. Het lijkt er misschien op dat Google in Europa gewillig aanpassingen doet ten behoeve van een eerlijkere concurrentie. Maar was dit niet een uitgelezen kans om veel meer over de ondoorzichtige werking van de grootste zoekmachine ter wereld te weten te komen? Kunnen de problemen die ontstaan in een markt die in zo hevige mate gedomineerd wordt door een enkel bedrijf, wel met de bestaande wetgeving met betrekking tot concurrentie aangepakt worden?

Het ontbrekende gebruikersperspectief

In 'Google Before the Law' stelt Angela Daly, Swinburne University of Technology in Melbourne, het monopolie en de handelswijze van Google in relatie tot Europese en Amerikaanse wetgeving aan de kaak. Volgens haar is de concurrentiewetgeving, die er is om misbruik van macht op de vrije markt tegen te gaan en een eerlijke concurrentie te handhaven, niet geschikt om dit soort problemen op te lossen. De wet met betrekking tot concurrentie is toegespitst op de economische effecten van marktdominantie en gaat niet letterlijk in op het gebruik van een dominante dienst, aldus Daly. De onderzoeken met betrekking tot concurrentiewetgeving hebben volgens haar een te smalle focus, waarin het perspectief van de gebruikers van online diensten grotendeels buiten beschouwing blijft. Net als overigens de rol van Google.



Er is veel te doen over de wijze waarop Google data verzamelt en profielen opstelt van personen



Google is in staat om bepaalde zoekresultaten te censureren, zoals links naar filesharing websites, en kan daarnaast op onnavolgbare wijze de ene informatiebron boven de andere plaatsen in de gepresenteerde ranking van zoekresultaten. In sommige gevallen kan Google bronnen zelfs helemaal weglaten, zonder daar een duidelijke reden voor te hoeven geven. Gebruikers van Google hebben het er maar mee te doen, zonder dat ze precies weten wat er achter de schermen gebeurt.

Er is ook veel te doen over de wijze waarop Google data verzamelt en profielen opstelt van personen. Volgens Daly is hierdoor niet alleen sprake van aantasting van privacy, de enorme berg data die Google heeft verzameld, zorgt er ook voor dat het haast onmogelijk is voor nieuwkomers om een serieuze concurrentiepositie in te nemen. Hoe meer data Google verzamelt, des te accurater de zoekresultaten worden. Daly wijst erop dat Google deze data op grootschalige wijze uitbuit. Dit soort problemen zijn niet opgenomen in de huidige wetgeving, terwijl de monopoliepositie van Google – waar de onderhavige rechtszaak om draaide – wel degelijk met dit soort praktijken te maken heeft.

Zoeken naar oplossingen

Terwijl Google in Europa wordt beschuldigd van machtsmisbruik en naar een schikking toewerkt, komt het bedrijf in Amerika makkelijker weg onder dezelfde beschuldigingen. De FTC, de Federal Trade Commission, deed negentien maanden lang onderzoek naar machtsmisbruik. De conclusie luidde, in 2013, dat Google helemaal geen concurrentiewerend gedrag had vertoond. Het bedrijf zou in plaats daarvan de kwaliteit hebben verbeterd van de eigen zoekdienst, ten gunste van de consumenten. Dat de zogenaamde 'innovatie' van de dominantste speler nadelig uitpakt voor de concurrenten beschouwde de FTC niet als machtsmisbruik.





In Europa wordt momenteel het verdrag voor databescherming herzien, wat misschien gevolgen heeft voor de wijze waarop Google omgaat met het verzamelen van data en kwesties rond privacy. Het gewijzigde verdrag zou er onder meer voor kunnen zorgen dat Google de gebruikers nadrukkelijk toestemming moet vragen voor het verzamelen en verwerken van data. Toch lijkt de kans dat dit gebeurt klein, omdat Google hier met vele andere grote internetbedrijven tegen zal lobbyen.

Censuur, bevoordelen van eigen diensten, aantasting van privacy, gebrek aan transparantie en uitbuiting van persoonlijke data zijn in de ogen van Daly problemen die lastig via de wet te adresseren zijn, gezien de snelheid waarin nieuwe technologie wordt ontwikkeld en geadopteerd. Het kost veel tijd en moeite om passende wetten te schrijven en te implementeren om deze problemen op te lossen. Google staat voor de rechter, maar soms ook boven de wet, aldus Daly. Zij oppert dat men zich daarom wellicht het beste kan richten op buitenwettelijke oplossingen, zoals alternatieve zoekmachines.

Lees het hele artikel van Angela Daly in de *Society of the Query Reader*:

http://networkcultures.org/query/wp-content/uploads/sites/4/2014/06/7.Angela_Daly.pdf

Dominating Search: Google Before the Law (pdf).





Discussiepunten:

- 1: Wat vind jij van de problemen die Daly noemt? Voel jij je hierbij aangesproken?
- 2: Ben je het eens met Daly dat de problemen die door Google zijn veroorzaakt nog niet op de juiste manier in de wetgeving zijn opgenomen?
- 3: Opdracht: In de Verenigde Staten lijkt Google veel makkelijker onder beschuldigingen van machtsmisbruik uit te komen dan in Europa. Is dat vaker het geval? Gebruik Google om te zoeken naar voorbeelden van aanklachten, rechtszaken, schikkingen en boetes in relatie tot Google in het verleden. Wat valt je hierbij op?

