



datum:  
23 december 2011

begeleider:  
Mirko Tobias Schäfer

MA Thesis

# Google's zoektocht naar u

Google zoeken

Ik doe een gok

Over Google's personaliseringsmechanismen, de gevaren ervan en de rol van Google+

Mijn dank gaat in het bijzonder uit naar:

Mirko T. Schäfer, Willy Poelhekke, Laura de Jongh, David van Toor, Thijs Hupkens en Michiel Stoter

1. Inleiding	4
2. De overgang naar gepersonaliseerd zoeken	6
2.1. De geschiedenis van zoekmachines	6
2.2. Search engine bias	8
2.3. Personalisering	9
2.3.1. Het belang van personalisering op individueel niveau	10
2.3.2. Het belang van personalisering op commercieel niveau	11
3. Personaliseren in de praktijk	13
3.1. Googlization	13
3.2. De ontwikkeling van een persoonlijk profiel	15
3.2.1. Het kennisprofiel van de gebruiker	16
3.2.2. Het sociale profiel van de gebruiker	18
3.2.3. Het fysieke profiel van de gebruiker	18
4. De gevaren van personaliseren	20
4.1. Beperking van privacy	20
4.2. Beperking van autonomie	22
5. De volgende stap in personalisering: Google+	25
5.1. Identiteitsconstructie	25
5.2. Google Circles	26
5.3. Google's +1 knop	28
6. Conclusie	31
7. Referenties	33

## 1. Inleiding

Na een aantal jaren gewerkt te hebben aan een algoritme dat een doorbraak moest zijn in de zoekmachinewereld, lanceerden Sergey Brin en Larry Page in 1998 de zoekmachine Google (google.com n.d. a). Zij stelden destijds: “Google is designed to crawl and index the Web efficiently and produce much more satisfying search results than existing systems” (Brin & Page 1998, p. 1). Google was als zoekmachine een relatieve laatkomer, maar onder andere vanwege haar overzichtelijke startpagina, het gebrek aan advertentiebanneren en de door Brin en Page beloofde kwaliteit van de zoekresultaten werd Google al snel de populairste zoekmachine ter wereld. Inmiddels is het bedrijf google.com met een merkwaarde van 114 miljard dollar uitgegroeid het meest waardevolle merk ter wereld (BrandZ 2010). Dit succes is naast hun zoekmachine ook te danken aan tal van andere diensten en producten, zowel on- als offline. Hierbij kan gedacht worden aan haar e-maildienst Gmail en de online atlas Google Maps, maar ook aan Google’s mobiele besturingssysteem Android waarmee Google een sprong heeft gemaakt naar de markt van de mobiele telefonie.

Google heeft binnen verschillende markten een bijna onaantastbare monopoliepositie. Denk naast de online-zoekmarkt bijvoorbeeld aan de online-advertentiemarkt (Ingram 2011a) en de opkomst van Android. Hierdoor wordt haar motto “Don’t be evil” (google.com 2009) door critici vaak als tegenstrijdig ervaren (Diaz 2005; Lovink 2009; Vaidhyanathan 2011). Door Google’s dominante positie als toegangspoort tot informatie en het feit dat Google in staat is om grote hoeveelheden aan (privé)informatie over haar gebruikers op te slaan en te gebruiken voor commerciële doeleinden, wordt Google vaak monopolisme en het plegen van privacyschending verweten (Hoofnagle 2009). De monopoliepositie en de tegenstrijdigheid in de uitstraling en het beleid van het bedrijf maken Google tot een interessant onderzoeksobject binnen deze thesis.

Door de jaren heen is Google’s algoritme PageRank eveneens aan de nodige kritiek onderhevig geweest (Diaz 2005; Introna & Nissenbaum 2000; Lovink 2008; Metahaven 2009; Pasquinelli 2010). Doordat het algoritme reeds populaire websites zou bevoordelen, zou er steeds meer sprake zijn van een centralisering en commercialisering/popularisering van het Web, het zogenaamde ‘rich get richer’ fenomeen. Het personaliseren van zoekresultaten is een door critici veel geopperde oplossing voor dit probleem, aangezien hierdoor beter rekening wordt gehouden met de heterogeniteit van de gebruikers (E. Goldman 2008; Keenoy & Levene 2005). Bovendien zou hierdoor de relevantie van zoekresultaten per gebruiker vergroot kunnen worden en marginale informatie beter vindbaar zijn.

Het verzamelen van persoonsgegevens is voor Google niet enkel van belang op het gebied van zoekresultaten, maar levert ook grote voordelen op voor Google als advertentiebedrijf. Zo is het in staat om gepersonaliseerde advertenties te tonen aan gebruikers waardoor de ‘clickthrough rate’<sup>1</sup> van deze advertenties kan worden vergroot. Alhoewel het personaliseren van zoekresultaten voor Google en haar gebruikers grote voordelen kan hebben, kleven er ook nadelen aan op het gebied van privacy en de autonomie van de gebruikers.

---

<sup>1</sup> Clickthrough rate: Het percentage van het aantal klikken op een advertentie gedeeld door het aantal keer dat de advertentie getoond is. Aan de hand van de clickthrough rate kan het succes van een advertentiecampagne worden afgelezen.

In september 2011 is Google's nieuwste dienst Google+ gelanceerd. Deze 'social network site'<sup>2</sup> kan vanuit het perspectief van personalisering gezien worden als een logische stap van Google om een nog nauwkeuriger beeld te krijgen van haar gebruikers. Door te focussen op de sociale connecties van haar gebruikers en hen steeds meer te positioneren binnen hun fysieke leefomgeving, maakt Google een stap voorwaarts in de constructie van een persoonlijk gebruikersprofiel. Dit gebruikersprofiel, als het ware de digitale representatie van een gebruiker, wordt door Google's algoritmen gebruikt om per gebruiker de relevantie van zoekresultaten te berekenen en heeft dus grote invloed op de toegankelijkheid van bepaalde informatie. De nauwkeurigheid ervan is dus van essentieel belang.

Het doel van deze thesis is om de personaliseringsmechanismen van Google te verkennen en te kijken of de kritiek binnen het kritisch discours rondom het personaliseren van zoekresultaten nog steeds valide is met de komst van Google+.

Allereerst zal in het tweede hoofdstuk vanuit een historisch perspectief gekeken worden naar de ontwikkeling die Google heeft doorgemaakt en zal het nut van personalisering op individueel en commercieel niveau uiteengezet worden. Vervolgens zal in hoofdstuk drie gekeken worden naar de wijze waarop Google door middel van haar diensten persoonlijke informatie van gebruikers verzamelt en aan de hand daarvan een persoonlijk gebruikersprofiel opstelt. Dit zogenaamde "service-for-profile" (Rogers 2009) model dat Google hanteert wordt aan de hand van de - binnen het Google discours veel voorkomende - term "googlization" (Salveker 2003) beschreven. Daarna wordt beschreven welke methoden en diensten Google hanteert binnen dit model. Naar aanleiding van de grootschalige verzameling en opslag van persoonlijke informatie zal in hoofdstuk vier ingegaan worden op de vraag naar welke fundamentele problemen personalisering met zich mee brengt. Binnen het kritisch discours hieromtrent wordt de nadruk gelegd op privacyschending en de afname van autonomie bij gebruikers tijdens hun zoektocht naar informatie. Tot slot zal in hoofdstuk 5 de nieuwe dienst Google+ en een aantal van haar belangrijkste functies worden beschreven. Hierbij wordt gekeken of deze dienst een positieve dan wel negatieve invloed heeft op de waarden privacy en autonomie.

---

<sup>2</sup> Social network site: Een website waar gebruikers een persoonlijke pagina aan kunnen maken en contact kunnen leggen met andere personen binnen dat sociale netwerk.

## 2. De overgang naar gepersonaliseerd zoeken

Om haar dominante positie te behouden is Google constant bezig met het verbeteren van haar algoritmes en diensten. In de afgelopen jaren is er een duidelijke verschuiving zichtbaar geweest binnen de zoekmachinewereld. Zo wordt de gebruiker door zoekmachines steeds meer betrokken in het berekenen van de relevantie en dus de positionering van zoekresultaten. Hierbij zijn niet enkel de interesses van de gebruiker van belang, maar ook de sociale connecties en de handelingen buiten de virtuele wereld worden meegenomen in het ontwikkelen van een persoonlijk gebruikersprofiel. Binnen dit hoofdstuk zal vanuit historisch perspectief de opkomst van Google en de overgang naar het gepersonaliseerd zoeken worden belicht.

### 2.1. De geschiedenis van zoekmachines

Om het succes van Google binnen de zoekmachinewereld te kunnen duiden zal eerst kort gekeken worden naar de context waarin zoekmachines hun opmars tot poortwachters binnen het World Wide Web (WWW) hebben kunnen realiseren. Vervolgens zal vanuit een technisch perspectief de dominantie van Google worden verklaard.

Door de snelle opkomst van het WWW in de jaren '90 en de enorme hoeveelheid aan informatie die daarmee beschikbaar kwam, ontstond er een steeds grotere behoefte aan een manier om al deze informatie te organiseren en bereikbaar te maken. De wijze waarop dat voorheen gebeurde in bibliotheken, namelijk het centraal indexeren en ordenen van documenten aan de hand van categorie, auteur of inhoud (middels de samenvatting), bleek niet meer toereikend. Ten eerste omdat het WWW geen lijst aanbiedt waarop alle webpagina's vermeld staan. En ten tweede aangezien de inhoud of het onderwerp van webpagina's niet vast staat (Diaz 2005). Tijdens de beginfase van het WWW werden er door verschillende internetgebruikers en instituties handmatig lijsten opgericht waarin gelinkt werd naar verschillende webpagina's, documenten en software, maar al snel bleek de enorme toevoer van informatie op deze manier niet meer bij te houden (Fuez, Fuller & Stalder 2011; Halavais 2009; Rieder 2009).

De oplossingen voor de problemen van de klassieke wijze van indexeren zijn vandaag de dag nog steeds zeer belangrijke punten waarop zoekmachines zich van elkaar onderscheiden, namelijk op het gebied van 'crawling' en 'ranking'. Crawling is het op systematische wijze indexeren van webpagina's met behulp van een crawler oftewel een computerprogramma dat als het ware het web rondkruipt. Een crawler begint zijn speurtocht op een bepaalde webpagina, slaat deze op in een database en volgt van daaruit alle hyperlinks naar andere pagina's die het tegenkomt en slaat deze vervolgens ook op in de database. Door dit proces continu te herhalen ontstaat er op den duur een uitgebreid netwerk aan webpagina's die door middel van hyperlinks met elkaar verbonden zijn. Hierbij moet vermeld worden dat slechts een relatief klein deel van het WWW geïndexeerd wordt door zoekmachines. Dit gedeelte van het Web wordt ook wel 'The indexable Web' genoemd (Bergman 2001). Het overgrote deel van het Web wordt echter niet bezocht door 'crawlers' omdat dit bestaat uit dynamisch gegenereerde pagina's die niet rechtstreeks toegankelijk zijn voor zoekmachines. Dit gedeelte van het Web wordt aangeduid met de term 'The deep Web' (ibid.).

Om tijdens een zoekopdracht van een gebruiker de meest relevante pagina's bovenaan te laten verschijnen moet de inhoud van iedere pagina geanalyseerd worden. Dit principe wordt aangeduid met de term ranking. Een zo goed mogelijk rankingsysteem is van essentieel belang voor het voortbestaan van een

zoekmachine en bepaalt het wezen van de zoekmachine. Gebruikers verwachten relevante zoekresultaten in de eerste 10 tot 20 resultaten en zodra ze deze niet krijgen ervaren zij hun zoekopdracht als mislukt (Goldman 2006; Introna & Nissenbaum 2000; Pan et al. 2007). De ranking-algoritmes die door zoekmachines gebruikt zijn, zijn door de jaren heen flink aan verandering onderhevig geweest.

Het eerste volledig geautomatiseerde ranking-systeem werd geïmplementeerd door Altavista. Deze baseerde de positionering van een webpagina aan de hand van bepaalde documenteigenschappen zoals bijvoorbeeld de aangetroffen hoeveelheid en de positie van zoektermen (Rieder 2009, p. 138). Ook focuste Altavista op zogenaamde meta tags: informatie in de 'header'<sup>3</sup> van de webpagina waarin de paginabeheerder relevante trefwoorden en documentinformatie kan toevoegen. Dit systeem bleek echter al snel fraudegevoelig te zijn omdat paginabeheerders geen relevante maar populaire zoektermen zoals Britney Spears en seks gingen gebruiken om hun pagina vaker gevonden te laten worden. Bovendien kende Altavista geen relevantiewaardering toe aan de verschillende webpagina's waardoor de zoekresultaten niet optimaal waren (van der Vlies 2008, p. 20-21). Lycos loste deze problemen een aantal jaar later grotendeels op door de inhoud van een webpagina vast te stellen aan de hand van de 'anchor text'<sup>4</sup> van een hyperlink en op grond daarvan de positionering van een webpagina te bepalen (Battelle 2005, p. 53). Daarnaast maakte Lycos gebruik van het 'backlink counting' principe. Door het aantal verwijzingen naar een bepaalde pagina op te tellen kon hiermee de importantie van een pagina bepaald worden (Diaz 2005, p. 64). Ook dit systeem bleek echter niet waterdicht omdat het redelijk spam gevoelig was. Zo konden er een groot aantal willekeurige pagina's worden gecreëerd die naar eenzelfde pagina verwezen om zo de importantie van die desbetreffende pagina te vergroten. Google heeft het backlink counting principe van Lycos verder uitgewerkt door er vanuit te gaan dat niet iedere backlink even belangrijk is. Google's recursieve PageRank algoritme is gebaseerd op het citeersysteem dat gebruikt wordt in de academische wereld. Daarin wordt de importantie van een academische tekst bepaald aan de hand van het aantal teksten dat daaraan refereert. Google berekent de PageRank van een pagina op soortgelijke wijze op basis van het aantal hyperlinks dat naar een pagina verwijst. Deze hyperlinks worden zwaarder meegewogen in het bepalen van de PageRank wanneer de pagina waarop deze hyperlinks geplaatst zijn zelf ook een hoge PageRank heeft. Een x aantal backlinks vanuit sites met een lage PageRank levert dus voor een pagina een minder hoge PageRank op dan dat zelfde aantal backlinks vanuit pagina's met een hoge PageRank (ibid., p. 64-65). Sergey Brin en Lawrence Page (1998) stellen dat hierdoor de PageRank van een pagina de waarschijnlijkheid dat een willekeurige bezoeker op deze pagina terecht komt representeert.

