

Quaero Quaestio

Een exploratief onderzoek naar online methoden van contentontsluiting en de handelswijzen en zoekdoelen van internetgebruikers.

Masterthesis Media en Journalistiek
Dennis Snijders 2009
Studentnummer 308508

Faculteit der Historische en Kunstwetenschappen
Erasmus Universiteit Rotterdam

Begeleiders: drs. Mijke Slot MA & drs. Silvain de Munck
Tweede lezer: Dr. Marc Verboord
2008/2009

Inhoudsopgave

1. INLEIDING	5
1.1 Uitwisseling van informatie.....	5
1.2 Effecten.....	7
1.2.1 Negatieve effecten	7
1.2.2 Positieve effecten	9
1.3 Toegangspoorten.....	11
1.4 Problematiek.....	15
1.5 Doelstelling	16
1.6 Hoofd en Deelvragen	17
1.6.1 Hoofdvraag:	17
1.6.2 Deelvragen.....	17
1.7 Relevantie	17
1.8 Leeswijzer.....	18
2. CONCEPTUEEL RAAMWERK	19
2.1 Het internet.....	19
2.2 De Long Tail.....	20
2.3 Wat zijn data, informatie en content?	23
2.3.1 Data	23
2.3.2 Informatie	24
2.3.3 Content	24
2.4 Metadata	26
2.5 De methoden.....	27
2.5.1 Introductie	27
2.5.2 Zoekmachines	28
2.5.3 Folksonomieën	29
2.5.4 Webdirectories	29
2.5.5 Weblogs	30
2.5.6 Andere methoden	30
2.6 Zoekdoelen.....	33
2.6.1 Navigationeel.....	33
2.6.2 Informationeel.....	34
2.6.3 Transactioneel	34
2.7 Conclusie.....	36
3. METHODOLOGISCHE VERANTWOORDING	38
3.1 Verantwoording dienstenanalyse.....	38
3.1.1 Operationalisatie	40
3.2 Verantwoording gebruikersonderzoek	42
3.2.1 Onderzoekstype.....	42
3.2.2 Verspreidingmethode.....	44

3.2.3 Representativiteit	45
3.2.4 Surveyconstructie	46
3.2.5 Dataverwerking	48
3.3 Verantwoording verdieping onderzoeksresultaten	49
4. RESULTATEN DIENSTENANALYSE	50
4.1 Zoekmachines.....	50
4.1.1 Diensten.....	50
4.1.2 Methode	56
4.1.3 Conclusie.....	58
4.2 Folksonomieën	61
4.2.1 Diensten.....	61
4.2.2 Methode	65
4.2.3 Conclusie.....	67
4.3 Webdirectories	71
4.3.1 Diensten.....	71
4.3.2 Methode	75
4.3.3 Conclusie.....	76
4.4 Weblogs	79
4.4.1 Diensten.....	79
4.4.2 Methode	81
4.4.3 Conclusie.....	82
4.5 Conclusie dienstenanalyse	84
5. RESULTATEN GEBRUIKERSONDERZOEK	85
5.1 Populatie.....	85
5.2 Stellingen	86
5.3 Open vraag.....	87
5.4 Naamsbekendheid en gebruiksfrequentie	89
5.4.1 Diensten.....	89
5.4.2 Methoden	91
5.4.3 Mannen versus vrouwen	92
5.4.4 Leeftijd	93
5.4.5 Tijdsbesteding.....	94
5.4.6 Aanvulling door respondenten	95
5.5 Navigatieel, informatieel en transactioneel gebruik.....	97
5.5.1 Navigatieel.....	98
5.5.2 Informatieel.....	100
5.5.3 Transactioneel	102
5.6 Conclusie gebruikersonderzoek.....	104
6. VERDIEPING ONDERZOEKSRESULTATEN	107
6.1 Zoekmachines.....	107
6.2 Folksonomieën	109
6.3 Webdirectories	111
6.4 Weblogs	113
6.5 Conclusie.....	115