Verder lezen:

- + Google belooft Europa beterschap, nos.nl, 5 februari 2014.
<http://nos.nl/artikel/606593-google-belooft-europa-beterschap.html>
- + Jeff Gould, 'Google - EU Privacy Dispute May Affect US Law', 24 januari 2014, InformationWeek.
<http://www.informationweek.com/government/cybersecurity/google---eu-privacy-dispute-may-affect-us-law/d/d-id/1113567>
- + Janneke Slöetjes, 'Google beslissing CBP: vaag privacybeleid niet voldoende', 28 november 2013, Bits of Freedom.
<https://www.bof.nl/2013/11/28/google-beslissing-cbp-vaag-privacybeleid-niet-voldoende/>
- + Claire Cain Miller en Mark Scott, 'Google Settles Its European Antitrust Case; Critics Remain', The New York Times, 5 februari 2014.
<http://www.nytimes.com/2014/02/06/technology/google-reaches-tentative-antitrust-settlement-with-european-union.html>





De ideologie van alternatieve zoekmachines

Astrid Mager



Google is al jaren de meest gebruikte zoekmachine. Maar dat is Google niet zomaar geworden. Als je Google gebruikt, kan het bedrijf data verzamelen, die het inzet om het gebruik nog efficiënter en dus groter te maken, wat nog meer data oplevert, et cetera. In die zin maakt Google optimaal gebruik van de kapitalistische vrijmarktideologie die snelheid, efficiëntie en schaalvergroting dicteert, zo betoogt Astrid Mager, verbonden aan de Austrian Academy of Sciences. Zij stelt dat zoekmachines – net als andere technologie – worden gevormd door de maatschappij waar ze in opereren. Met andere woorden, ontwikkelaars van een zoektechnologie zijn niet als enige verantwoordelijk voor het succes ervan, wij hebben met z'n allen Google dominant gemaakt. Mager ziet onder meer voor internetgebruikers, providers, marketeers, SEO-specialisten, journalisten, beleidsmakers en juristen een belangrijke rol bij het stabiliseren van de macht van Google. Als iedereen Google de rug zou toekeren, kan geen businessmodel, geraffineerd algoritme of database het bedrijf nog helpen. Hoewel de kans dat dit gebeurt klein lijkt te zijn, is er na de onthullingen van Edward Snowden steeds meer aandacht voor alternatieve zoekmachines. Maar wat zijn die zogeheten alternatieve zoekmachines precies? Kunnen we ze wel 'echte' alternatieven noemen? Of opereren ze ook binnen de dominante ideologie?

In het artikel 'Is Small Really Beautiful? Big Search and Its Alternatives' analyseert Mager diverse alternatieve zoekmachines en hun onderliggende ideologieën. Alternatieve zoekmachines zijn technologieën waarvan de ontwikkelaars claimen dat zij op een fundamenteel andere wijze opereren dan de grote bekende tegenhangers als Google, Bing en Yahoo. De grote zoekmachines zijn stuk voor stuk schatplichtig aan de kapitalistische ideologie, waarin het draait om het omzetten van handelingen en data van gebruikers in financiële winst. Hoe zit dat met de kleinere? Mager laat zien dat de alternatieve zoekopties vaak voor hun resultaten afhankelijk zijn van grotere zoekmachines en allianties met hen vormen. Hoewel zij zich aan de oppervlakte lijken te onderscheiden van de grotere zoekmachines en hun aanpak, opereren ze in de meeste gevallen dus in dezelfde kapitalistische geest, aldus Mager. Het opbouwen en onderhouden van een webindex kost immers veel tijd en geld (zie ook het artikel van Dirk Lewandowski). Sommige alternatieven werken daarom samen met commerciële partijen die hun index tegen (financiële) voorwaarden beschikbaar stellen.



DuckDuckGo: de privacyvriendelijke zoekmachine



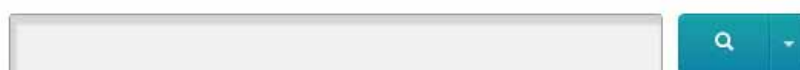
Afbeelding 1: DuckDuckGo

(Bron: <http://donttrack.us>)

Misschien heb je wel eens van DuckDuckGo gehoord. DuckDuckGo profileert zich als een privacyvriendelijke zoekmachine en slaat in tegenstelling tot Google geen zoekgedrag van gebruikers op. Zij spelen dus ook geen data door naar diensten zoals de Amerikaanse NSA. Het aantal zoekopdrachten met DuckDuckGo nam in de tweede helft van 2013, na de onthullingen over de surveillancopraktijken van de NSA, fors toe. Inmiddels verwerkt de zoekmachine dagelijks meer dan 4 miljoen opdrachten. Hoewel de gebruikers inderdaad niet 'gevolgd' worden en data niet wordt verkocht om winst te maken, begeeft DuckDuckGo zich wel op het terrein van commerciële, kapitalistische techbedrijven. DuckDuckGo toont namelijk contextuele advertenties van Bing (Microsoft) en gebruikt andere zoekmachines, waaronder Bing, als bron voor de eigen zoekresultaten.

Hoewel privacy het ideologische fundament is, maakt DuckDuckGo zich daarmee afhankelijk van grotere commerciële zoekmachines en hun verdienmodellen, waaronder het opslaan en doorverkopen van data. Volgens Mager zou DuckDuckGo niet kunnen bestaan zonder de grotere spelers. Dit geldt ook voor soortgelijke alternatieven als ixquick.com en metager.de, die zijn geprogrammeerd om hetzelfde te doen: zoekresultaten van andere zoekmachines ophalen en tonen, zonder persoonsgebonden gegevens van gebruikers op te slaan of door te spelen.

Groene zoekmachines



Afbeelding 2: <http://Ecosia.org>

Groene zoekmachines zijn gebaseerd op wat Mager de 'groene ideologie' noemt. Ze geven via de opbrengst uit zoekopdrachten financiële ondersteuning aan ecologische projecten. Ecosia doneert bijvoorbeeld 80% van de advertentie-inkomsten aan een partij die het Braziliaanse regenwoud bebost. 'You've helped plant trees!' is dan ook hun slogan. Ze werken samen met Yahoo voor advertenties en met Bing voor zoekresultaten. Dat lijkt geen probleem, tot je naar de 'achterkant' van de bedrijven kijkt. De servers van Ecosia draaien op groene stroom, maar die van Bing niet. Hierdoor draagt deze groene zoekmachine indirect bij aan de uitstoot van CO₂.