We assume there is a 'random surfer' who is given a web page at random and keeps clicking on links, never hitting "back" but eventually gets bored and starts on another random page. The probability that the random surfer visits a page is its PageRank (Brin & Page 1998, p. 4).

Op deze manier wordt het Web op hiërarchische wijze in kaart gebracht waardoor de ranking van relevante zoekresultaten grotendeels gebaseerd wordt op de positie van een webpagina binnen deze hiërarchie. Er

---

<sup>3</sup> Header: Het gedeelte van een webpagina waarin de stijlen, scripts en meta informatie worden gedefinieerd.

<sup>4</sup> Anchor text: De tekst van een link die zichtbaar is voor de gebruiker en waarop geklikt kan worden.

moet hier opgemerkt worden dat er naast het PageRank algoritme en de personalisering van zoekresultaten tal van andere factoren meespelen bij het bepalen van de ranking - zoals bijvoorbeeld het aantal in- en uitgaande links, de leeftijd van het domein en de technische opbouw van de webpagina (Stamoulis 2010). Vanwege de scope van dit onderzoek zal hier niet verder op worden ingegaan.

## 2.2. Search engine bias

Google vergelijkt de wijze waarop zoekresultaten geordend worden met een democratisch proces omdat een backlink gezien kan worden als een democratische stem van een pagina beheerder op de pagina waar hij naar verwijst (Rieder 2009, p. 138). Desalniettemin is er veel kritiek geuit op de wijze waarop het PageRank algoritme functioneert.

Alejandro Diaz (2005), wetenschapper aan de Stanford Universiteit, stelt dat Google hier een belangrijke voorwaarde voor democratie over het hoofd ziet, namelijk dat iedere stem even zwaar weegt (Diaz 2005, p. 69). Doordat Google meer waarde hecht aan de backlinks die van webpagina's met een hoge PageRank afkomen, zorgt ze ervoor dat reeds populaire websites een grotere invloed hebben in het bepalen van welke andere websites een hoge ranking krijgen. Diaz stelt dat "[t]his may allow a prominent site to dominate discourse on a given subject by linking extensively to self-reinforcing, like-minded pages" (ibid. p. 74).

Een ander veel gehoord punt van kritiek dat dit proces versterkt is het rich get richer fenomeen. Doordat grote reeds populaire websites vanwege hun hoge PageRank eerder in beeld zijn, neemt de kans ook toe dat er nog vaker naar hen verwezen wordt. Hierdoor worden populaire websites steeds populairder en verdwijnen minder populaire websites die marginale standpunten vertegenwoordigen uit het zicht (ibid., p. 71-73). Uit onderzoek van de Hongaarse netwerkwetenschapper Albert-László Barabási (2002) blijkt dat de spreiding van backlinks voldoet aan de zogenaamde 'power law'. Hij stelt dat:

the architecture of the World Wide Web is dominated by a few very highly connected nodes, or hubs. These hubs [...] are extremely visible - everywhere you go, you see another link pointing to them. In the network behind the Web many unpopular or seldom noticed nodes with only a small number of links are held together by these few highly connected Websites [sic] (Barabási 2002, p. 58).

Dit houdt in dat een klein aantal websites in staat is om beeldvorming rondom bepaalde onderwerpen te domineren en dat de notie van een vrij, open en democratisch internet oftewel het internet als een publiek goed - wat door vele theoretici en 'cyber libertarians' tijdens de opkomst van het WWW werd toegejuicht (Barlow 1996, Kapor 1993; Katz 1995; Rheingold 1993) - achterhaald is gebleken. Lucas Introna en Helen Nissenbaum (2000) stellen in hun invloedrijke artikel rondom 'search engine bias': *Shaping the Web: Why the Politics of Search Engines Matters* dat het door zoekmachines systematisch bevoordelen van websites met een mainstream of commercieel karakter ten koste gaat van websites die hier niet onder vallen. De keuzevrijheid op het internet wordt hierdoor beperkt. Dit is niet enkel een probleem voor degenen die op zoek zijn naar informatie, maar eveneens voor hen die via het web gevonden willen worden (Introna & Nissenbaum 2000, p. 19-20).



Volgens Introna en Nissenbaum zijn zoekmachines gebaat bij het focussen op populariteit, aangezien zij hierdoor een zo groot mogelijke groep aanspreken met hun zoekresultaten. Dit komt ten goede aan hun populariteit en dus aan hun marktaandeel. Daarbij komt het feit dat invloedrijke spelers binnen het web extra macht kunnen uitoefenen door een toppositie binnen de zoekresultaten te kopen (ibid., p. 14).

De jurist Eric Goldman (2006) gaat in zijn artikel *Search Engine Bias and the Demise of Search Engine Utopianism* in op deze veel gehoorde kritiek op het PageRank algoritme. Hij stelt dat search engine bias een “unavoidable consequence of search engines exercising editorial control over their databases” (E. Goldman 2006, p. 195) is. Wanneer er niet een dergelijke vorm van categorisering plaats vindt, dan wordt het systeem overgenomen door spammers en fraudeurs (Gilbert in E. Goldman 2006, p. 196). Goldman stelt daarbij dat de marktwerking tussen zoekmachines sterk genoeg is en ervoor zorgt dat gebruikers overstappen naar een andere zoekmachine wanneer zij het idee hebben dat zij niet de resultaten krijgen die ze wensen (E. Goldman 2006, p. 196). Goldman gaat hier echter voorbij aan het feit dat zoekmachines niet enkel gebruikt worden om te zoeken, maar ook gebruikt worden door paginabeherders om gevonden te worden door het grote publiek. Daarnaast moet opgemerkt worden dat Google inmiddels een monopoliepositie heeft en dat haar technieken en services door andere bedrijven als voorbeeld worden gebruikt (Rogers 2009), waardoor de keuzevrijheid voor gebruikers steeds verder verkleind wordt. Goldman erkent desalniettemin dat het soort ‘one-size-fits-all’ algoritmes zoals PageRank niet meer voldoet aan de eisen die de zeer heterogene groep van zoekmachinegebruikers stelt. Hij voorziet de oplossing in het personaliseren van zoekresultaten door zogenaamde ‘personalized ranking algorithms’ (E. Goldman 2006, p. 198-9). Volgens Goldman zorgt personalisering ervoor dat er voor een bepaalde zoekterm meerdere topresultaten zijn, in plaats van dat er één website altijd bovenaan staat. Daarnaast wordt de invloed op zoekresultaten van ranking-algoritmes zoals PageRank verkleint doordat er meer waarde gehecht wordt aan de voorkeuren van de gebruiker. Hierdoor kan de dominantie van populaire websites enigszins worden verkleint (ibid.).

### [2.3. Personalisering](#)

Goldman staat niet alleen in dit standpunt. Een veel geopperde oplossing voor search engine bias is het ordenen van zoekresultaten aan de hand van een persoonlijk gebruikersprofiel (Keenoy & Levene 2005; Zimmer 2008a). Dit persoonlijk profiel wordt onder andere opgebouwd uit persoonsgegevens, sociale gegevens en locatiegegevens<sup>5</sup>. De zoekresultaten worden als het ware door een tweede filter gehaald waardoor ze worden geordend aan de hand van het bestaande profiel van de gebruiker. De mediawetenschappers Felix Stalder en Christine Mayer (2009) beschrijven dit fenomeen als “The Second Index”. Alhoewel een persoonlijk gebruikersprofiel de positionering van zoekresultaten relevanter kan maken voor de gebruiker, wordt het door zoekmachines ook vooral toegepast om advertenties toe te spitsen op de gebruiker. Stalder en Mayer stellen dat:

What might be relevant on an aggregate, topological level could well be irrelevant on the level of the individual search interest. This problem is compounded by the fact that the actual customers of the search engines, the advertisers, are not interested in the topology

---

<sup>5</sup> In het volgende hoofdstuk zal dieper worden ingegaan op de personaliseringmethoden die Google hanteert.

of the network either, but in the individual users journeying through the network. These two forces, one intrinsic to the task of search engines themselves, the other intrinsic to what has become their dominant business model - advertising - drive the creation of the second index (Stalder & Mayer 2009, p. 98-99).

Hiermee wordt ook meteen duidelijk wat het belang van zoekmachines is bij het opstellen van een zo compleet mogelijk gebruikersprofiel. Alhoewel veel critici stellen dat het Google voornamelijk om gepersonaliseerde advertenties te doen is (Lovink 2008; Röhle 2007), zal hieronder het belang van het opstellen van een gebruikersprofiel door Google op zowel het individueel niveau van de gebruiker als het commercieel niveau van Google worden besproken.

### 2.3.1. Het belang van personalisering op individueel niveau

Het filteren van informatie is een proces dat in allerlei verschillende segmenten van de maatschappij voorkomt. Binnen de media gebeurt dit niet enkel op het internet. Ook binnen de traditionele media wordt informatie gefilterd voor de eindgebruiker door bijvoorbeeld redacties en journalisten. Goede informatievoorziening is essentieel voor mensen om een rationeel bestaan te leiden en om volwaardig te kunnen participeren in de maatschappij en draagt dus bij aan de vrijheid en autonomie van een individu (Hoven & Rooksby in Bozdag & Timmermans 2011). Om tot goede informatievoorziening te komen is het noodzakelijk dat het overvloedige aanbod versmald wordt. Van belang is hierbij de wijze waarop deze informatie gefilterd wordt en in hoeverre de informatie die door de filter heen komt relevant en betrouwbaar is. Dit geldt eveneens voor het filtermechanisme voor zoekmachines.

De Amerikaanse filosoof en cognitieve wetenschapper Alvin Goldman (2008) gaat in zijn artikel *The Social Epistemology of Blogging* in op de wijze waarop informatie vergaart wordt en democratische standpunten gevormd worden. Hierbij beschrijft hij het belang van correcte informatie en het belang van een goed filtersysteem dat deze informatie kan verzorgen. Hij stelt het volgende:

Which factors determine how well citizens acquit themselves in getting relevant information or knowledge on political matters [...]? This partly depends on citizens themselves, [...] [b]ut it also depends partly on the institutional structures used in the communication or transmission of information and misinformation (A. Goldman 2008, p. 3).

Goldman onderscheidt drie verschillende rollen binnen het standaardconcept van filteren met een communicatiekanaal en een groep mensen: “senders”, mensen die een boodschap willen versturen; “receivers”, mensen die wellicht een verstuurd boodschap ontvangen; en tot slot de “filterer” of “gatekeeper”, een individu of groep die de macht heeft om te bepalen welke boodschappen door het communicatiekanaal verstuurd worden en welke niet. Zodra een gatekeeper een bericht niet doorlaat dan is dit filteren (ibid. 6). Hierbij moet opgemerkt worden dat in het geval van het ordenen van zoekresultaten het niet per definitie het geval hoeft te zijn dat deze boodschappen letterlijk worden weggefilterd, maar veeleer

dat deze lager in de zoekresultaten komen te staan, waardoor de kans significant kleiner is dat deze door de gebruiker worden opgemerkt (Introna & Nissenbaum 2000).

Goldman gaat verder in op de wijze waarop mensen een rationeel en kritisch oordeel vormen op basis van de informatie die voorhanden is, hetgeen vervolgens leidt tot “belief” oftewel het geloven en innemen van een bepaald standpunt. Hij benoemt drie stadia binnen dit proces: de “reporting stage”; de “reception stage” en uiteindelijk de “acceptance stage”. Het filteren van boodschappen kan zowel in de reporting als in de reception stage plaatsvinden. Wanneer er gefilterd wordt tijdens de reporting stage dan wordt er gefilterd door de gatekeeper voordat een boodschap bij de ontvanger aankomt. Deze vorm van filteren vindt van oudsher plaats binnen de media aangezien traditionele media veelal beperkt waren in omvang en dus hun informatie moesten selecteren. Filteren tijdens de reception stage houdt in dat de ontvanger zelf een selectie maakt van de boodschappen die op hem afkomen. Dit doet hij door bepaalde boodschappen te selecteren en te consumeren om vervolgens in de acceptance stage deze boodschappen tegen elkaar af te wegen en te besluiten welke boodschappen te geloven (A. Goldman 2008, p. 7).

Het probleem van een beperkte omvang zoals traditionele media dat kennen geldt niet voor het Web, waardoor het filteren ook tijdens de twee laatste stadia kan plaatsvinden. Er kan gesteld worden dat er op het Web door het personaliseren van zoekresultaten een samensmelting zichtbaar is van het filteren tijdens de reporting stage met het filteren tijdens de reception stage. Dit komt omdat ontvangers (ongemerkt) steeds meer invloed hebben op de filteralgoritmes van de gatekeeper oftewel de zoekmachine. Desalniettemin stelt Goldman dat: “[i]f receivers are poor at estimating the reliability of the channels over which messages are sent, they won’t do a very good filtering job at the reception stage” (ibid.). Ook stelt hij dat “[i]f receivers are poor at estimating the accuracy of the specific messages they receive and read, they also won’t do a very good filtering job at the acceptance stage” (ibid.). Hierbij komt ook het gevaar van zogenaamde “echo chambers”, een metafoor voor het feit dat mensen geneigd zijn om informatie te filteren op basis van waar ze het wel of niet mee eens zijn (Sunstein 2007, p. 51-7). Enerzijds moeten filters dus goed bestand zijn tegen deze invloeden vanuit de ontvanger. Maar anderzijds is het van groot belang dat het filterproces geen relevante en afwijkende informatie achterhoudt voor de ontvanger omdat deze anders wellicht tot een ander standpunt zou zijn gekomen.