7. CONCLUSIE	117
7.1 Onderzoeksverloop.....	117
7.2 Resultaten.....	118
7.3 Kritische reflectie en vervolgonderzoek.....	120
LITERATUURLIJST	122
BIJLAGEN	127
B1. Lijst van figuren en tabellen	127
B2. Topiclist dienstenanalyse	129
B3. Survey.....	131

1. Inleiding

“In the historical perspective, more information has almost always been a very good thing. Information made possible the dissemination of culture, the development of commerce and technology - and it was one of the main driving forces behind the large-scale establishment of democracy and human rights. There was quite literally no downside to having more information. The dissemination of information empowered people. Oddly enough, the Information Age has been named for something which once conferred only benefits, and which is now increasingly seen as a problem.” (Cristopher Carlson, 2003:169-173)

1.1 Uitwisseling van informatie

De mens heeft sinds het ontstaan van de mensheid de drang om kennis te vergaren en deze te delen met anderen.

Een vroeg voorbeeld hiervan zijn de grotschilderingen die wereldwijd gevonden zijn. Met de toename in kennis die de mens bezat, groeide ook de middelen om deze kennis uit te wisselen; de gesproken taal ontstond en later volgde ook geschreven taal. Met de uitvinding van het schrift en, minstens net zo belangrijk, de uitvinding van papier (papyrus) kon kennis bewaard blijven en had de mens de mogelijkheid om informatie vast te leggen en uit te wisselen over grote afstanden. Een deel van deze geschreven kennis werd verzameld in bibliotheken. In de grootste bibliotheek uit de oudheid, de Koninklijke Bibliotheek van Alexandrië (300 BC), waren al 700.000 boekrollen opgeslagen. In deze bibliotheken werd door bibliothecarissen gedocumenteerd welke boekrollen er aanwezig waren en waar deze terug te vinden waren.

Er is sinds die tijd veel veranderd. De mens heeft enorm veel kennis en informatie verworven. Ook het arsenaal aan media waarop deze opgeslagen en uitgewisseld kan worden, is sterk uitgebreid. Tot voor kort hebben mensen deze kennis en informatie altijd op moeten slaan op een fysiek medium. Hierdoor zijn er relatief hoge kosten verbonden geweest aan de (re)productie, verspreiding en de opslag van al deze kennis.

Deze fysieke beperking kwam grotendeels te vervallen met de intrede van de digitale revolutie, de uitvinding van de computer en later het internet. De kosten van het vastleggen en uitwisselen van informatie namen drastisch af. Alsof iemand de doos van Pandora had geopend, nam de productie en verspreiding van informatie hierdoor onvoorstelbaar snel toe.

In eerste instantie was het voor de gebruikers alleen mogelijk om de informatie die zij op internet vonden te consumeren. Deze informatie was afkomstig van een relatief beperkt aantal producenten. De manier waarop internet vandaag de dag is ingericht, maakt het voor gebruikers mogelijk en makkelijk om zelf informatie toe te voegen.

Deze vereenvoudiging van productie en uitwisseling van informatie, ook wel *Web2.0* (O'Reilly, 2005) genoemd, zorgt voor grote veranderingen en heeft de ontwikkeling van internet in een stroomversnelling gebracht.

De hoeveelheid informatie die op internet te vinden is neemt nog sneller toe.

Deze toename wordt mede veroorzaakt door het stijgende aantal gebruikers die een producerende rol innemen en gegevens op internet zetten. Deze gebruikers staan ook wel bekend onder de naam *prosumers*, een samentrekking van producer en consumer (Toffler, 1981). Deze gegevens, veelal in de vorm van tekst, beeld, geluid, video en animaties, worden ook wel *content* genoemd.

Een bekend voorbeeld hiervan is de website Wikipedia, een volwaardige encyclopedie op internet geschreven en bijgehouden door haar gebruikers.

Andere sites zoals Yousendit en Rapidshare verschaffen hun gebruikers de mogelijkheid om op grote schaal bestanden te verzenden. Op websites als Youtube en Flickr delen en becommentariëren gebruikers ongekeerde hoeveelheden beeldmateriaal vanuit de gehele wereld.