In de praktijk zijn de groen-ideologische zoekmachines afhankelijk van zoekresultaten en advertentieverdiensten van grotere partijen. Dat vormt een bedreiging in ideologische en in praktische zin. Dit legt Mager uit aan de hand van drie stellingen: ten eerste steunen groene zoekmachines op deze manier commerciële verdienmodellen, waardoor ze de kapitalistische ideologie versterken (die op punten indruist tegen de groene, duurzame ideologie). Ten tweede dragen zij, bijvoorbeeld door indirect vervuilende servers te gebruiken, bij aan milieuvervuiling. Ten derde kunnen de grotere bedrijven wanneer ze maar willen uit de kleine, groene projecten stappen waardoor het kleine alternatief meteen ophoudt te bestaan. Dit gebeurde bijvoorbeeld bij de groene zoekmachine [Ecocho](http://en.wikipedia.org/wiki/Ecocho), die moest stoppen als gevolg van een aanpassing van Google's advertentiebeleid. Al met al zijn de groene zoekmachines net als privacyvriendelijke zoekmachines zeer afhankelijk van grote commerciële partijen.

Kennismachines

Dat geldt in mindere mate van de zogeheten kennismachines. Die nemen het overdragen van nieuwe kennis als uitgangspunt. In plaats van externe bronnen als zoekresultaten te leveren, waarbij gebruikers zelf de meest toepasselijke informatie moeten opzoeken, beantwoorden kennismachines vragen direct op een 'wetenschappelijke' manier, gebaseerd op facts and figures. Wolfram|Alpha is de bekendste kennismachine die rekenkundige en wetenschappelijke zoekvragen beantwoordt, bijvoorbeeld: 'Waar stierf Achilles?', 'Wat is de huidige aandelenkoers van Google' of 'Hoeveel neutronen zitten in een zuurstofatoom?'

Wolfram|Alpha gebruikt zijn eigen algoritme en verwerkt geen zoekresultaten van grote commerciële zoekmachines. Toch valt dit alternatief volgens Mager binnen de ideologie van het kapitalisme, want het oogmerk is winst, die wordt behaald door advertenties en de mogelijkheid een pro-account te nemen. Een ander punt dat Mager problematisch vindt, is dat Wolfram|Alpha copyright claimt op alle resultaten die het systeem levert op de ingevoerde zoekopdrachten. Dat betekent dat zelfs informatie die 'free' en open beschikbaar is, via de verwerking tot zoekresultaat tot 'gesloten' bezit van Wolfram|Alpha wordt gemaakt zodra iemand ernaar zoekt.

The screenshot shows the Wolfram|Alpha interface for the query 'cyan + yellow'. The search bar at the top contains the text 'cyan + yellow'. Below the search bar, the 'Input interpretation' section shows 'color blend' with two color swatches: cyan and yellow. The 'Result' section displays 'color red 0.5 green 1. blue 0.5'. A 'Color swatch' section shows a green color swatch. The 'Representations' section includes a table with the following data:

fractions	red 0.5 green 1. blue 0.5
24-bit RGB	128 255 128
HSB	120° 50% 100%
hexadecimal	#80FF80

The 'Nearest named HTML colors' section includes a table with the following data:

	HTML name	hexadecimal	24-bit RGB
	pale green	#98FB98	152 251 152
	light green	#90EE90	144 238 144

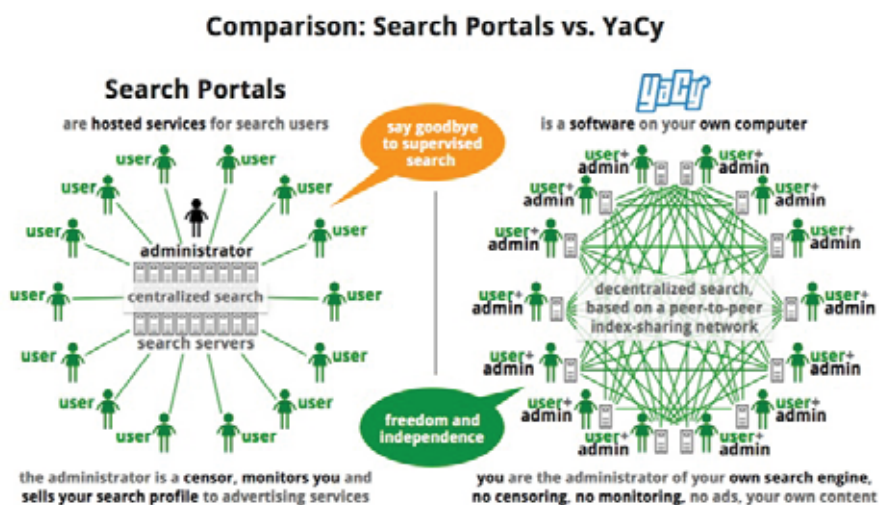
The 'Related colors' section shows 'complementary colors' with a green and purple swatch, and 'color triad' with green, blue, and red swatches. At the bottom, it says 'Computed by: Wolfram|Mathematica' and 'Download as: PDF | Live Mathematica'.

Afbeelding 3: Wolfram Alpha -

Some rights reserved,

<http://www.flickr.com/photos/roba/>

Peer-to-peer alternatieven: YaCy



Afbeelding 4: YaCy

Dit brengt ons bij YaCy, een alternatieve zoekmachine die in Magers optiek als het enige werkelijk onafhankelijke alternatief voor commerciële zoekmachines en hun kapitalistische ideologie kan worden beschouwd. YaCy is een gratis open source zoekmachine die vrije toegang tot informatie vooropstelt. Er is geen sprake van tracking, advertenties, censuur of surveillance. De zoekmachine werkt gedecentraliseerd en heeft een peer-to-peer netwerkstructuur. De index van het web, waaruit de zoekresultaten worden getoond, is opgebouwd en doorzoekbaar gemaakt door de individuele computers van YaCy-gebruikers. Al die computers staan met elkaar in verbinding zonder dat er een centrale server nodig is.