### 2.3.2. Het belang van personalisering op commercieel niveau

Naast het feit dat het personaliseren van zoekresultaten voordelige effecten kan hebben op de zoekopdracht van de gebruiker, wordt het implementeren van een persoonlijk gebruikersprofiel door zoekmachines ook voornamelijk ingegeven door het feit dat zoekmachines daardoor in staat zijn om gerichtere advertentiemethodes toe te passen. In 2000 introduceerde Google haar eigen businessmodel dat gebaseerd was op ‘targeted advertisement’. Met de succesvolle diensten AdWords en AdSense stelt Google gebruikers in staat om heel eenvoudig op basis van het Pay-per-click<sup>6</sup> (PPC) systeem advertenties te plaatsen naast zoekresultaten en op websites die betrekking hebben op de door de adverteerder ingevoerde trefwoorden. Bij de laatstgenoemde ontvangt de beheerder van de website waarop advertenties geplaatst worden zelf ook

---

<sup>6</sup> Pay Per Click: Adverteerders hoeven pas te betalen wanneer er op hun advertentie geklikt wordt. Adverteerders kunnen zelf per advertentie een maximumbedrag opgeven dat betaald moet worden wanneer er op de advertentie geklikt wordt. Hoe hoger dit bedrag, des te beter de adverteerder kan concurreren met andere adverteerders op hetzelfde trefwoord.

een deel van de advertentieopbrengsten. In 2010 produceerden de online advertenties van Google 96% van de totale inkomsten van het bedrijf (google.com 2011a). Hieruit kan opgemaakt worden dat het beleid van Google voornamelijk gericht is op het vertonen van advertenties. Martin Fuez, Matthew Fuller en Felix Stalder (2011) stellen dan ook dat Google zichzelf heeft getransformeerd naar een advertentiebedrijf dat niet zozeer zoekresultaten, maar veeleer haar gebruikers als primaire handelswaar heeft. De zoekresultaten (en alle overige diensten) worden volgens Fuez et al. door Google enkel gebruikt om deze gebruikers aan zich te binden (p. 2).

De Duitse mediawetenschapper Theo Röhle (2007) benoemt in zijn artikel *Desperately seeking the consumer: Personalized search engines and the commercial exploitation of user data* twee volgens hem fundamentele problemen van traditionele vormen van adverteren zoals die in kranten en op TV. Hij stelt dat de traditionele methoden uit gaan van onnauwkeurige demografische modellen die lastig zijn aan te passen aan de steeds verdere gedifferentieerde consumptiepatronen. Daarnaast ontstaat er volgens Röhle bij consumenten een steeds grotere weerstand tegen dit soort advertenties, hetgeen zich uit in het structureel vermijden en tegengaan van advertenties, bijvoorbeeld door te zappen of door het gebruik van 'pop-up blockers' (p. 2).

Online advertenties van zoekmachines kunnen deze problemen te boven komen doordat ze zich steeds beter kunnen richten op de behoeften van hun gebruikers. Voorheen berustte de plaatsing van dit soort advertenties voornamelijk op de termen in de zoekopdracht van de gebruikers. Deze bestonden bovendien veelal uit maar een paar termen en hadden vaak een onduidelijke betekenis (Spink & Jansen in Röhle 2007, p. 5). Door het verzamelen van persoonlijke informatie van gebruikers kunnen advertenties steeds beter op hun doelgroep gefocust worden en is er een grotere affiniteit met de getoonde advertenties. Bovendien kunnen de advertenties worden getoond naast de standaard zoekresultaten, waardoor gebruikers minder snel ontwijkend gedrag vertonen. De effectiviteit van de advertenties wordt hiermee sterk vergroot waardoor de 'return on investment' van de zoekmachines verhoogd wordt.

Er is dus zowel op het individuele niveau van de gebruiker als op commercieel niveau belang bij het opstellen van een zo nauwkeurig mogelijk profiel van de gebruiker. Personalisering biedt eveneens een oplossing voor het veel bekritiseerde search engine bias probleem doordat de relevantie van webpagina's niet meer enkel wordt berekend aan de hand van populariteit. Op individueel niveau zorgt het ervoor dat de gebruiker invloed heeft op het filtermechanisme van Google, waardoor de zoekresultaten beter op de gebruiker kunnen worden afgestemd. Daarnaast is personalisering voor Google op commercieel niveau van groot belang omdat advertenties beter op de gebruiker gericht kunnen worden. Het zo nauwkeurig mogelijk vaststellen van een gebruikersprofiel is dan ook een taak waar Google veel aandacht aan besteedt. In het volgende hoofdstuk zal gekeken worden naar de methoden die Google hanteert om haar gebruikers zo goed mogelijk te kunnen analyseren.

### 3. Personaliseren in de praktijk

Er zijn legio systemen waarbij informatie van de gebruiker wordt opgeslagen. Het opslaan van informatie dient in ieder systeem weer een ander specifiek doel. Dit geldt eveneens voor het type gebruikersinformatie dat wordt opgeslagen en voor de wijze waarop dat gebeurt (Keenoy & Levene 2005). Ook zoekmachines verschillen in de wijze waarop ze een zogenaamd 'user profile' opstellen. Bovendien is dit proces continu aan verandering onderhevig. Binnen dit hoofdstuk zal echter de focus liggen op de methoden die Google hanteert om een zo nauwkeurig mogelijk persoonlijk profiel van haar gebruikers op te stellen. Doordat Google steeds meer nieuwe (gratis) diensten op de markt brengt, is Google in staat steeds meer en steeds diversere gebruikersinformatie te verzamelen. Hierbij kan gedacht worden aan gebruiksinformatie zoals browse-gegevens, maar ook aan sociale interactie zoals verzonden en ontvangen e-mails tot aan de fysieke locatiegegevens van de gebruiker. Allereerst zal dit fenomeen, ook wel "googlization" genoemd uitvoerig worden besproken om vervolgens te kijken hoe googlization gebruikt wordt voor het opstellen van een gebruikersprofiel. Felix Stalder en Christine Mayer (2009) zetten in hun artikel *The Second Index: Search Engines, Personalization and Surveillance* drie verschillende lagen binnen het personaliseringsproces - waarin Google een filter op basis van een persoonlijk profiel, oftewel de second index, opstelt - uiteen. Aan de hand van deze lagen zullen met behulp van het werk van Kevin Keenoy en Mark Levene (2005) de verschillende personaliseringsmethoden die Google hanteert uiteengezet worden.

#### 3.1. Googlization

De term googlization wordt voor het eerste gebruikt door Alex Salkever (2003) in zijn column op de website Bloomberg Businessweek getiteld *Google Here, There, and Everywhere*. Salkever beschrijft hierin de steeds verdere verspreiding en overname van online diensten door Google. Hij stelt dat bedrijven die qua inkomsten afhankelijk zijn van hun online diensten steeds meer gedomineerd worden door de zoekgigant.

The distinction between getting info from Google or going to eBay or FedEx for it are somewhat subtle, as is the creeping nature of the Google universe. The search engine's innovations change how people use the Web at the margins, making it easier to weave it into the fabric of their everyday life and more convenient. For companies that rely on the Internet as a key part of their business, Googlization [sic] means they have to deal with a new, powerful intermediary seeking to embed itself ever deeper into the habits of Web surfers. (Salkever 2003, p. 2).

Door de jaren heen is binnen het Google discours de term googlization gebruikt om kritiek te leveren op de dominante positie van Google op het Web. Onder googlization kan de wijze waarop Google invloed heeft op het gebruik van het Web verstaan worden. Google is voor steeds meer mensen de toegangspoort tot het Web geworden. Zij starten hun surfessie via Google en maken tijdens het browsen niet meer zozeer gebruik van de hypertextuele linkstructuur, maar vallen veeleer keer op keer terug op Google als vertrekpunt naar nieuwe informatie. Ook wordt met googlization gerefereerd aan de hiërarchische vorm waarop Google het Web gestructureerd heeft en naar de macht die Google heeft in het bepalen van deze hiërarchie (Van der Vlies 2008, 40-1). De Amerikaanse cultuurhistoricus en mediawetenschapper Siva Vaidhyanathan (2007)

ziet googlization echter als de steeds verdere overname van diensten en industrietakken door Google. Hij stelt: "Google.com has radically altered the rules of the game for at least six major industries: Advertising, software applications, geographic services, e-mail, publishing, and Web commerce itself" (Vaidhyanathan 2007, p. 1). Een goed voorbeeld dat Vaidhyanathan (2011) geeft in zijn boek *The Googlization of Everything (and Why We Should Worry)* is de dienst Google Books. Google heeft met Google Books de klassieke taak van bibliotheken voor een deel overgenomen door miljoenen boeken te scannen en deze inzichtelijk en doorzoekbaar te maken voor internetgebruikers (p. 150). Een achterliggende motivatie voor de overname van dit soort industrietakken is echter het verzamelen van persoonsgegevens van gebruikers van deze diensten.

De Amsterdamse mediawetenschapper en oprichter van het Institute of network cultures Geert Lovink (2008) gaat in zijn artikel *The society of the query and the Googlization of our lives* in op deze overname van traditionele industrietakken door Google, waarbij ook volgens hem de overgenomen diensten enkel in het teken staan van het verzamelen van gebruikersinformatie. Hij betoogt tamelijk cynisch over het voorbeeld van Google Books:

The prime objective of this cynical enterprise is to monitor user behaviour in order to sell traffic data and profiles to interested third parties. Google is not after the ownership of Emile Zola; its intention is to lure the Proust lover away from the archive. Whereas for the French, Balzac's collected works are the epiphany of French language and culture, for Google they are abstract data junk, a raw resource whose sole purpose it is to make profit (Lovink 2008, p. 5).

Richard Rogers (2009), ook een mediawetenschapper aan de Universiteit van Amsterdam, vat dit principe samen in wat hij het "service-for-profile" model noemt. Googlization is volgens hem niet enkel het toepassen van dit model door Google binnen haar eigen diensten, maar ook het imiteren van dit model door andere bedrijven en Google adepten. Dit komt voornamelijk door het grote succes van Google onder haar gebruikers. In zijn artikel *The Googlization Question, and the Inculpable Engine* maakt Rogers onderscheid tussen twee vormen van googlization: *Back-end googlization* en *Front-end googlization*. Met de term back-end googlization duidt Rogers op de politieke invloed die het algoritme van Google heeft op het rangschikken van zoekresultaten en dus op de specifieke beeldvorming die daardoor tot stand komt. Google heeft hierdoor volgens Rogers de taak van traditionele opiniemakers zoals uitgevers en journalisten overgenomen en geautomatiseerd. Ook het basisprincipe van dit automatiseringsproces wordt overgenomen door andere bedrijven. Zo stopte bijvoorbeeld Yahoo in 2001 met Yahoo Directory - een overzicht van gecategoriseerde websites die handmatig werden toegevoegd - en lanceerde het in 2004 een eigen zoekmachine met een algoritme dat vergelijkbaar is met dat van Google (Rogers 2009, p. 176-8). Front-end googlization houdt volgens Rogers het kopiëren of imiteren van de interface(-elementen) van Google door derden. Ook deze overname heeft volgens hem info-politieke consequenties aangezien kleine aanpassingen in de interface directe gevolgen kunnen hebben op de informatievoorziening. Zo stelt Rogers bijvoorbeeld dat "[t]he burying of the directory [door de knop van de homepage te verwijderen] in both Yahoo and Google signals a much larger transformation -- the demise of the expert human editors of the Web" (ibid. p. 8).

De term googlization wordt door theoretici op verschillende manieren gebruikt, maar in wezen houdt het dus de verspreiding van (eigenschappen van) Google op het Web in. Deze verspreiding staat in dienst van de commerciële werkwijze - oftewel het service-for-profile model - van Google, dat daardoor op het Web een steeds belangrijkere plaats inneemt. De focus van Google ligt hierdoor niet meer enkel op haar zoekmachine maar richt zich op zoveel mogelijk informatieprocessen zowel on- als offline. De behoefte van Google om de gebruiker steeds beter te kunnen analyseren uit zich dan ook onder andere in de komst van Google+. Doordat dit een platform is waarop gebruikers hun eigen identiteit steeds beter zelf kunnen vastleggen is dit voor Google een waardevolle bron van informatie. In hoofdstuk 5 zal verder op Google+ en de identiteitsconstructie die hierop plaatsvindt worden ingegaan.

### 3.2. De ontwikkeling van een persoonlijk profiel

De methoden die Google met zijn veelvoud aan diensten hanteert om een zo nauwkeurig mogelijk gebruikersprofiel op te stellen zijn legio. Zo wordt op de achtergrond iedere klik of server-aanroep die gebruikers maken binnen een van Google's diensten opgeslagen in logbestanden. Daarin wordt standaardinformatie zoals IP adres, locatie, datum, tijd, etcetera verzameld. Daarnaast maakt Google bijvoorbeeld ook gebruik van zogenaamde *cookies*<sup>7</sup> die op de computer van de gebruiker worden geïnstalleerd. Aan de hand van deze cookies is Google in staat om de gebruiker te volgen tijdens diens surfessie over het Web (Stalder & Mayer 2009, p. 100).

Het verzamelen van gebruikersgegevens en -voorkeuren kan onderverdeeld worden in twee vormen: het expliciet en het impliciet verzamelen van gegevens. Tijdens het expliciet verzamelen van gegevens wordt van de gebruiker verwacht een of meerdere handelingen te verrichten, zoals bijvoorbeeld het invullen van een formulier of het aanklikken van like of dislike knop. Het impliciet verzamelen van gebruikersvoorkeuren vindt plaats tijdens reguliere interacties met het systeem, dus door bijvoorbeeld op bepaalde hyperlinks te klikken of door het opslaan van een pagina in de favorieten. Beide opties hebben voor- en nadelen. Zo is expliciet verzamelde informatie zeer waarheidsgetrouw, maar wordt de handeling die er voor nodig is door de gebruiker vaak als vervelend ervaren. Impliciet verzamelde informatie is minder betrouwbaar, maar vergt geen verdere acties van de gebruiker (Keenoy & Levene 2005, p. 5). Beide vormen kunnen eveneens naast elkaar gebruikt worden.

Stalder en Mayer (2009) onderscheiden drie verschillende lagen binnen het gebruikersprofiel dat wordt opgesteld door Google, namelijk dat van de gebruiker als "knowledge being", waarbij de persoonlijke voorkeuren van de gebruiker zo goed mogelijk worden achterhaald; als "social being", waarbij de identiteit van de gebruiker in relatie wordt gezien met zijn sociale contacten; en als "physical being", waarbij het profiel van gebruiker wordt gekoppeld aan zijn fysieke locatie. Aan de hand van deze drie lagen zal gekeken worden naar de personaliseringsmethoden die hierbinnen worden toegepast door Google. De informatiewetenschappers Kevin Keenoy en Mark Levene (2005) beschrijven aan de hand van een aantal systemen die zij analyseren in hun artikel *Personalisation of Web Search* zes verschillende methoden die gebruikt worden voor het personaliseren van zoekresultaten. Zij gaan in op "relevance feedback", "content

---

<sup>7</sup> Cookie: Een databestandje dat op de computer van de gebruiker wordt opgeslagen bij het bezoeken van een website. Wanneer de bezoeker de website een volgende keer bezoekt, dan wordt deze informatie door de browser naar de server gestuurd. Hierdoor kan de gebruiker door de website worden herkend en kan informatie over vorige bezoeken worden gebruikt.

analysis”, “recommender systems”, “link analysis”, “social systems” en “mobile and context awareness”. Hieronder zal gekeken worden in hoeverre deze methodes reeds door Google wordt toegepast en hoe dit soort systemen bijdraagt aan het samenstellen van een gebruikersprofiel.

### 3.2.1. Het kennisprofiel van de gebruiker

De diensten van Google zijn continu bezig met het verzamelen van data. Bovendien komen er constant nieuwe diensten bij om het type data waaruit het persoonlijk profiel bestaat steeds verder uit te breiden. Door diensten als AdSense, AdWords en Analytics is het Google gelukt om zichzelf te integreren binnen een enorm aantal websites, waardoor het niet enkel binnen haar eigen websites, maar ook op die van derden data kan verzamelen. Ook de interactie van de gebruiker met het systeem van Google zorgt ervoor dat het systeem continu met een leerproces bezig is. Alle opgeslagen data dragen bij aan het vormen van een “knowledge profile” (Stadler & Mayer 2009). Binnen dit knowlegde profile wordt specifiek informatie opgeslagen die betrekking heeft op de voorkeuren en het surfgedrag van de gebruiker zelf.

Één manier om te achterhalen welke webpagina’s de gebruiker wel en niet interessant of relevant vindt is het gebruik van relevance feedback. “Relevance feedback is the process of collecting user assessments of retrieved documents and using these assessments to try to improve search results” (Keenoy & Levene 2005, p.7). Deze informatie wordt vervolgens gebruikt om bij de volgende zoekopdrachten bepaalde resultaten een hogere ranking te geven of deze er juist uit te lichten. Er bestaan een hoop verschillende wijzen waarop het verzamelen van relevance feedback in systemen wordt geïmplementeerd. Zo kunnen gebruikers bijvoorbeeld expliciet bepaalde documenten aanvinken die zij waardevol of juist waardeloos achten. Maar ook impliciet kan er informatie worden verzameld. Wanneer gebruikers bijvoorbeeld met hun muis over bepaalde links gaan, waardoor er een korte samenvatting verschijnt en zij vervolgens toch niet op de link klikken, kan worden aangenomen dat die bepaalde pagina minder relevant is (ibid. p. 9).

Ook Google maakt gebruik van relevance feedback. Voorheen gebeurde dit bijvoorbeeld in de vorm van SearchWiki (google.com 2008) en later in de vorm van Google Stars. Gebruikers konden op een sterretje achter de zoekresultaten klikken. Wanneer er de volgende keer een soortgelijke zoekopdracht werd uitgevoerd, dan verschenen de aangevinkte items bovenaan de zoekresultaten, zoals te zien is in afbeelding 1 (google.com 2010). Inmiddels zijn de Google Stars vervangen door de Google +1 button<sup>8</sup>. Daarnaast is het mogelijk om bepaalde pagina’s te verwijderen uit de zoekresultaten door deze domeinen geheel te blokkeren via een link bij ieder zoekresultaat (google.com 2011b).

Een andere methode voor het personaliseren van zoekresultaten is het personaliseren op basis van content analysis. Hiervoor wordt de informatie die is opgeslagen in het persoonlijk profiel van de gebruiker gebruikt om deze te matchen met de zoekresultaten en aan de hand daarvan de relevantie van pagina’s te bepalen. Hierbij wordt het persoonlijk profiel opgebouwd uit bijvoorbeeld de zoekhistorie van gebruikers. Alle zoekopdrachten die worden gegeven worden opgeslagen in een database en worden onderverdeeld in categorieën, waardoor er een persoonlijke lijst met interesses gevormd wordt (Keenoy & Levene 2005, p. 11). Maar ook allerlei andere soorten informatie die bekend zijn over de gebruiker kunnen worden gebruikt om de meest relevante pagina’s te vinden op basis van hun inhoud.

---

<sup>8</sup> In hoofdstuk 5 zal op de werking van de Google +1 knop verder worden ingegaan



[Advanced Search](#)

Web [Show options...](#)

Results 1 - 10 of about 92,600,000 for nfl [definition]. (0.17 seconds)

Starred results for nfl

★ [The Official Website of the Indianapolis Colts](http://www.colts.com/) - [www.colts.com/](http://www.colts.com/)

★ [New Orleans Saints - Super Bowl Champions](http://www.neworleanssaints.com/) - [www.neworleanssaints.com/](http://www.neworleanssaints.com/)

[NFL.com](http://www.nfl.com/) - Official Site of the **National Football League** ☆

Official site of the **National Football League**. It delivers in-depth team pages for all clubs, game-day coverage with real time statistics and play-by-play ...

[Schedules](#) - [Fantasy](#) - [Scores](#) - [Teams](#)

[www.nfl.com/](http://www.nfl.com/) - [Cached](#) - [Similar](#)

Afbeelding 1: Weergave van zoekresultaten, waarbij de bovenste twee zijn aangeraden en dus aangeduid met ster.

Google verzamelt hiervoor dan ook tal van verschillende browsegegevens, door middel van haar eigen online diensten, zoals Google Docs, Gmail, Google Translate, Blogger, etcetera, maar ook door middel van offline diensten zoals bijvoorbeeld haar browser Google Chrome. Hierdoor worden niet meer enkel online- maar ook offlinegegevens verzameld door Google. Google slaat bijvoorbeeld iedere toetsaanslag die een gebruiker maakt in Chrome's Omnibox - de geavanceerde adresbalk van Chrome - op (Fried 2008). Daarnaast wordt de historie van Chrome opgeslagen, waardoor Google iedere webpagina kan bijhouden die bezocht is. Dit geldt zelfs voor webpagina's die met een SSL certificaat<sup>9</sup> beveiligd zijn (Cheung 2008). Al deze informatie wordt vergeleken met de inhoud van webpagina's om zo de relevantie en ranking te bepalen. Dit kan bijvoorbeeld aan de hand van meta-tags, maar tegenwoordig maakt Google veeleer gebruik van de gehele semantische structuur van de HTML pagina, waarbij kopteksten en bepaalde datatypen - zoals bijvoorbeeld persoonsgegevens of filminformatie gedefiniëerd in Microformats<sup>10</sup> - worden geanalyseerd.

Deze informatie wordt eveneens gebruikt als basis voor de de recommender systems van Google. Recommender systems worden meestal gebruikt om bijvoorbeeld boeken of films aan te raden op basis van eerder bekeken items. Het doel is vaak om consumenten nieuwe producten te laten leren kennen uit een enorm aantal keuzes (Resnick & Varian 1997, p. 56). Maar recommender systems worden ook gebruikt om bijvoorbeeld relevante nieuwsitems aan te raden. Een interessant verschil met dit soort systemen en de eerder genoemden is dat ze werken op basis van de het zogenaamde 'collaborative filtering' principe. Hierbij wordt niet alleen gebruik gemaakt van het profiel van de gebruiker zelf, maar worden soortgelijke profielen van andere gebruikers gebruikt om informatie aan te raden (Ujjin & Bentley 2001, p. 1). Google vergelijkt dus gebruikersprofielen met elkaar op basis van overeenkomsten en koppelt hier een bepaald verwachtingspatroon aan. Wanneer bijvoorbeeld gebruiker A en B beiden het boek X hebben aangeschaft en

<sup>9</sup> SSL certificaat: SSL oftewel Secure Sockets Layer is een encryptiemethode waarmee gegevens veilig over het internet verzonden kunnen worden. Een SSL certificaat zorgt ervoor dat de data versleuteld worden en dat identiteit van de te bezoeken website wordt gevalideerd.

<sup>10</sup> Microformats: Een semantische uitbreiding voor HTML tags. Door middel van de Microformats standaard kan informatie aan de structuur van de website worden toegevoegd, waardoor zoekmachines weten om wat voor een soort informatie het gaat. Dit kan bijvoorbeeld gebruikt worden om aan te geven dat een bepaalde HTML tag een auteursnaam of een datum van een blogpost bevat. (Zie: <http://microformats.org>)

gebruiker A heeft ook boek Y aangeschaft dan is de kans dat gebruiker B boek Y ook interessant zal vinden groter. Dit is uiteraard een versimpelde voorstelling van zaken, maar het is wel de basis waarop collaborative filtering is gebaseerd. Hieruit blijkt dus dat het knowledge profiel dat Google opstelt van zijn gebruikers onder andere wordt beïnvloed door het profiel van soortgelijke gebruikers.

Recommender systems worden in tal van Google diensten gebruikt, sommige wat opzichtiger dan anderen. Een bekend voorbeeld is Google's "Did you mean" functie. Wanneer een gebruiker een bepaalde zoekopdracht uitvoert, herkent Google soortgelijke zoekopdrachten die veel vaker worden uitgevoerd. Op basis hiervan verwacht Google dat de gebruiker eigenlijk deze zoekopdracht bedoelt heeft en raadt deze aan. Hierdoor is Google bijvoorbeeld in staat om spelfouten in de zoekopdrachten te herkennen. Een andere dienst die gebruikt maakt van een recommender system is Google News, waarbij de gebruiker nieuwsberichten aangeraden krijgt op basis van het leesgedrag van soortgelijke gebruikers (Price 2006). Maar misschien wel het belangrijkste maar minder opvallend recommender system is Google's Pagerank algoritme dat aan de hand van link analysis bepaald welke pagina's relevant zijn voor de gebruiker. Met de komst van Google+ heeft Google er nog een sociaal recommender system bij in de vorm van de Google +1 knop. Hiermee kunnen gebruikers expliciet hun voorkeur delen met personen binnen hun Google Circles.

### [3.2.2. Het sociale profiel van de gebruiker](#)

Met de komst van het Web 2.0 werd het steeds duidelijker dat gebruikers niet alleen informatie consumeren, maar dat zij ook informatie willen delen met anderen. Hierdoor werden gebruikers steeds meer gezien als sociale individuen die zowel on- als offline intensieve relaties aangaan met anderen (Stalder & Mayer 2009, p. 103). Google is daarom meer rekening gaan houden met de sociale connecties van haar gebruikers bij het definiëren van een gebruikersprofiel. Door de overname en lancering van grote social network sites zoals YouTube en Orkut is Google steeds beter in staat om de connecties tussen gebruikers in kaart te brengen. Maar ook diensten als Google Groups en Gmail worden gebruikt om de sociale interacties tussen gebruikers te monitoren. Hele conversaties kunnen geanalyseerd worden om na te gaan waar de interesses van de gebruikers liggen en welke overeenkomsten zij met elkaar vertonen (ibid p. 104). Met deze informatie is Google in staat om aannames over een bepaalde gebruiker te doen op basis van zijn sociale connecties. Dit principe is vergelijkbaar met dat van recommender systems, maar in dit geval worden de aannames gebaseerd op verbindingen die de gebruiker zelf expliciet tot stand heeft gebracht in plaats van dat er sprake is van soortgelijke profielen.

### [3.2.3. Het fysieke profiel van de gebruiker](#)

De derde laag in de door Stalder en Mayer beschreven second index is gebaseerd op de fysieke context van de gebruiker. Door fysieke gegevens zoals de geografische locatie van een gebruiker toe te voegen aan het gebruikersprofiel is Google in staat om de gebruiker van informatie te voorzien die betrekking heeft op de geografische positie van de gebruiker. Daarnaast stelt Google zichzelf in staat om gerichtere advertenties te plaatsen. Zo is een gebruiker bijvoorbeeld meer geïnteresseerd in restaurants in diens eigen woonplaats dan in restaurants die zich elders bevinden. De komst van smartphones heeft de mogelijkheden voor Google om locatiegerichte informatie te verzamelen sterk vergroot. Door diensten zoals Google Latitude is Google in staat om de fysieke verplaatsing van gebruikers in kaart te brengen en te achterhalen welke locaties zoals

restaurants en andere openbare ruimtes zij bezoeken (google.com 2011c). Ook de zoekfunctie in Google Maps kan worden ingezet om te bepalen waar gebruikers zich naartoe willen verplaatsen en wat hun favoriete locaties zijn (ibid.).

Niet enkel locatiegegevens worden door Google verzameld, maar ook andere zaken die betrekking hebben op de gebruiker als fysiek persoon. Zo had Google met de - in januari 2012 te beëindigen - dienst Google Health inzicht in de medische gegevens van haar gebruikers. Daarnaast kunnen gebruikers betalingen verrichten met de dienst Google Wallet, waardoor de aankopen en de aankooplocaties van gebruikers kunnen worden geregistreerd (google.com 2011d).

Uit bovenstaande blijkt dat de mogelijkheden van Google voor het verzamelen van persoonlijke informatie van haar gebruikers zich zelfs tot buiten het online domein uitstrekken. De hoeveelheid aan persoonlijke informatie die Google verzamelt is dan ook enorm<sup>11</sup>. Alhoewel er voor zowel Google als haar gebruikers vele voordelen kleven aan een dergelijk grote informatie- database, moet er ook kritisch gekeken worden naar de problemen die hiermee gepaard gaan. Het op grote schaal verzamelen van persoonlijke informatie roept onmiddellijk vraagtekens op wanneer er gekeken wordt naar de privacy van Google's gebruikers. Daarnaast dreigt de autonomie van de gebruikers in het gedrang te komen omdat Google steeds meer beslissingen voor hen gaat maken op basis van hun eerder verzamelde voorkeuren. In het volgende hoofdstuk zullen deze twee thema's en hun consequenties voor de gebruiker uiteengezet worden.

---

<sup>11</sup> SEOMoz biedt een overzicht van de verschillende gebruikersdata die verzameld wordt per Google dienst: <http://www.seomoz.org/blog/the-evil-side-of-google-exploring-googles-user-data-collection>

## 4. De gevaren van personaliseren

Zoals gebleken is in hoofdstuk 2 is het filteren van zoekresultaten op basis van een persoonlijk profiel van groot belang voor het vinden van relevante informatie. Marginale standpunten die op het internet weinig vertegenwoordigd worden, komen hierdoor eerder naar voren en gebruikers hebben invloed op de wijze waarop het filtermechanisme opereert. Alhoewel de voordelen van personalisering voor de gebruiker veelal positief lijken, kleven er ook een aantal grote problemen aan deze steeds verdergaande ontwikkeling. Binnen het kritisch discours rondom het personaliseren van zoekresultaten is een aantal belangrijke kritiekpunten te onderscheiden op het gebied van privacy, autonomie en transparantie (Fuez et al. 2011; Rieder 2009; Röhle 2007; Stalder & Mayer 2009; Zimmer 2008b). In dit hoofdstuk zal aan de hand van deze waarden kritisch worden gekeken naar de problemen van personalisering. Aangezien de transparantie van een systeem nauw verbonden is met de autonomie van de gebruiker zal, mede door de beperkte scope van dit onderzoek, transparantie worden behandeld binnen het thema autonomie.