Het gebruik van gedecentraliseerde zoekmachines vereist knowhow, inspanning en geduld

Zoekmachines met een centrale netwerkstructuur, zoals Google, genereren zoekresultaten uit een index die is opgeslagen in massale datacentra, beheerd door een commerciële partij. Bij YaCy kan je zelf aan een decentrale globale zoekindex bijdragen door een webcrawler te installeren en op je computer te laten draaien.



Een webcrawler, ook wel spider (<http://nl.wikipedia.org/wiki/Spider>) genoemd, is een computerprogramma dat automatisch websites doorbladert en indexeert zodat deze met een zoekmachine kan worden doorzocht. Daarnaast biedt YaCy ook de mogelijkheid om je eigen zoekportaal te beheren en zelf te bepalen welke websites je op je computer indexeert en er doorzoekbaar maakt (zie Afbeelding 4). Een goede optie voor instellingen zoals bibliotheken die hun lokale database doorzoekbaar willen maken. In tegenstelling tot grote zoekmachines heeft YaCY alleen de stroom en rekenkracht van de computers van gebruikers nodig om te functioneren. Hierdoor draagt dit initiatief ook een steentje bij aan de groene agenda.

De onderliggende ideologie van YaCy is die van de 'commons'. De commons is een term die verwijst naar culturele en natuurlijke middelen die voor iedereen toegankelijk zijn en waar geen eigendom op berust. De ontwikkelaars van YaCy zijn duidelijk voorstanders van gratis software en de vrije toegang tot informatie. Mager stelt dat YaCy door deze visie en aanpak eerder aansluit bij een communistische dan een kapitalistische geest.

Het gebruik van een gedecentraliseerde zoekmachine als YaCy is voor de meeste mensen vrij complex. Technische knowhow, inspanning en geduld zijn nodig. Vooral dit laatste is ver te zoeken in de snel bewegende consumentencultuur, waarin iedereen gewend is aan het gemak van Google. Of YaCy uiteindelijk zal kunnen concurreren met Google zal afhangen van het aantal gebruikers, met andere woorden de human resources. Maar tijd en geld zijn ook belangrijke schaarse factoren. YaCy is in de ogen van Astrid Mager een prachtig alternatief voor Google. De vraag is: hoe springt de maatschappij er mee om?

Lees het hele artikel van Astrid Mager in de *Society of the Query Reader*:
http://networkcultures.org/query/wp-content/uploads/sites/4/2014/06/5.Astrid_Mager.pdf

Is Small Really Beautiful? Big Search and Its Alternatives (pdf).





Discussiepunten:

1: Opdracht: Bezoek de website <http://dontrack.us/> van DuckDuckGo. Wat zijn de belangrijkste voordelen van anoniem zoeken op het web en vind je die ook belangrijk?

2: Volgens Astrid Mager zijn privacyvriendelijke zoekinitiatieven niet alleen zeer afhankelijk van de praktijken van de grotere zoekmachinebedrijven, ze willen vaak ook net als hen geld verdienen. Denk je dat er veel mensen zijn die hiermee zitten? Wat is jouw mening over deze kwestie?

<http://trisquel.info/en/forum/problems-exist-duckduckgo-and-startpage>

3: Opdracht: Download YaCy en doe een poging om het programma te installeren en hiermee een zoekopdracht uit te voeren.

<http://yacy.net/en/index.html>

4: Het lijkt erop dat voor veel mensen de technische drempel te hoog ligt om YaCy te gebruiken. Om hier wat aan te doen heeft YaCy recentelijk samen met Obsidian Security Services een crowdfunding campagne opgezet voor de ontwikkeling van YaCyPi: een stukje hardware gebaseerd op Raspberry Pi, die het installeren en het beheren van YaCy zoekmachine een stuk makkelijker zou kunnen maken. Zou je overwegen om een eigen zoekmachine willen beheren?

https://www.kickstarter.com/projects/1455616350/yacypi-turnkey-raspberry-pi-based-internet-search/creator_bio

<https://www.kickstarter.com/projects/1455616350/yacypi-turnkey-raspberry-pi-based-internet-search>

<http://www.raspberrypi.org/>





Verder lezen en kijken:

- +** De lezing die Astrid Mager gaf op Society of the Query #2 is op het blog beschreven en terug te zien: 'Astrid Mager – Is Small Beautiful? Big Search and its Alternatives', Institute of Network Cultures, 11 november 2013.
<http://networkcultures.org/wpmu/query/2013/11/11/astrid-mager-is-small-beautiful-big-search-and-its-alternatives/>
<http://vimeo.com/79283301>

- +** Jeroen Geuens, 'YaCy daagt Google-zoekmachine uit', ZDNet, 29 november 2011.
<http://www.zdnet.be/internet/133718/yacy-daagt-google-zoekmachine-uit/>

- +** Tyler Handley, 'P2P Search as an Alternative to Google: Recapturing Network Value Through Decentralized Search', Journal of Peer Production: The Critical Power of Free Software.
<http://peerproduction.net/issues/issue-3-free-software-epistemics/peer-reviewed-papers/p2p-search-as-an-alternative-to-google-recapturing-network-value-through-decentralized-search/>

- +** Dustin Verburg, 'DuckDuckGo: The Slow Waddle of Progress', Squawk, 31 juli 2013.
<http://squawk.im/industry-news/duckduckgo/>





Pleidooi voor een onafhankelijke index van het web

Dirk Lewandowski



Wat is de beste manier om alternatieve zoekmachines ontwikkelen? Daarvoor zul je allereerst een goed werkende index van het hele internet moeten hebben. Het opzetten én bijhouden van zo'n index is echter zo kostbaar en tijdrovend, dat niemand zich er nog aan waagt, behalve Google en Bing (Microsoft). Dirk Lewandowski, professor aan de Hogeschool van Hamburg stelt in het artikel 'Why We Need an Independent Index of the Web' dat een onafhankelijke zoekmachine-index onmisbaar is om het monopolie in online zoeken te doorbreken. Die zou dienen als kweekvijver voor alternatieve zoekapplicaties en zo leiden tot meer diversiteit en eerlijke concurrentie.