### 4.1. Beperking van privacy

De komst van nieuwe media heeft een grote verschuiving veroorzaakt tussen de informatie die wij als privé en publiek ervaren (Halavais 2009, p. 139). Op social network sites, blogs en andere Web 2.0 toepassingen maken individuen allerhande (voorheen) persoonlijke informatie publiekelijk. Daarnaast is het door de digitale media mogelijk geworden om grote hoeveelheden aan persoonlijke informatie te verzamelen, op te slaan en te distribueren (Lyon 2007, p. 85). De verzameling van persoons- en gebruiksgegevens door zoekmachines voor commercieel gebruik ontkomt niet aan de nodige kritiek op de manier waarop er met de privacy van gebruikers wordt omgegaan. Zo zou Google Europese privacywetgeving schenden omdat het gebruiksinformatie en cookies te lang zou bewaren (Palmer 2007). Een ander voorbeeld is het feit dat Google Streetview in verschillende landen tot verantwoording is geroepen en in Tsjechië zelfs is verboden omdat het inbreuk zou maken op privacy van de burgers (msn.com). Daarnaast zou het illegaal informatie van beschikbare WiFi netwerken hebben opgeslagen (Shiels 2010).

Het verzamelen van persoonlijke informatie kan op zichzelf al gezien worden als een gevaar voor de privacy van Google's gebruikers. Er ontstaat echter nog een groter maatschappelijk probleem doordat mensen aan de hand van hun persoonsgegevens en voorkeuren door Google worden gecategoriseerd. Dit kan grote gevolgen hebben voor de informatievoorziening van een individu. Om deze kritiek rondom het categoriseren van gebruikers en de consequenties hiervan te kunnen duiden zal vanuit het veld van surveillance studies gekeken worden naar de bezwaren die hieraan kleven.

Volgens de socioloog en directeur van het Surveillance Studies Centre David Lyon (2001) valt onder surveillance te verstaan: "the collection and processing of personal data, whether identifiable or not, for the purposes of influencing or managing those whose data have been garnered" (Lyon 2001, p 2). De wijze waarop surveillance plaatsvindt en wordt uitgevoerd is door de jaren heen behoorlijk aan verandering onderhevig geweest. Daar waar voorheen nog sprake was van face-to-face surveillance in de vorm van bijvoorbeeld een fabrieksopzichter die zijn werknemers in de gaten hield, worden inmiddels grote groepen mensen met behulp van geavanceerde digitale systemen gesurveilleerd en geclassificeerd. Dit gebeurt aan de hand van persoonskenmerken zoals opleidingsniveau, kleur van de ogen of bijvoorbeeld in het geval van zoekmachines hun zoekhistorie (Lyon 2007, p. 73-93).

Maar niet alleen de wijze waarop surveillance wordt uitgevoerd is aan verandering onderhevig, ook het doel van surveillance is met de komst van digitale media verschoven. Traditionele vormen van surveillance hadden als doel een zelfdisciplinerende werking te hebben op de gesurveilleerden. Dit principe van machtsuitoefening door middel van surveillance is goed te illustreren aan de hand van Jeremy Bentham's gevangenismodel, het Panopticon. Binnen dit Panopticon werden gevangenen in aparte cellen geplaatst. In het midden van de gevangenis stond een wachttoren van waaruit de gevangenen geobserveerd konden worden. Het centrale idee van Bentham's ontwerp was dat gevangenen niet wisten op welk moment zij geobserveerd werden omdat er een fel licht vanaf de toren naar hen scheen. "Hence the major effect of the Panopticon: to induce in the inmate a state of conscious and permanent visibility that assures the automatic functioning of power" (Foucault 1979, p. 201). De Franse filosoof Michel Foucault zag het Panopticon als metafoor voor allerlei soorten instituties binnen onze maatschappij - zoals ziekenhuizen, scholen en fabrieken - waarbinnen de autodisciplinaire werking van surveillance zichtbaar is (ibid. p. 205).

Tegenwoordig wordt surveillance veeleer gebruikt voor commerciële doeleinden om bepaald gedrag of de voorkeuren van individuen te voorspellen (Haggerty & Ericson 2000; Lyon 2007, p. 89; Röhle 2007). In het geval van zoekmachines worden persoonsgegevens opgeslagen en aan de hand van deze gegevens worden gepersonaliseerde advertenties getoond. Een groot verschil tussen de twee genoemde vormen van surveillance is de wijze waarop individuen worden gekarakteriseerd. De Franse filosoof Gilles Deleuze (1992) beschrijft in zijn invloedrijke manifest *Postscript on the Societies of Control* de overgang van een disciplinaire maatschappij naar een "society of control", waarin surveillance alomtegenwoordig is. Hij stelt dat in onze huidige maatschappij "individuals have become dividuals" (Deleuze 1992, p. 3). Hiermee bedoelt Deleuze dat individuen worden opgedeeld en digitaal opgeslagen in stukjes data die hun persoonskenmerken weergeven. Een opeenstapeling van deze stukjes data wordt gezien als de digitale representatie oftewel het persoonlijk profiel van het individu. Haggerty en Ericson (2000) beschrijven dit fenomeen als de "data-double".

Het voornaamste probleem van de data-double is dat het geen reële weergave van het individu is maar door computersystemen wel als zodanig wordt gebruikt. Zoals socioloog Bart Simon (2005) stelt: "the reference of the panoptic gaze is no longer the body but its double, and indeed this is no longer a matter of looking but rather one of data analysis" (Simon 2005, p. 15). De data-double heeft hierdoor dus ook directe invloed op het individu zelf. Dit kan - in het geval van zoekmachines - zijn op het gebied van informatievoorziening, maar kan ook zorgen voor discriminatie en sociale uitsluiting (Lyon 2007, p. 94-117). In het geval van zoekmachines staat de data-double aan de basis van de gepersonaliseerde zoekresultaten en advertenties, waardoor de gebruiker constant met zijn vermeende identiteit te maken heeft.

Door de fragmenten van verschillende data-doubles met elkaar te vergelijken is het mogelijk om gebruikers te categoriseren aan de hand van hun eigenschappen. Binnen surveillance studies wordt er dan gesproken van 'social sorting'. David Lyon omschrijft en bekritiseert dit fenomeen op de volgende wijze:

Codes, usually processed by computers, sort out transactions, interactions, visits, calls and other activities. They are invisible doors that permit access to, or exclude from participation in a myriad of events, experiences and processes. The resulting classifications are designed to influence and manage populations and persons thus

directly and indirectly affecting the choices and chances of data subjects (Lyon in Stalder & Mayer 2009, p. 107).

Social sorting wordt door zoekmachines toegepast om op basis van de sociale categorieën het gedrag en de voorkeuren van individuen te voorspellen (Röhle 2007). Ook de beroemde socioloog Richard Jenkins (2000) stelt dat alhoewel het categoriseren van individuen noodzakelijk is om de wereld om ons heen te kunnen begrijpen, deze (vaak bij het individu onbekende) categorisering van grote invloed kan zijn op het leven van het individu. Dus de wijze waarop een individu wordt geïdentificeerd is volgens Jenkins minstens zo belangrijk als de identiteit van de persoon zelf. Er schuilt hierdoor een grote macht in het categoriseren van individuen. Zoekmachines als Google oefenen deze macht uit door de gebruiker zoekresultaten te tonen op basis van diens persoonlijk profiel. Hierdoor wordt de toegang tot informatie beperkt aan de hand van de data-double van de gebruiker, waar deze geen directe invloed op kan uitoefenen. Daardoor wordt eveneens de autonomie van de gebruiker beperkt.

#### 4.2. Beperking van autonomie

Om als individu tot een persoonlijk en individueel standpunt te kunnen komen is het van belang dat een individu autonoom op zoek kan gaan naar informatie zonder hierbij beperkt te worden door het systeem. Philip Brey (2000), professor in de filosofie van technologie aan de Universiteit Twente, beschrijft autonomie als:

the ability to construct one's own goals and values, and to have the freedom to make choices and plans and act in ways that are believed by one to help achieve these goals and promote these values. [...] [I]t is a freedom one must have in order to create a life that is experienced by oneself as meaningful and fulfilling (Philip Brey 2000, p. 14).

Doordat zoekmachines toegang verlenen tot een enorme hoeveelheid van informatie, zorgen zij enerzijds voor een toename van autonomie bij de gebruikers. De gebruikers zijn door deze informatie immers beter in staat hun eigen keuzes te bepalen en waarden te vormen. Het filtermechanisme van zoekmachines is hierbij van essentieel belang, aangezien dit voor de gebruiker een selectie maakt van de enorme hoeveelheden informatie. Echter doordat het filtermechanisme van Google niet inzichtelijk oftewel niet transparant is, wordt anderzijds de autonomie van de gebruiker verkleint aangezien deze niet weet op basis van welke factoren het filteralgoritme opereert. Hierdoor is de gebruiker minder goed in staat om weloverwogen keuzes te maken tussen informatie en wordt de keuzevrijheid beperkt. De gebruiker heeft dankzij de personalisering van het filtermechanisme weliswaar invloed op de uitkomsten, maar de manier waarop invloed wordt uitgeoefend is onbekend. De jurist Frank Pasquale beschrijft dit autonomieprobleem op de volgende manier:

At a minimum, autonomy requires a meaningful variety of choices, information of the relevant state of the world and of these alternatives, the capacity to evaluate this information and the ability to make a choice. If A controls the window through which B sees the world - if he systematically exercises power over the relevant information about

the world and available alternatives and options that reaches B - then the autonomy of B is diminished. To control one's informational flows in ways that shape and constrain her choice is to limit her autonomy, whether that person is deceived or not (Pasquale in Stalder & Mayer 2009, p. 109).

Daarbij komt dat op basis van het persoonlijk profiel van de gebruiker informatie als relevant wordt aangemerkt op een wijze die voor de gebruiker zelf niet inzichtelijk is (Lobet-Maris 2009, p. 81). Het algoritme bepaalt dus als het ware wat wel en niet interessant is voor de gebruiker, waardoor er een eenzijdig beeld van de werkelijkheid kan ontstaan. Internet activist Eli Pariser (2011) beschrijft dit fenomeen in zijn populair-wetenschappelijke boek *The Filter Bubble: What the Internet is Hiding from You* als de "filter bubble". Hij stelt dat zoekmachines constant bezig zijn met het verfijnen van het persoonsprofiel van hun gebruikers en op basis daarvan hun toekomstig handelen proberen te voorspellen. Hierdoor ontwikkelen zoekmachines volgens Pariser per individu een unieke informatiebubble die grote invloed heeft op de wijze waarop mensen hun ideeën en standpunten ontwikkelen (Pariser 2011, p. 9). Pariser stelt dus impliciet dat zoekmachines constant informatie die niet past binnen het persoonlijk profiel van de gebruiker wegfilteren. Hierdoor ontstaat een beperkt wereldbeeld en een beperkte keuzevrijheid van informatie. De autonomie van de gebruiker wordt hierdoor dus aangetast. Alhoewel er getwijfeld wordt aan de mate waarin deze filter bubble zich manifesteert en de invloed die het heeft op de beeldvorming (Weisberg, 2011), stellen Stalder en Mayer dat "[t]he differences between the personalized and the non-personalized version of the search results [...] are initially subtle, but likely to increase over time" (Stalder & Mayer 2009, p. 110). Daarnaast stelt Google dat het mogelijk is om de personalisering van zoekresultaten te beperken door uit te loggen en de door Google geplaatste cookies te verwijderen (google.com n.d. b). Veel gebruikers zijn zich echter niet bewust van deze mogelijkheid. Bovendien vereisen verschillende diensten van Google dat de gebruiker is aangemeld om van de dienst gebruikt te kunnen maken, waardoor het uitgelogd blijven ook geen reële optie is.

De impact van het beperkte wereldbeeld dat de gebruiker door zoekmachines wordt voorgespiegeld wordt naast het ontstaan van filter bubbles nog eens versterkt door het feit dat mensen geneigd zijn eerder informatie te consumeren die binnen hun eigen interesses valt die overeenkomt met hun eigen denkbeelden. Cass R. Sunstein (2007) stelt in zijn boek *The Republic.com 2.0* dat het filteren van informatie onvermijdelijk is om een 'information overload'<sup>12</sup> te voorkomen. Filteren op zich is volgens hem geen probleem, het probleem zit volgens Sunstein in het feit dat "when options are so plentiful, many people will take the opportunity to listen to those points of view that they find most agreeable. [...] [T]here is a natural human tendency to make choices with respect to entertainment and news that do not disturb our preexisting view of the world" (Sunstein 2007, p. 52). De denkbeelden van individuen worden hierdoor alsmaar versterkt en minder genuanceerd. Dit proces wordt binnen het media discours ook wel het echo chamber effect genoemd. Het probleem van dit soort echo chambers is volgens Sunstein dat wanneer verschillende groepen zich alleen maar focussen op specifieke standpunten en perspectieven, het onderlinge begrip verdwijnt waardoor het gezamenlijk oplossen van maatschappelijke problemen bemoeilijkt wordt (ibid. p. 56).

---

<sup>12</sup> Information overload: Een term die vaak gebruikt wordt in relatie met het internet vanwege de enorme hoeveelheid aan informatie die beschikbaar is. Er is sprake van een information overload wanneer iemand meer informatie tot zich krijgt dan hij kan verwerken.

Filter bubbles zijn niet zozeer de oorzaak van dit soort echo chambers, maar werken wel als versterkende factor doordat gelijkgestemde informatie nog eerder voorhanden is.

Om te voorkomen dat gebruikers als het ware door hun data-double in een informatie-isolement belanden is het van belang dat het algoritme niet enkel opereert op basis van de individuele interesses van de gebruiker, maar dat er ook rekening wordt gehouden met de sociale connecties van deze gebruiker. Deze connecties kunnen bruggen vormen tussen de verschillende filter bubbles en zo het isolement van de gebruiker doorbreken (Metahaven 2009). Google's nieuwe social network site Google+ biedt Google de mogelijkheid om de connecties tussen gebruikers beter in kaart te brengen. Bovendien biedt het gebruikers de mogelijkheid om zelf expliciet hun persoonlijk profiel uit te breiden en te bewerken. Dit heeft positieve gevolgen voor de problemen rondom privacy en autonomie. In het volgende hoofdstuk zal gekeken worden naar de rol die Google+ speelt bij het verkleinen van deze problemen. Daarnaast zal er ingegaan worden op de mogelijkheden die Google+ aan Google biedt om gebruikersinformatie beter te kunnen ontsluiten.



## 5. De volgende stap in personalisering: Google+

Op 28 juni 2011 lanceerde Google de betaversie van haar nieuwste dienst, de social network site Google+, op basis van Google's bekende 'invite-only' principe. Dit principe werd door Google al eerder toegepast bij onder andere de dienst Gmail. De eerste gebruikers mochten 150 uitnodigingen versturen naar anderen (Pinola 2011), waardoor de dienst in eerste instantie een besloten karakter had. Op 20 september 2011 werd de dienst publiekelijk gemaakt voor alle gebruikers boven de 18. Rond eind oktober 2011 had Google+ rond de 40 miljoen gebruikers, wat de site de snel groeiende social network site in de geschiedenis maakt (Evans, 2011a).

De komst van Google+ kan gezien worden als een logische stap in Google's steeds verder gaande personalisering en socialisering. Via Google's - vooral in Zuid-Amerika populaire - social network site Orkut en de overname van - en de nauwe samenwerking met verschillende social network sites zoals respectievelijk YouTube en Twitter was Google reeds in staat om grote hoeveelheden gebruikersinformatie te verzamelen (google.com n.d. a). Het ontwikkelen van een social network site die voornamelijk gericht is op de interactie en verhoudingen tussen de gebruikers zelf en die een belangrijke tegenhanger kan zijn van de immens populaire social network site Facebook biedt Google de mogelijkheid om de interactie tussen gebruikers nog nauwkeuriger in kaart te brengen en de zoekresultaten en advertenties te socialiseren (Siegler 2011a). Daarnaast heeft het positieve effecten op de privacy en autonomie van de gebruikers. In dit hoofdstuk zal gekeken naar de wijze waarop Google+ gebruikers in staat stelt hun identiteit en voorkeuren zoals deze bekend zijn bij Google enigszins te beïnvloeden, waardoor hun autonomie vergroot wordt. Daarnaast zullen de functies Google Circles en de Google's +1 knop uitvoerig worden besproken en zal gekeken worden naar de effecten die deze hebben op het gebied van privacy en autonomie.

### 5.1. Identiteitsconstructie

Google CEO Eric Schmidt stelde in een interview op het MediaGuardian Edinburgh International TV Festival dat het gebrek aan een goede 'identity service' op het internet een van de voornaamste redenen was voor de ontwikkeling van Google+ (Carvin 2011). Een dergelijke dienst staat haaks op de idealen zoals anonimiteit en een egalitair internet die de cyber libertarians voor ogen hadden. Schmidt stelt echter dat er vele voordelen kleven aan vaststellen van de 'ware' identiteit van de gebruikers. Zo stelt hij bijvoorbeeld dat:

if we knew that it was a real person, then we could sort of hold them accountable, we could check them, we could give them things, we could you know bill them, [...] there are people who do really really evil and wrong things on the Internet, and it would be useful if we had strong identity so we could weed them out, I'm not suggesting eliminating them, what I'm suggesting is if we knew their identity was accurate, we could rank them. Think of them like an identity rank (Schmidt in Carvin 2011, § 1).

Aan 'user generated content'<sup>13</sup> kan dan op basis van dit soort identity ranks, oftewel reputatiescores, een kwaliteitswaardering worden toegekend. Een bijdrage van iemand met een goede reputatie krijgt dan een hogere kwaliteitswaardering al die van iemand met een lagere reputatie. Dit geldt voor artikelen, maar ook voor bijvoorbeeld de reacties daarop (Slawski 2011). Daarnaast benoemt Schmidt het feit dat een accuraat persoonsprofiel Google helpt om haar diensten zoals de zoekmachine te verbeteren.

Een van de spraakmakende eisen die Google ten behoeve van haar identity service stelde aan het aanmaken van een Google+ account, is dat de gebruiker zich enkel onder zijn echte naam mocht registreren. Pseudoniemen werden door Google niet geaccepteerd en de gebruikers werden geblokkeerd (Evans 2011b, Goodwin 2011). Dit is Google echter op een hoop kritiek komen te staan, aangezien veel gebruikers een valide reden hebben om niet hun echte naam op het internet te gebruiken. Hierbij kan bijvoorbeeld gedacht worden aan gebruikers die onder een dictatuur leven, homoseksuelen en medische patiënten (Cowan 2011). Eerder stelde Schmidt dat het eenieder vrij staat om Google+ te gebruiken zolang ze zich houden aan de door Google gestelde gebruiksvoorwaarden. Hij zag daarom geen noodzaak om af te wijken van het beleid (Carvin 2011). Desalniettemin is Google twee maanden na de lancering teruggekomen op haar strikte beleid en staat het inmiddels weer pseudoniemen toe (Empson 2011).

Een ander middel om de identiteit van de gebruiker beter vast te stellen is de Google+ Sparks functie. Gebruikers kunnen in een zoekmachine - die door Google's senior vice president of Social Vic Gundotra wordt aangeduid als een "sharing engine" - een van hun interesses invoeren. De zoekmachine retourneert aan de hand van de ingevoerde interesse een lijst met artikelen, video's en andere multimediale toepassingen welke vervolgens kunnen worden bekeken, opgeslagen en gedeeld met contacten binnen je Google Circles (Siegler 2011a). Het interessante aan deze functie is dat gebruikers hierdoor in staat zijn expliciet hun interesses in het systeem in te voeren. Zij krijgen dus meer macht over de constructie van hun eigen persoonlijk profiel, waardoor hun autonomie toeneemt. Daarnaast neemt voor Google de noodzaak om impliciet interesses van gebruikers te extraheren uit de verzamelde data af. Dit kan ten goede komen aan hun diensten en advertentiemodel. Alhoewel Google op dit moment stelt dat "Google+ is not part of the Google Display Network" (Singel 2011), wat inhoudt dat gebruikersinformatie niet gebruikt wordt voor gerichte advertenties, betekent dit niet dat gebruikersinformatie in de toekomst hier niet voor gebruikt zal worden. Dit soort uitspraken zijn in het verleden bijvoorbeeld bij Youtube ook onwaar gebleken (ibid.).

## 5.2. [Google Circles](#)

Een van de belangrijkste functies binnen Google+ is Google Circles. Gebruikers kunnen hiermee hun contacten indelen in bepaalde groepen zoals vrienden, familie en collega's. Vervolgens kunnen statusupdates, artikelen en andere webcontent worden gedeeld binnen deze Circles. Google had hiermee als doel om de informatie op dezelfde manier te delen als in de fysieke wereld gebeurt. Zoals Google zelf stelt: "From close family to foodies, we found that people already use real-life circles to express themselves, and to share with precisely the right folks. So we did the only thing that made sense: we brought Circles to

---

<sup>13</sup> User generated content: Webcontent die door de gebruikers van een online medium wordt toegevoegd. Hierbij kan gedacht worden aan multimediale content zoals afbeeldingen en video, maar ook aan reacties en de content van websites zoals Wikipedia.

software” (google.com 2011e). Het delen van informatie gaat dus op basis van Circles, maar ook de informatie die binnenkomt via de berichten-stream van iemands contacten kan gefilterd worden aan de hand van de Circles.

Ook hier stelt Google haar gebruikers weer in staat om expliciet hun persoonlijk profiel aan te passen door hun contacten in te delen binnen bepaalde groepen. Hierdoor kan Google een nauwkeuriger beeld krijgen van de sociale relaties tussen gebruikers. Richard Jenkins (2000) merkt een belangrijk verschil op tussen groepen en categorieën. Hij stelt dat groepen als zodanig worden erkend door de leden ervan en dat categorieën worden gevormd zonder medeweten van de leden door middel van sociale analyse oftewel social sorting. Doordat Google+ gebruikers de mogelijkheid geeft om zichzelf binnen groepen in te delen wordt een deel van het categorisatieproces bij Google uit handen genomen, wat kan leiden tot vermindering van de problemen op het gebied van social sorting. Hier geldt evenwel dat het onduidelijk is in hoeverre de Google Circles gebruikt worden bij het optimaliseren van zoekresultaten en advertenties, maar aannemelijk is dat deze informatie invloed heeft op het persoonlijk profiel van de gebruiker.

De komst van Google Circles en de mogelijkheid om informatie met een selecte groep contacten te delen kan gezien worden als een goede stap in de richting van een betere waarborging van de privacy van gebruikers. Deze verbetering kan goed worden geduid met behulp van Helen Nissenbaums (2004) artikel *Privacy as Contextual Integrity*. Hierin beschrijft Nissenbaum een nieuw raamwerk om te kijken in hoeverre er sprake is van privacyschending bij het delen van bepaalde informatie. Nissenbaum levert kritiek op het traditionele privacy-model dat uitgaat van drie basisprincipes, namelijk: “(1) limiting surveillance of citizens and use of information about them by agents of government, (2) restricting access to sensitive, personal, or private information, and (3) curtailing intrusions into places deemed private or personal” (Nissenbaum 2004, p. 107). Nissenbaum stelt dat het traditionele model een aantal grijze gebieden heeft waarbij het in sommige gevallen onduidelijk is of er nu wel of niet sprake is van privacyschending. Een belangrijk voorbeeld dat Nissenbaum noemt is dat van ‘online profiling’, zoals het geval is bij de personalisering door Google. Zij stelt:

Online profiling is troubling, even when the information gathered is not sensitive (excludes credit card information, for example) and when it takes place on the public. According to the [traditional] framework, therefore, it seems that public surveillance is determined not to be a privacy problem. Because this conclusion is at odds with the intuition and judgment of many people, it warrants more than simple dismissal. In this disparity lie the grounds for questioning the three-principle framework as a universal standard for public deliberations over privacy (ibid. p. 116).

In plaats daarvan stelt Nissenbaum het “contextual integrity” raamwerk voor. Binnen dit raamwerk ligt de nadruk op de context waarin privacykwesties plaatsvinden. “Almost everything—things that we do, events that occur, transactions that take place—happens in a context not only of place but of politics, convention, and cultural expectation” (ibid. p. 119). Omdat deze context sterk van invloed kan zijn op de vraag of er sprake is van privacyschending of niet baseert Nissenbaum haar raamwerk op twee normatieve waarden: “appropriateness” en “distribution”. Wanneer een van de twee waarden wordt geschonden is er volgens Nissenbaum sprake van privacyschending.

Of het delen van informatie appropriate is hangt af van het type informatie en de context waarin de uitwisseling plaatsvindt. Zo deelt een patiënt bijvoorbeeld wel informatie over zijn fysieke gesteldheid met zijn dokter, maar niet met anderen; en bespreekt iemand zijn financiën wel met zijn partner en de bank, maar bijvoorbeeld niet met zijn klanten. Distribution heeft betrekking op de vraag of de informatie binnen de sociale context blijft waarin het is vrijgegeven. Nissenbaum stelt in het geval van online profiling dat het appropriate is om als verkoper gebruikersgegevens te verzamelen om gebruikers/klanten beter te leren kennen, maar wanneer deze gegevens vervolgens gebruikt worden voor marketingdoeleinden en dus buiten de oorspronkelijke context treden, dan is er sprake van een schending van het distribution principe (ibid. p. 134).

Wanneer vanuit het perspectief van contextual integrity gekeken worden naar de wijze waarop informatie wordt gedeeld binnen Google+ dan valt op dat voordat een bericht geplaatst kan worden eerst gekozen moet worden binnen welke Circles die informatie gedeeld wordt. In tegenstelling tot Facebook waar informatie standaard met iedereen gedeeld wordt, laat Google+ de gebruiker expliciet kiezen binnen welke context deze informatie toepasselijk is. Private informatie blijft hierdoor besloten tenzij de gebruiker specifiek aangeeft dat het publiekelijk gemaakt moet worden. Bovendien biedt Google+ de mogelijkheid om statusupdates te vergrendelen. Contacten binnen de Circles waarin bepaalde informatie wordt gedeeld zijn dan niet meer in staat om deze informatie te 're-sharen'. Het distribueren van private informatie buiten de sociale context waarin het geplaatst is wordt hierdoor bemoeilijkt, waardoor de distribution waarde onder gebruikers minder snel geschonden wordt en dus de privacy van de gebruiker beter gewaarborgd blijft. Wanneer Google echter de in besloten kring gedeelde informatie toch gaat gebruiken ten behoeve van hun advertenties dan maakt Google zich zelf wel schuldig aan het schenden van de privacy van haar gebruikers.

### 5.3. Google's +1 knop

Een andere belangrijke functie van Google+ is de +1 knop. Deze knop is vergelijkbaar met de Facebook's 'like' knop die inmiddels alomtegenwoordig is op het Web. De +1 knop is te vinden achter Google's zoekresultaten, afbeeldingen, advertenties en andere Google diensten zoals Blogger en Youtube. Daarnaast kan de +1 knop door websitebeheerders middels een stukje code aan een website worden toegevoegd (google.com 2011f). Wanneer er door een gebruiker op de knop gedrukt wordt, wordt de gedeelde content weergegeven op het Google+ profiel van de gebruiker en wordt bij contacten binnen de gebruikers' Circles deze content in de zoekresultaten gemarkeerd als zijnde gedeeld door een Google+ connectie (zie afbeelding 2). Een nadeel van de +1 knop op het gebied van privacy is echter dat het niet mogelijk is om bepaalde content enkel binnen bepaalde Circles te delen. Wanneer een gebruiker de +1 knop activeert wordt deze zichtbaar binnen de zoekresultaten van alle contacten. Hierdoor kan Nissenbaums' appropriateness in het gedrang komen.