Op dit moment is Google de marktleider als het gaat om zoeken op het web. Google gebruikt een bepaalde methode om data te verzamelen, verwerken en doorzoekbaar te maken, maar biedt geen transparantie over de gebruikte algoritmes. Het is onduidelijk hoe Google precies werkt en dat is problematisch, juist omdat we Google zoveel gebruiken en geen andere optie hebben. Ook volgens Lewandowski is er geen echt alternatief voor Google. Zelfs Microsoft kan maar met moeite haar zoekmachine Bing draaiende houden. Kleinere alternatieven kunnen realistisch gezien niet eens een begin maken met een goede index.

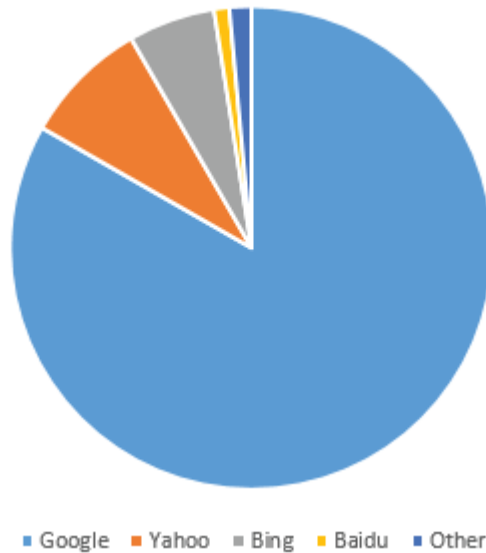
Alternatieve zoekmachines

Alternatieve zoekmachines, zoals beschreven door Astrid Mager, gebruiken vaak de web-index van andere, grotere zoekmachines en bieden dus eigenlijk alleen een alternatieve interface voor de zoekresultaten van de grote spelers. Vaak delen dit soort alternatieven in de inkomsten uit advertenties. Deze constructie noemt Lewandowski het 'partner index-model'. Alle grote webportalen passen dit model toe, zoals bijvoorbeeld Yahoo. Het is een gunstige situatie voor zowel webportal als zoekmachine, die echter ten gevolge heeft dat er niet veel concurrentie overblijft in de zoekmachine-industrie (zie Afbeelding 1).





Search Engine Market Share July 2013



Afbeelding 1: Zoekmachine marktaandeelen juli 2013

(bron <http://www.techwyse.com/blog/internet-marketing/search-engine-market-share-july-2013/>)

Er zijn veel pogingen ondernomen om alternatieve zoekmachines te bouwen, maar volgens Lewandowski is het werken aan één alternatieve zoekmachine niet zinvol. Die zal om verschillende redenen snel mislukken, bijvoorbeeld vanwege een verkeerde ontwerpbeslissing of een marketingstrategie die niet aanslaat. Bovendien is het opbouwen van een grootschalige zoekmachine-index een dure aangelegenheid en is het zeer moeilijk om vanaf nul te beginnen. Alternatieve zoekmachines die een eigen web-index willen beheren lopen een groot risico.

Eindelijk zou er eerlijke concurrentie op de zoekmachinemarkt kunnen ontstaan

Een index voor iedereen

Lewandowski pleit voor het bouwen van een alternatieve index die als basis kan dienen voor de ontwikkeling van diverse zoekmachines. We zouden niet aangewezen moeten zijn op maar een paar aanbieders van zoekmachines en daarom ook niet moeten focussen op het bouwen van een enkel alternatief. Het is beter om de omstandigheden te creëren die het mogelijk maken voor individuen, bedrijven en instellingen om hun eigen zoekmachines te bouwen. Dat betekent het faciliteren van een openbare web-index, waarbij zelf bepaald kan worden hoe de index wordt verwerkt. Dit zou voor een grotere diversiteit aan kleine zoekprojecten zorgen, die zelfs voorbij het reguliere zoeken op het web voeren; zoekapplicaties waarvan het nu nog onmogelijk is om ze te bedenken, aldus Lewandowski. Bovendien zou er ook eindelijk eerlijke concurrentie op de zoekmachinemarkt kunnen ontstaan.





De transnationale, door overheden gefinancierde index die Lewandowski voor ogen heeft is voor iedereen toegankelijk tegen redelijke voorwaarden. Het verwerken van een kleine hoeveelheid data uit deze index zou kosteloos moeten zijn, waardoor ook NGO's en non-profit organisaties ervan kunnen profiteren. Grootschalig gebruik van de data zou wel een prijs hebben.

Wat zou het kosten om de index op te tuigen? De exacte kosten zijn moeilijk in te schatten, maar het zou zeker gaan om een paar honderd miljoen euro. Het project kan dan ook niet door één land worden gefinancierd, maar vraagt om samenwerking door meerdere overheden in bijvoorbeeld EU-verband. Lewandowski vergelijkt de investering in een onafhankelijke index die informatie voor iedereen openbaar en toegankelijk maakt met het financieren van infrastructuur zoals snelwegen. De constructie van een publieke infrastructuur voor data is een taak die volgens hem is weggelegd voor de overheid.

Een nieuw op te richten of bestaand onderzoeksinstituut zal dan index moeten beheren. Lewandowski benadrukt dat de beheerders niet als enige beslissingen zouden mogen maken over de manier waarop de documenten worden geïndexeerd en beschikbaar worden gemaakt. Een democratische structuur is nodig, met bijvoorbeeld een adviesorgaan, bestaande uit (potentiële) gebruikers van de index.

Lees het hele artikel van Dirk Lewandowski in de *Society of the Query Reader*: http://networkcultures.org/query/wp-content/uploads/sites/4/2014/06/4.Dirk_Lewandowski.pdf

Why We Need an Independent Index of the Web (pdf).





Discussiepunten:

- 1: Dirk Lewandowski noemt de openbare web-index die hij voorstelt onafhankelijk. Kan een dergelijke index wel onafhankelijk zijn? Lijkt het jou verstandig om de overheid bij de ontwikkeling van een dergelijke web-index te betrekken?
- 2: Astrid Mager noemt in het artikel 'Is Small Really Beautiful? Big Search and Its Alternatives' de gedecentraliseerde zoekmachine YaCy. Sluit deze alternatieve zoekmachine wel of niet goed aan op de visie van Lewandowski?
- 3: Wat vind je van het voorstel dat voor intensief gebruik van de index betaald moet worden? Hoe zou Google hierop kunnen reageren?
- 4: Lewandowski gaat niet in op internetcensuur. Vind je het aannemelijk dat meerdere Europese overheden gaan samenwerken om een web-index te bouwen, zonder zeggenschap te eisen over de inhoudelijke informatie?

Verder lezen en kijken:

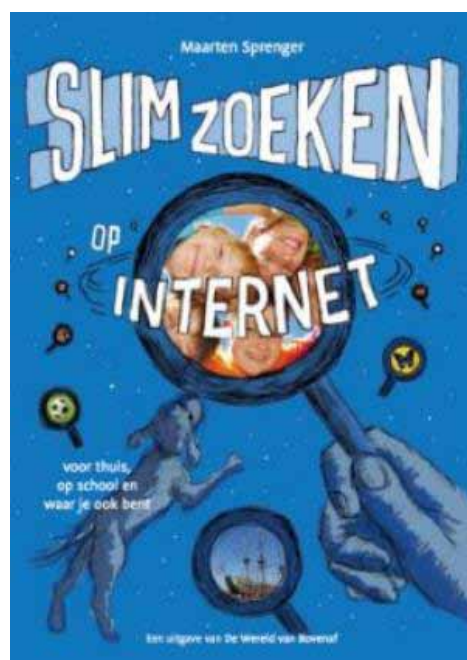
- + De lezing die Dirk Lewandowski gaf op Society of the Query #2 is op het blog beschreven en terug te zien: 'Dirk Lewandowski: Why We Need an Independent Index of the Web', Institute of Network Cultures, 11 november 2013.
<http://networkcultures.org/wpmu/query/2013/11/11/dirk-lewandowski-why-we-need-an-independent-index-of-the-web//>
<http://vimeo.com/79286779>
- + Dirk Lewandowski, 'New Perspectives on Web Search Engine Research', 2012.
http://www.bui.haw-hamburg.de/fileadmin/user_upload/lewandowski/doc/New_perspectives_on_Web_search_engine_research_preprint.pdf
- + Philipp Mayr en Dirk Lewandowski, 'Exploring the Academic Invisible Web', 2007.
<http://arxiv.org/ftp/cs/papers/0702/0702103.pdf>
- + Kevin J. O'Brien, en Thomas Crampton, 'Germany Quits Search Engine Project', International Herald Tribune, 2 januari 2007.
<http://www.nytimes.com/2007/01/02/business/worldbusiness/02iht-search.4081237.html>



Slim zoeken?!

Miriam Rasch

Alleen al omdat we zo vaak iets opzoeken online, denken we dat we daar heel goed in zijn. Maar klopt dat wel? Maarten Sprenger schreef het boek *Slim zoeken op internet – voor thuis, op school en waar je ook bent*, boordevol tips en tricks om veel meer uit het internet te halen. Ook beheert hij de zoekmachine voor kinderen 8-12.info (<http://www.8-12.info/>). Slim zoeken, stelt hij in het artikel Children's Information – Who Cares? (<http://networkcultures.org/wpmu/query/2014/02/13/childrens-information-who-cares-by-maarten-sprenger/>), betekent vooral ook: slim met informatie omgaan. De observaties die Sprenger maakt over online informatie voor kinderen kan iedereen zich ter harte nemen.



Slim zoeken op internet - voor thuis, op school en waar je ook bent

Het gaat dus niet zozeer om slim zoeken, maar om dat wat je vindt. Want je kunt nog zoveel links op je scherm krijgen – hoe weet je nu wat een goed resultaat is waar je wat aan hebt en waar je op kunt vertrouwen? Het allerbelangrijkste is het kunnen beoordelen van de informatie die je krijgt voorgeschoteld: is die betrouwbaar, begrijpelijk, waar komt de informatie vandaan, zijn er geen commerciële of politieke belangen? Pas als je op die manier al die links weet te evalueren, wordt het web een waardevolle bron van kennis.



Hoe doe je dat dan? Op school krijg je vaak alleen maar te horen: 'Google het maar'. De vraag stellen hoe dan wel, voelt al gauw een beetje dom. Als je de juiste zoektermen gebruikt moet dat toch geen probleem zijn? Maar wat zijn dan de juiste zoektermen? Het stellen van de juiste vraag is vaak even moeilijk als het vinden van het antwoord, dat weet iedereen die wel eens onderzoek heeft moeten doen.

Zelfs al heb je goede, neutrale en toch specifieke zoektermen te pakken, dan kun je makkelijk in de problemen raken. Een mooi voorbeeld daarvan geeft het onderzoek van René König en Erik Borra naar googelen op 9/11 (<http://networkcultures.org/wpmu/query/2012/10/01/googling-911-a-cross-cultural-comparison-of-suggestions-for-a-loaded-term/>), de aanslagen op onder andere de Twin Towers in New York. Het blijkt dat de resultaten en de volgorde daarvan in westerse landen sterk verschilt van die in Arabische landen. Waarom is dat relevant? Stel dat je onderzoek doet naar de vorming van complottheorieën na die aanslagen, dan is het noodzakelijk is om de verschillende bronnen van elkaar te kunnen onderscheiden. Je stelt jezelf vragen, bijvoorbeeld: gelooft de auteur zelf in de theorie, staat de informatie op een website van een universiteit of staan er links in naar andere bronnen? Bij 9/11 ligt dat misschien voor de hand, maar het is belangrijk om naar alle online informatie zo kritisch te kijken. Dat begint vaak al met de url van de website, die vertelt vaak meer dan je denkt.