De +1 knop werkt als een sociaal recommender system. Gebruikers kunnen expliciet aangeven welke content hen interesseert en wat hen mogelijk interessant lijkt voor anderen. De handeling om vrienden interessante webcontent aan te raden wordt hiermee heel eenvoudig. Bovendien functioneert de +1 knop ook als een soort van favorieten systeem, doordat het alle webcontent die de gebruiker eerder heeft aangeraden opslaat en archiveert. Hiermee wordt niet enkel het persoonlijk profiel verder opgebouwd, maar hebben gebruikers bovendien invloed op de ordening van zoekresultaten van hun contacten en de

gebruikers die zich niet binnen hun Circle's bevinden (Siegler 2011b). Google had al eerder de stap gemaakt naar deze vorm van 'social search' door een deal aan te gaan met bijvoorbeeld de 'microblogging' dienst Twitter en de fotosite Flickr (ibid.). De komst van de +1 button kan echter gezien worden als een nieuwe stap op het gebied van social search aangezien Google hiermee zelf de data betreffende interactie tussen gebruikers in beheer heeft.

### [Google To Begin Encrypting Searches & Outbound Clicks By Def...](http://searchengineland.com/google-to-begin-encrypting-searches-by-default)

[searchengineland.com/google-to-begin-encrypting-searches...](http://searchengineland.com/google-to-begin-encrypting-searches-by-default) 



by [Danny Sullivan](#) - In 70,307 circles

6 days ago – **Google** will now **begin encrypting searches** that people do by **default**, if they are logged into **Google.com** already through a secure connection.

33 comments on [Danny Sullivan's Google+ post.](#)



[Matt Siltala](#), [Danny Sullivan](#), [Eric Ward](#) and 3 other people +1'd this

Afbeelding 2: Google +1 button (rechtsboven) en een opsomming van het aantal personen dat de webpagina aanraadt (linksonder).

Interessant aan deze vorm van social search is dat hiermee de informatiebubbles waarin gebruikers zich bevinden kunnen worden doorbroken. De onderzoeks- en ontwerpstudio Metahaven (2009) beschrijft aan de hand van het werk van Ronald Burt de werking van "structural holes". Burt beschrijft deze als volgt:

Structural holes are the gaps between non-redundant contacts [contacten met verschillende sociale achtergronden en opvattingen]. As a result of the hole between them, once bridged the two contacts provide network benefits that are in some degree additive rather than overlapping (Burt in Metahaven 2009, p. 188).

Metahaven stelt dat in het geval van zoekmachines er te weinig aandacht besteed wordt aan dit soort structural holes, waardoor informatie die beschikbaar is binnen andere sociale sferen gemarginaliseerd wordt. Metahaven stelt daarom een alternatief zoekmachinemodel voor dat zich focust op de zogenaamde "weak ties" in de periferie van sociale sferen. Deze weak ties zijn verbindingen tussen actoren binnen twee verschillende sociale sferen die als het ware de structural holes overbruggen. Mark Granovetter (1973) stelt dat een persoon met "strong ties" binnen een sociale sfeer min of meer over dezelfde informatie en hetzelfde gedachtengoed beschikt als de andere actoren binnen die sfeer. De actoren met weak ties aan de rand van de sociale sfeer zijn minder beïnvloed door het centrale gedachtengoed en dus eerder vatbaar voor ideeën van buiten de sociale sfeer. Hierdoor zijn weak ties essentieel voor de overdracht van informatie tussen verschillende sferen. Door de focus te leggen op deze weak ties komen de verschillende ideeën en het gedachtegoed dat heerst binnen verschillende sociale sferen beter naar voren en ontstaat er in de zoekresultaten een gedifferentieerder beeld van de werkelijkheid. Metahaven stelt dan ook dat "network power induced by ranking models that value the spanning of social bridges, will hold an incentive for resolving polarized views instead of reinforcing them" (ibid., p. 195).

Wanneer vanuit het perspectief van structural holes gekeken wordt naar het ontwerp van Google+ en het recommender systeem achter de +1 knop, valt op dat het sociale netwerk voornamelijk gericht is op

personen in de naaste omgeving van gebruikers (vrienden, familie, bekenden). Dit zijn over het algemeen personen waar de gebruiker een sterke band mee heeft, waardoor er minder snel alternatieve informatie voorhanden komt. Wanneer deze contacten gebruik maken van de +1 knop is de waarschijnlijkheid dat de gedeelde informatie reeds binnen de informatiebubbel van de gebruiker valt groot en is de informatie dus waarschijnlijk minder vernieuwend of afwijkend. Google+ biedt echter ook standaard een Circle "Following". Deze Circle heeft als ondertitel "People you don't know personally, but whose posts you find interesting". Hierin kunnen dus gebruikers geplaatst worden die niet binnen de sociale sfeer van de gebruiker vallen en waarmee de gebruiker dus een weak tie heeft. Webcontent die middels de +1 knop door deze contacten wordt aangeraden zal dan ook eerder een aanvulling zijn op de informatiebubbel waarin de gebruiker zich begeeft. Een grote diversiteit aan contacten binnen het sociale netwerk kan daardoor dus leiden tot een vermindering van het filter bubble probleem. Door de gebruiker zelf de mogelijkheid te geven om zijn sociale netwerk samen te stellen en dus de relevantie van zoekresultaten enigszins te beïnvloeden neemt de autonomie van de gebruiker dan ook toe.

## 6. Conclusie

Binnen het persoonlijk profiel van de gebruiker dat Google opstelt zijn drie lagen te onderscheiden: het kennisprofiel, het sociale profiel en het fysieke profiel van de gebruiker. Aan de hand van haar service-for-profile model stelt Google allerlei (veelal gratis) diensten en producten beschikbaar voor haar gebruikers die door het verzamelen van gebruiksinformatie bijdragen aan het opbouwen van deze drie lagen. Wat opvalt is dat Google constant bezig is om nieuwe diensten te ontwikkelen en oudere diensten aan te passen of te beëindigen. Hierdoor vindt er ook continu een verschuiving plaats tussen de waarde die Google hecht aan de verschillende lagen binnen het persoonlijk profiel. Daar waar voorheen voornamelijk gefocust werd op het kennisprofiel van de gebruiker, hebben tegenwoordig het sociale profiel en het fysieke profiel van de gebruiker steeds meer invloed op de wijze waarop Google haar gebruikers karakteriseert en van informatie voorziet.

De ontwikkeling van een persoonlijk gebruikersprofiel is voor Google van groot belang voor hun advertentiemodel. Doordat middels het gebruikersprofiel advertenties op de interesses van de gebruiker kunnen worden toegespitst, wordt de effectiviteit ervan sterk verhoogt. Dit is voor Google dan ook een belangrijke commerciële motivatie om het gebruikersprofiel constant te perfectioneren. Maar personalisering heeft ook voordelige effecten voor de gebruiker in diens zoektocht naar informatie. Zo biedt het een antwoord op het door vele critici beschreven rich get richer fenomeen - waarbij reeds populaire websites alsmaar populairder worden - doordat marginale informatie voor bepaalde gebruikers eerder als relevant wordt aangemerkt. Daarnaast zorgt het personaliseren van zoekresultaten ervoor dat het filteren van informatie steeds meer gebeurt op grond van de persoonlijke interesses van de gebruiker, waardoor de voor de gebruiker minder interessante zoekresultaten ook minder hoog gepositioneerd worden. Hier ligt echter ook meteen een van de gevaren van personalisering.

Binnen het kritisch discours rondom personalisering zijn voornamelijk twee belangrijke waarden te onderscheiden die volgens critici in het gedrang zijn: privacy en autonomie. Naast dat er door Google een grote hoeveelheid aan persoonlijke informatie wordt opgeslagen, schuilt er nog een groot privacyprobleem in Google's personaliseringsmechanisme. Doordat Google gebruikers classificeert aan de hand van hun persoonlijke kenmerken, oefent Google een grote macht uit op haar gebruikers. Dit fenomeen wordt ook wel social sorting genoemd. Door aan de hand van een digitale representatie van de gebruiker voorspellingen te doen over diens interesses en voorkeuren en hierop de informatietoever te baseren, worden op den duur de keuzes en keuzemogelijkheden van de gebruiker beïnvloed. Dit heeft dus directe gevolgen op de autonomie van de gebruiker. De macht die uitgaat van social sorting kan zich uiten in de vorm van een zogenaamde filter bubble. De gebruiker zit als het ware gevangen in een informatiebubbel die gebaseerd is op wat Google voor de gebruiker interessant of relevant acht. Dit kan resulteren in het zogenaamde echo chamber effect. Dit houdt in dat individuen door een selectieve informatietoever steeds stelliger gaan geloven in hun eigen standpunten en geen waarde meer hechten aan alternatieve denkwijzen.

Om antwoord te krijgen op de vraag in hoeverre de bovenstaande problemen nog steeds aanwezig zijn met de komst van Google+, is gekeken naar de wijze waarop Google met de identiteit van de gebruiker omgaat en is een aantal belangrijke nieuwe functies uitgelicht. Het sociale en fysieke profiel van de gebruiker spelen een steeds grotere rol binnen de constructie van het persoonlijk gebruikersprofiel. Dit blijkt onder andere uit het feit dat Google+ is ontwikkeld vanuit het gebrek aan een goede identity service op het

Web en vanuit de waarde die Google hecht aan de sociale connecties van gebruikers. Deze verschuiving geeft aan dat Google ook op die vlakken een zo accuraat mogelijk beeld probeert te krijgen van haar gebruikers en dus de verschillen met hun data-double probeert te verkleinen. Dit gaat echter wel gepaard met het verzamelen van nog meer privacygevoelige informatie.

Alhoewel een groot deel van de kritiek rondom personalisering - mede daardoor - nog steeds valide is, valt op dat met de komst van Google+ gebruikers beter in staat zijn om expliciet informatie aan hun persoonlijk profiel toe te voegen. Zo kunnen door middel van Google Sparks persoonlijke interesses aan het profiel worden toegevoegd en stelt Google Circles gebruikers in staat om hun sociale contacten in te delen in bepaalde groepen. Door middel van de +1 knop hebben contacten binnen deze Circles de mogelijkheid om bepaalde webcontent aan te raden. Gebruikers kunnen zich zo abonneren op andere gebruikers die zij interessant achten. Hierdoor komen webpagina's die binnen de Circles van de gebruiker worden aangeraden ook eerder in de zoekresultaten naar voren. Door gebruikers de mogelijkheid te geven om een groot deel van hun persoonlijk profiel zelf in te vullen en invloed uit te oefenen op het categoriseringsproces wordt de autonomie van de gebruiker vergroot. Bovendien kan de invloed van sociale contacten ervoor zorgen dat het filter bubble probleem afneemt. Op het gebied van privacy is de rol van Google+ echter nog steeds discutabel. Door gebruikers de mogelijkheid te geven om informatie enkel binnen bepaalde Circles te delen wordt het gevaar van het schenden van privacy door andere gebruikers desalniettemin enigszins verkleint. Bovendien stelt Google dat informatie die gedeeld wordt op Google+ niet voor haar advertentienetwerk gebruikt wordt. Het blijft echter onduidelijk in hoeverre Google discreet met de informatie omgaat en het distribution principe niet schendt.

Met de komst van Google+ is een groot deel van kritiek binnen het kritisch discours rondom personalisering nog steeds valide. Google blijft grote hoeveelheden aan informatie verzamelen om zo de gebruiker steeds beter te kunnen analyseren en interpreteren. De wijze waarop de informatie verzameld wordt is echter aan verandering onderhevig. De gebruiker heeft met de komst van Google+ een steeds groter aandeel in de constructie van diens eigen profiel. Door gebruikers expliciet de mogelijkheid te geven om persoonlijke gegevens en voorkeuren aan het systeem toe te voegen, is het ontsluiten van persoonlijke interesses door Google niet meer volledig geautomatiseerd. Dit kan de verschillen tussen de gebruiker en diens data-double verkleinen en dus zo de problemen op het gebied van social sorting laten doen afnemen.

Het is waarschijnlijk dat Google haar sociale netwerk in de toekomst nog meer zal gaan inzetten om een goede online identity service te bieden. Google heeft hiermee waardevolle informatie in handen voor adverteerders. Bovendien is het met een dergelijke service mogelijk voor gebruikers om - zoals nu al het geval is met onder andere OpenID en Facebook - met hun Google account in te loggen op andere websites. Google krijgt daardoor inzage in het handelen van gebruikers op websites buiten hun eigen netwerk. Bovendien is het met een goede identity service in staat om een online reputatieservice op te zetten. De toekomst zal uitwijzen of Google+ net zo groot wordt als Facebook en de race tot "identity gatekeeper" (Ingram 2011b) zal winnen. Maar één ding is zeker: Google zal haar zoektocht naar u voortzetten.



## 7. Referenties

- Barabási, A. L. (2002). *Linked: the new science of networks*. Cambridge: Perseus Publishing.
- Barlow, J. P. (1996) "A Cyberspace Independence Declaration." Bezocht op November 11, 2011, from [http://w2.eff.org/Censorship/Internet\\_censorship\\_bills/barlow\\_0296.declaration](http://w2.eff.org/Censorship/Internet_censorship_bills/barlow_0296.declaration)
- Battelle, J. (2003). "The "Creeping Googlization" Meme." Bezocht op Oktober 18, 2011, from [http://battellemedia.com/archives/2003/12/the\\_creeping\\_googlization\\_meme.php](http://battellemedia.com/archives/2003/12/the_creeping_googlization_meme.php).
- Battelle, J. (2005). *The search: how Google and its rivals rewrote the rules of business and transformed our culture*. New York: Portfolio.
- Bergman, M. K. (2001) The Deep Web: Surfacing Hidden Value. *JEP*, 7, Bezocht op September 28, 2011, from <http://quod.lib.umich.edu/cgi/t/text/text-idx?c=jep;view=text;rgn=main;idno=3336451.0007.104>
- Bozdag, E., & Timmermans, J. (2011). "Values in the Filter Bubble: Ethics of Personalization Algorithms in Cloud Computing." In Proceedings of the 13th IFIP TC13 Conference on Human-Computer Interaction (Interact), Lisbon.
- BrandZ (2010, 22-11-2011). "Top 100 Most valuable global brands 2010." Bezocht op November 22, 2011, from [http://c1547732.cdn.cloudfiles.rackspacecloud.com/BrandZ\\_Top100\\_2010.pdf](http://c1547732.cdn.cloudfiles.rackspacecloud.com/BrandZ_Top100_2010.pdf).
- Brey, P. (2000). "Disclosive computer ethics." *Computers and Society* 30(4), 10-16.
- Brin, S., & Page, L. (1998). The Anatomy of a Large-Scale Hypertextual Web Search Engine. In Seventh International World-Wide Web Conference, Brisbane, Australia
- Carvin, A. (2011). "Regarding my prior post on G+ and real names." Bezocht op November 16, 2011, from <https://plus.google.com/117378076401635777570/posts/CjM2MPKocQP>.
- Cheung, H. (2008). "Chrome is a security nightmare, indexes your bank accounts." Bezocht op Oktober 20, 2011, from <http://www.tgdaily.com/security-features/39176-chrome-is-a-security-nightmare-indexes-your-bank-accounts>.
- Cowan, D. (2011). "On Pseudonyms: Transparency And Free Expression Are Not Mutually Exclusive." Bezocht op November 16, 2011, from <http://techcrunch.com/2011/09/05/pseudonyms-transparency-free-expression/>.
- Deleuze, G. (1992). "Postscript on the Societies of Control." *October*, 59, 3-7.