Ook is het belangrijk om te begrijpen, in elk geval ongeveer, hoe een zoekmachine werkt en wat de relatie is tussen een zoekmachine en het internet. Google wordt maar al te vaak gezien als synoniem met internet, terwijl het een portaal is dat je toegang biedt tot een deel van het web, na uitgebreide selectie en filtering. Hoe Google selecteert en filtert weet niemand precies. Er zijn ook andere zoekmachines, die op een andere manier werken en het loont om voor een ander perspectief eens Bing of DuckDuckGo te gebruiken.





Maarten Sprenger schrijft vooral om kinderen te helpen goede informatie te vinden en om hun leerkrachten te helpen bij de ondersteuning van die zoektocht. Maar zijn conclusies zijn waardevol voor iedereen die soms verdwaalt op het wereldwijde web en toch niet zonder kan.

Dat is in vijf statements samen te vatten:

1. *Onderzoek de content die op het web wordt aangeboden, kwantitatief en kwalitatief.*

Hoeveel resultaten krijg je op een bepaalde vraag? Wat zou een normale hoeveelheid eigenlijk zijn? Soms krijg je geen enkel resultaat, soms zijn het miljoenen hits. En zijn al die hits even waardevol? Waarom zou de bovenste link eigenlijk als eerste genoemd worden? Is die link dat waard en wie bepaalt dat?

2. *Wees kritisch op de online bronnen die je krijgt aangeboden, ook van school.* Kritisch nadenken is altijd belangrijk, maar bij stof die online aangeboden wordt helemaal. Van welke site komt een artikel dat in je leeslijst staat? Kijk eens verder rond, is het een professionele kennisbank of een hobbyproject van een amateur? Of juist van een bedrijf dat iets probeert te verkopen? Een artikel lijkt misschien wetenschappelijk, maar voldoet het ook aan standaarden als controleerbaarheid en verwijfsbaarheid? Ook als de website goed is, moet je het bericht zelf kritisch bekijken. Vergeet niet dat zelfs serieuze nieuwsbronnen wel eens een zogenaamd 'nieuwsbericht' van satirische website De Speld voor waar aannemen en herplaatsen. Maar wees vooral beducht voor veelgebruikte sites waar iedereen zomaar informatie op kan zetten zonder ergens verstand van te hebben of zonder iets te onderbouwen, zoals bijvoorbeeld infonu.nl.





3. *Laat je niet afschepen met 'Google het maar'.*

'Google het maar' is een luie manier om iemand af te schepen en zegt vooral iets over de mate waarin Google inmiddels is ingeburgerd in het sociale verkeer. Er is zelfs een website met geanimeerde Google-zoekopdrachten, getiteld *Let me Google that for you* (<http://lmgty.com/>). Maar stel eens een tegenvraag: wat wil iemand precies weten, wat is de vraag? Hoe zou je docent of iemand anders zo'n zoekopdracht dan aanpakken? En waarom?

4. *Voeg zelf waardevolle content toe als dat kan (met bronvermelding).*

Scripties van studenten worden opgenomen in een scriptiedatabank, zodat andere studenten daar weer gebruik van kunnen maken. Als je zelf een onderzoek hebt gedaan naar een bepaald onderwerp en daar goede bronnen voor verzameld hebt, waarom zou je die dan niet delen? Veel van wat hier beschreven staat komt neer op het presenteren van informatie in een goede context. Welke vraag lag ten grondslag aan deze bronnen? Waarom zijn ze bruikbaar en waarom niet? Presenteer je een onderzoek, jouw mening, een opinie gebaseerd op expertise of weer iets heel anders? Alle informatie kan waardevol zijn, als maar duidelijk is wat de context is.

5. *Sla goede informatie op.*

Als iets eenmaal online staat, komt het er nooit meer vanaf, hoor je weleens. Een goede reminder voor als je op het punt staat een gênante foto op Facebook te zetten. Terwijl aan de andere kant een van de nadelen van online content juist is dat het ook zo weer kan verdwijnen. Discontinuïteit is een groot probleem bij het zoeken en vinden van goede informatie. Het is belangrijk dat goede informatie die kennis kan genereren beschikbaar blijft in de toekomst. Daar kun je zelf een steentje aan bijdragen door goede informatie op te slaan. Er zijn allerlei handige tools voor, van lees later-apps zoals Instapaper tot simpelweg de printknop.