- Diaz, A. M. (2005). *Through the Google Goggles: Sociopolitical Bias in Search Engine Design*. *Science, Technology and Society*. Stanford: Stanford University. Interdisciplinary Honors.
- Empson, R. (2011). "Google Apps Coming To Google+ 'Within Days'; Company Taking A 'Cautious Approach' To APIs." Bezocht op November 16, 2011, from <http://techcrunch.com/2011/10/19/google-apps-coming-to-google-within-days-company-taking-a-cautious-approach-to-apis/>.
- Evans, J. (2011a). "I Believe In Google Plus." Bezocht op November 16, 2011, from <http://techcrunch.com/2011/10/22/i-believe-in-google-plus/>.
- Evans, J. (2011b). "Google Plus Has A Problem. Fear Not: I Have A Solution." Bezocht op November 16, 2011, from <http://techcrunch.com/2011/07/24/google-plus-has-a-problem-fear-not-i-have-a-solution/>.
- Foucault, M. (1979). *Discipline and punish: the birth of the prison*. New York: Vintage Books.
- Fried, I. (2008). "Google's Omnibox could be Pandora's box." Bezocht op Oktober 20, 2011, from [http://news.cnet.com/8301-13860\\_3-10031661-56.html](http://news.cnet.com/8301-13860_3-10031661-56.html).
- Fuez, M., & Fuller, M., & Stalder, F. (2011) Personal Web searching in the age of semantic capitalism: Diagnosing mechanisms of personalisation. *First Monday*, 16(2), Bezocht op September 28, 2011, from <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3344/2766>
- Goldman, A. I. (2008). The Social Epistemology of Blogging. In J. van den Hoven & John Weckert (eds.) *Information Technology and Moral Philosophy*. (pp. 111-122). Cambridge: Cambridge University Press
- Goldman, E. (2006). "Search Engine Bias and the Demise of Search Engine Utopianism." *Yale Journal of Law and Technology*, 8, 188-200.
- Goodwin, D. (2011). "The Google+ Identity Service Project." Bezocht op November 16, 2011, from <http://searchenginewatch.com/article/2105026/The-Google-Identity-Service-Project>.
- Haggerty, K. D. and R. V. Ericson (2000). "The surveillant assemblage." *Journal of Sociology*, 51(4), 605-622.
- Halavais, A. (2009). *Search Engine Society*. Cambridge: Polity Press.

Hoofnagle, C. J.. (2009) Beyond Google and evil: How policy makers, journalists and consumers should talk differently about Google and privacy. *First Monday*, 14(4), Bezocht op September 28, 2011, from <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2326/2156>

google.com (2008). "SearchWiki: make search your own." Bezocht op Oktober 20, 2011, from <http://googleblog.blogspot.com/2008/11/searchwiki-make-search-your-own.html>.

google.com (2009). "Code of Conduct." Bezocht op December 20, 2011, from <http://investor.google.com/corporate/code-of-conduct.html>.

google.com (2010). "Stars make search more personal." Bezocht op Oktober 20, 2011, from <http://googleblog.blogspot.com/2010/03/stars-make-search-more-personal.html>.

google.com (2011a). "2011 Financial Tables." Bezocht op Oktober 14, 2011, from <http://investor.google.com/financial/tables.html>.

google.com (2011b). "Hide sites to find more of what you want." Bezocht op Oktober 20, 2011, from <http://googleblog.blogspot.com/2011/03/hide-sites-to-find-more-of-what-you.html>.

google.com (2011c). "Check in with Google Latitude." Bezocht op November 1, 2011, from <http://googleblog.blogspot.com/2011/02/check-in-with-google-latitude.html>.

google.com (2011d). "Google Wallet Privacy Policy." Bezocht op Oktober 15, 2011, from <http://www.google.com/wallet/privacy.html>.

google.com (2011e). "Introducing the Google+ project: Real-life sharing, rethought for the web." Bezocht op November 17, 2011, from <http://googleblog.blogspot.com/2011/06/introducing-google-project-real-life.html>.

google.com (2011f). "The +1 button for websites: recommend content across the web." Bezocht op November 18, 2011, from <http://googleblog.blogspot.com/2011/06/1-button-for-websites-recommend-content.html>.

google.com (n.d. a). "Google-geschiedenis." Bezocht op November 22, 2011, from <http://www.google.nl/intl/nl/about/corporate/company/history.html>.

google.com (n.d. b). "Hoe kan ik de personalisatie van zoekresultaten op basis van mijn zoekgeschiedenis beperken?." Bezocht op November 10, 2011, from <http://www.google.com/support/accounts/bin/answer.py?answer=54048>.

- Granovetter, M. S. (1973). "The Strength of Weak Ties." *American Journal of Sociology*, 78(6), 1360-1380.
- Ingram, M. (2011a). "A Google Monopoly Isn't the Point." Bezocht op December 20, 2011, from <http://www.businessweek.com/technology/a-google-monopoly-isnt-the-point-09232011.html>
- Ingram, M. (2011b). "Google Confirms It Aims to Own Your Online ID." Bezocht op December 20, 2011, from <http://www.businessweek.com/technology/google-confirms-it-aims-to-own-your-online-id-08292011.html>
- Introna, L. D., Nissenbaum, Helen (2000). "Shaping the Web: Why the politics of search engines matters." *Information Society*, 16(3), 169-185.
- Jenkins, R. (2000). "Categorization: Identity, Social Process and Epistemology." *Current Sociology*, 48(7), 7-25.
- Kapor, M. (1993) Where Is the Digital Highway Really Heading? The Case for a Jeffersonian Information Policy. *Wired*, 1.03, Bezocht op November 21, 2011, from [http://www.wired.com/wired/archive/1.03/kapor.on.nii.html?pg=2&topic=&topic\\_set=](http://www.wired.com/wired/archive/1.03/kapor.on.nii.html?pg=2&topic=&topic_set=)
- Katz, J. (1995) The Age of Paine. *Wired*, 3.05, Bezocht op November 21, 2011, from [http://www.wired.com/wired/archive/3.05/paine.html?pg=1&topic=&topic\\_set=](http://www.wired.com/wired/archive/3.05/paine.html?pg=1&topic=&topic_set=)
- Keenoy, K. and M. Levene (2005). "Personalisation of web search." *Intelligent Techniques for Web Personalisation*, 3169, 201-228.
- Lobet-Maris, C. (2009). From Trust to Tracks: A technology Assessment Perspective Revisited. In K. Becker and F. Stalder (eds.) *Deep Search: the politics of search beyond Google*. (pp. 73-84). Wenen: Studien Verlag.
- Lovink, G. (2009) Society of the Query. The Googlization of Our Lives. In K. Becker and F. Stalder (eds.) *Deep Search: the politics of search beyond Google*. (pp. 45-53). Wenen: Studien Verlag.
- Lyon, D. (2001). *Surveillance society: monitoring everyday life*. Philadelphia: Open University Press.
- Lyon, D. (2007). *Surveillance Studies: An Overview*. Cambridge: Polity Press.
- msn.com. (2010). "Czech Republic bans Google 'Street View'." Bezocht op December 22 2011, from [http://www.msnbc.msn.com/id/39302384/ns/technology\\_and\\_science/t/czech-republic-bans-google-street-view/#.TvMEAyOP2IU](http://www.msnbc.msn.com/id/39302384/ns/technology_and_science/t/czech-republic-bans-google-street-view/#.TvMEAyOP2IU).

- Metahaven. (2009) *Peripheral Forces: On the Relevance of Marginality in Networks*. In K. Becker and F. Stalder (eds.) *Deep Search: the politics of search beyond Google*. (pp. 185-197). Wenen: Studien Verlag.
- News, B. (2007). "Google ranked 'worst' on privacy." Bezocht op November 22 2011, from <http://news.bbc.co.uk/2/hi/technology/6740075.stm>.
- Nissenbaum, H. (2004). "Privacy as Contextual Integrity." *Washington Law Review*, 79(1), 101-158.
- Pan, B., & Hembrooke, H., & Joachims, T., & Lorigo, L., & Gay, G., & Granka, L. (2007). "In Google We Trust: Users' Decisions on Rank, Position, and Relevance." *Journal of Computer-Mediated Communications*, 12, 801-823.
- Palmer, M. (2007). "EU probes Google grip on data." Bezocht op December 22 2011, from <http://www.ft.com/cms/s/2/dc89ec96-0a24-11dc-93ae-000b5df10621.html#axzz1hG8XounU>.
- Pariser, E. (2011). *What the Internet is Hiding From You*. New York: The Penguin Press.
- Pasquinelli, M. (2009) Google's PageRank Algorithm: A Diagram of the Cognitive Capitalism and the Rentier of the Common Intellect. In K. Becker and F. Stalder (eds.) *Deep Search: the politics of search beyond Google*. (pp. 152-162). Wenen: Studien Verlag.
- Pinola, M. (2011). "Google+ Floodgates Open -- Everyone Gets 150 Invites and an Invite Page." Bezocht op November 23, 2011, from [http://www.pcworld.com/article/237454/google\\_floodgates\\_open\\_everyone\\_gets\\_150\\_invites\\_and\\_an\\_invite\\_page.html](http://www.pcworld.com/article/237454/google_floodgates_open_everyone_gets_150_invites_and_an_invite_page.html).
- Price, G. (2006). "Google News Launches Recommendation Service, List of Popular Stories Also Now Available." Bezocht op November 1, 2011, from <http://searchenginewatch.com/article/2059819/Google-News-Launches-Recommendation-Service-List-of-Popular-Stories-Also-Now-Available>.
- Resnick, P., & Varian, H. R. (1997). "Recommender Systems." *Communications of ACN* 40(3), 56-58.
- Rheingold, H. (1993). *The virtual community: homesteading on the electronic frontier*. Cambridge: MIT press.
- Rieder, B. (2009) Democratizing Search? From Critique to Society-oriented Design. In K. Becker and F. Stalder (eds.) *Deep Search: the politics of search beyond Google*. (pp. 133-151). Wenen: Studien Verlag.

- Rogers, R. (2009) The Googlization Question, and the Inculpable Engine. In K. Becker and F. Stalder (eds.) *Deep Search: the politics of search beyond Google*. (pp. 173-184). Wenen: Studien Verlag.
- Röhle, T. (2007). Desperately seeking the consumer: Personalized search engines and the commercial exploitation of user data. *First Monday*, 12(9), Bezoct op OkBezoct optober 14, 2011, from <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2008/1883>
- Salkever, A. (2003). "Google Here, There, and Everywhere." Bezoct op November 18, 2011, from [http://www.businessweek.com/technology/content/dec2003/tc20031216\\_9018\\_tc047.htm](http://www.businessweek.com/technology/content/dec2003/tc20031216_9018_tc047.htm).
- Shiels, M. (2010). "Google admits wi-fi data collection blunder." Bezoct op December 22, 2011, from <http://news.bbc.co.uk/2/hi/technology/8684110.stm>.
- Siegler, M. (2011b). "Google Search Finally Going Fully Social With Shared Twitter Links And Even Quora Data." Bezoct op November 18, 2011, from <http://techcrunch.com/2011/02/17/google-social-search/>.
- Simon, B. (2005). "The Return of Panopticism: Supervision, Subjection and the New Surveillance." *Surveillance Society*, 3(1), 1-20.
- Singel, R. (2011). "Google+ vs. Facebook on Privacy: + Ahead On Points — For Now." Bezoct op November 17, 2011, from <http://www.wired.com/epicenter/2011/06/google-facebook-privacy/>.
- Slawski, B. (2011). "How Google Might Rank User Generated Web Content in Google + and Other Social Networks." Bezoct op November 16, 2011, from <http://www.seobythesea.com/2011/07/how-google-might-rank-user-generated-web-content-in-google-and-other-social-networks/>.
- Stalder, F., & Mayer, C. (2009). The Second Index: Search Engines, Personalization and Surveillance. In K. Becker and F. Stalder (eds.) *Deep Search: the politics of search beyond Google*. (pp. 98-115). Wenen: Studien Verlag.
- Stamoulis, N. (2010). "Top Ten Google SEO Ranking Factors." Bezoct op November 20, 2011, from <http://www.searchengineoptimizationjournal.com/2010/01/14/google-ranking-factors/>.
- Sunstein, C. R. (2007). *Republic.com*. Princeton: Princeton University Press.
- Ujjin, S. B., P. J. (2001) Building a Lifestyle Recommender System. In Poster Proceedings of the 10th International World Wide Web Conference, Hong Kong.

Vaidhyanathan, S. (2011). *The Googlization of everything: (and why we should worry)*, Los Angeles: University of California Press.

Vaidhyanathan, S. (2007). "The Googlization of Everything: How one company is disrupting commerce, culture, and community." Bezocht op Oktober 18, 2011, from [http://www.googlizationofeverything.com/2007/09/where\\_is\\_this\\_book\\_going.php](http://www.googlizationofeverything.com/2007/09/where_is_this_book_going.php).

van der Vlies, L. C. D. J. (2008). Googlization: A new form of mass media critique. *Media Studies*. Amsterdam, University of Amsterdam. M.A.

Weisberg, J. (2011). "Bubble Trouble: Is Web personalization turning us into solipsistic twits?" *Slate Magazine*. Bezocht op November 10, 2011, from [http://www.slate.com/articles/news\\_and\\_politics/the\\_big\\_idea/2011/06/bubble\\_trouble.html](http://www.slate.com/articles/news_and_politics/the_big_idea/2011/06/bubble_trouble.html).

Zimmer, M. (2008a) The Externalities of Search 2.0: The Emerging Privacy Threats when the Drive for the Perfect Search Engine meets Web 2.0. *First Monday*, 13(3), Bezocht op August 15, 2011, from <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/2136/1944>

Zimmer, M. (2008b). "The Gaze of the Perfect Search Engine: Google as an Infrastructure of Dataveillance." *Information Science and Knowledge Management* 14(2): 77-99.