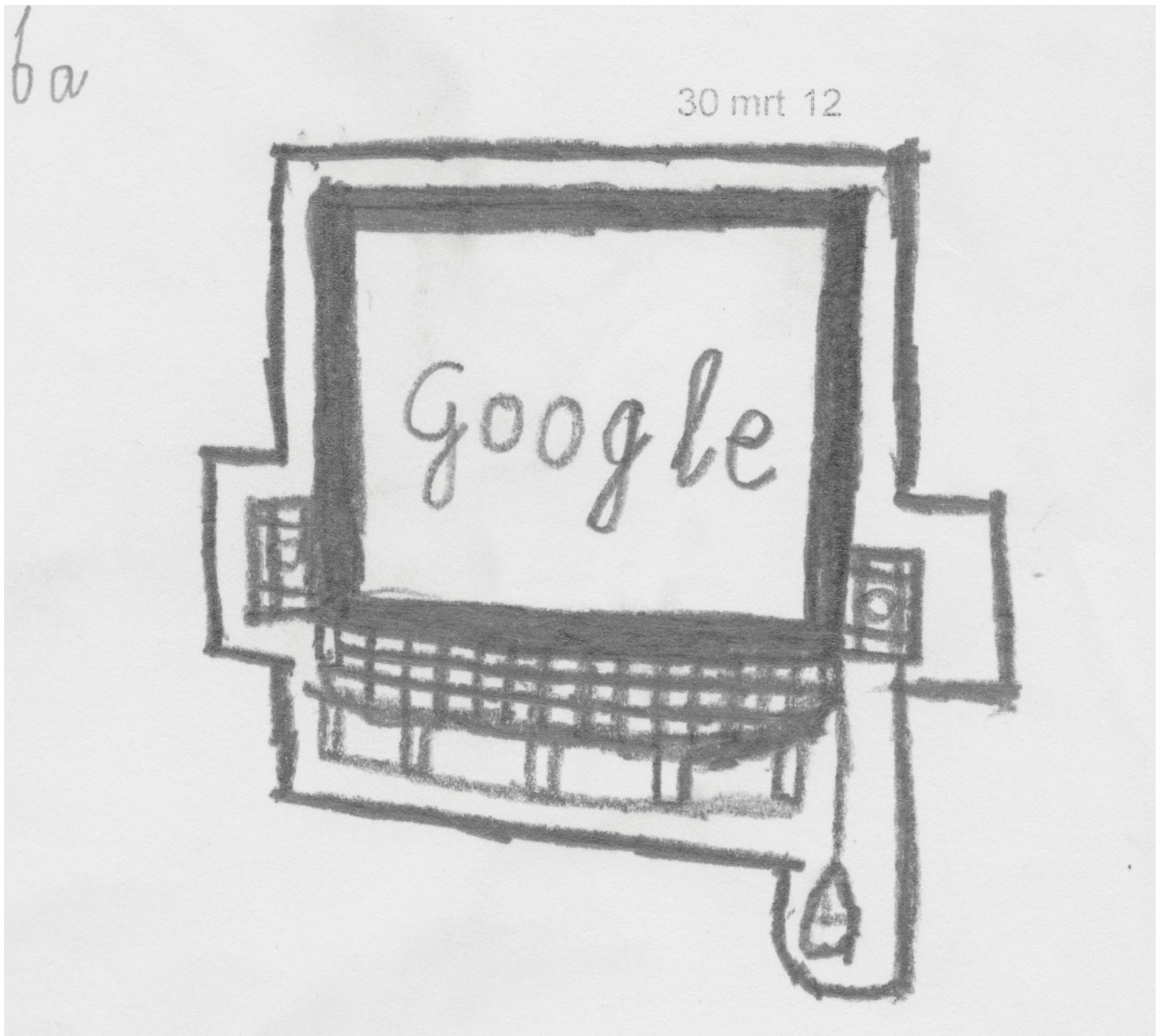




Meer slim zoeken-tips?

In het boek van Maarten Sprenger vind je er heel veel. Het is te bestellen via deze link (<http://www.sprengeronderwijs.nl/bestel/>). Het uitgebreide artikel is online beschikbaar.

<http://networkcultures.org/wpmu/query/resources/readings/childrens-information-who-cares-by-maarten-sprenger/>





Meer weten?

Miriam Rasch



Natuurlijk is het hiermee niet afgelopen. We blijven zoeken in de zee van online informatie, waarschijnlijk met gebruik van Google en hopelijk met gebruik van ons verstand. Society of the Query betekent ook: we zijn allemaal onderdeel van een samenleving die op zoek gaat. Naar antwoorden, maar ook naar de juiste manier om vragen te stellen.

Op de website van Society of the Query vind je heel veel materiaal over online zoeken en zoekmachines. Hier staan alvast wat dingen op een rijtje om je nieuwsgierig te maken.

De artikelen die voor dit magazine zijn bewerkt komen uit het boek *Society of the Query Reader: Reflections on Web Search* (red. René König en Miriam Rasch). Het boek is gratis te downloaden of te bestellen (<http://networkcultures.org/wpmu/portal/publication/society-of-the-query-reader-reflections-on-web-search/>), of lees online de pdf (http://issuu.com/instituteofnetworkcultures/docs/sotqreader_def_issuu?e=0/7587356).

Veel van de auteurs spraken tijdens de conferentie op 7 en 8 november 2013 over online zoeken. Al die lezingen zijn opgenomen en er is bovendien over elk van de praatjes een korte blogpost geschreven. Kijk bijvoorbeeld de actuele en meeslepende lezing terug van Siva Vaidhyathan over Google na Snowden: <http://vimeo.com/79282416>





Over Edward Snowden en de NSA is online heel veel te vinden. Een goede case study om slim zoeken op uit te proberen! Maar kijk bijvoorbeeld hier als je meer wilt weten (en vergelijk als je aan het lezen bent eens de berichten op kwaliteit en kwantiteit):

<http://www.volkskrant.nl/vk/nl/13524/Afluisterpraktijken-NSA/index.dhtml>

<http://zoeken.nrc.nl/?q=De+NSA-dossiers&sort=date>

Als je zelf waardevolle informatie vindt, vergeet die dan niet te delen! Je kunt contact opnemen met Miriam Rasch van het Instituut voor Netwerkcultuur via miriam@networkcultures.org





Colofon

Society of the Query Magazine. 10 artikelen over online zoeken

Redactie: Miriam Rasch en Marc Stumpel

Productie: Lars Böhm en Elin Wassenaar

Design: Studio Inherent

App-developer: MediaLAB Amsterdam in samenwerking met Martijn Walraven

Uitgever: Instituut voor Netwerkcultuur, Amsterdam

Ondersteund door: Hogeschool van Amsterdam, Amsterdam Creative Industries Publishing en Stichting Democratie en Media

Veel dank gaat uit naar de auteurs: Kylie Jarrett, Andrea Miconi, Mél Hogan, M.E. Luka, Dave Crusoe, Anton Tantner, Anna Jobin, Olivier Glassey, Min Jiang, Angela Daly, Astrid Mager, Dirk Lewandowski en Maarten Sprenger.

De oorspronkelijke artikelen verschenen in René König and Miriam Rasch (eds), *Society of the Query Reader: Reflections on Web Search*, Amsterdam: Institute of Network Cultures, 2014. ISBN 978-90-818575-8-1.

Lees, download of bestel het boek via networkcultures.org/publications

Deze publicatie verschijnt onder Creative Commons Attribution NonCommercial ShareAlike 3.0 Unported (CC BY-NC-SA 3.0), <http://creativecommons.org/licenses/by-nc-sa/3.0/>.

Institute of
network cultures



AMSTERDAM
CREATIVE
INDUSTRIES
CENTRE OF EXPERTISE



STICHTING
DEMOCRATIE
EN MEDIA