1.2 Effecten

Internet is pas sinds het begin van de jaren negentig beschikbaar voor de consument. Dit is relatief erg kort. Er wordt veel gespeculeerd over de mogelijke (toekomstige) effecten van de informatietoename en de omgangswijze van de gebruikers hiermee.

1.2.1 Negatieve effecten

Eén van de sceptici op dit gebied is Andrew Keen. Hij stelt dat door de inmenging van amateurs in de productie van informatie, de kwaliteit van content achteruit gaat. Doordat de gebruiker de mogelijkheid krijgt om zijn of haar eigen content te publiceren, wordt het internet gevuld met *“de nonsens van amateurs; mensen die denken dat zij muzikant, journalist of wetenschapper zijn”* (Martijn & Nihom, 2008). Keen vindt het vooral problematisch dat er geen *gatekeepers* meer zijn om goede en slechte content te onderscheiden.

Gatekeepers zijn de personen, in het verleden voornamelijk professionals, die controle hebben over de stroom van informatie, meestal vanuit de massamedia richting het individu. *“De enige poortwachter op internet is de zogenaamde wijsheid van de massa.”* (Martijn & Nihom 2008) Dit heeft volgens Keen tot gevolg dat de resultaten van Google en inhoud van Wikipedia niet de objectieve waarheid bevatten, maar slechts de meest populaire opvattingen.

Naast invloed op het soort informatie dat beschikbaar is, heeft de grote hoeveelheid beschikbare informatie ook gevolgen voor de gebruikers van deze informatie. Voor de komst van internet was al bekend dat grote informatiestromen negatieve gevolgen voor mensen kunnen hebben. Vrijwel iedereen heeft dit wel eens ervaren. Van scholieren met het leren van een omvangrijk proefwerk tot advocaten die zich moeten inlezen voor een belangrijke zitting. De belangrijkste beperking is eenvoudigerwijs die van tijd; een mens heeft slechts 24 uur per dag te besteden. Er kan dus een probleem ontstaan als de hoeveelheid informatie de beschikbare verwerkingstijd overstijgt. Als dit punt bereikt is, wordt er gesproken van een *information overload*. (O'Reilly, 1980:684-696) Dit wordt ook aangestipt door Marshall

McLuhan (1967). Hij zegt hierover: *“One of the effects of living with electric information is that we live habitually in a state of information overload. There's always more than you can cope with.”*

Toch is het probleem niet zo simpel te stellen als *teveel informatie, of te weinig tijd*. Het is belangrijk om de factoren die bij kunnen dragen aan de potentiële overbelasting verder te uit te diepen. Pas als de oorzaak begrepen wordt, kan het (mogelijke) probleem effectief aangepakt worden. De overbelasting van informatie die een persoon kan ervaren kan samengevat veroorzaakt worden door de invloed van een vijftal elementen. (Eppler & Mengis, 2004)

De eerste drie hiervan zijn de eigenschappen van de informatie zelf, de taken die verbonden zijn aan de ontvangen informatie en de structuur waarbinnen de informatie wordt aangeboden.

De laatste twee elementen vormen de kern van dit onderzoek, namelijk allereerst de invloed van de ontvanger van de informatie en vervolgens de technologie die gebruikt wordt om informatie te verwerken en hoe deze technologie gebruikt wordt.

De overbelasting die veroorzaakt kan worden door digitale informatie en technologie wordt er ook wel *technostress* (Rosen & Weil, 1997) genoemd.

Er zijn een aantal problemen die een individu kan ondervinden wanneer deze wordt blootgesteld aan een (te) grote hoeveelheid informatie. (Eppler et al., 2004)

Het eerste gevolg van teveel informatie kan zijn dat een persoon moeite krijgt met het onderscheid maken tussen de relevante en irrelevante zaken.

Ten tweede kan het ook tot gevolg hebben dat een groot deel van de informatie aan de kant geschoven wordt en genegeerd wordt vanwege een toename van de selectiviteit van de persoon.

Als derde kan het zien van het verband tussen de details en het geheel van de informatie als steeds moeilijker ervaren wanneer de kwantiteit ervan toeneemt. Meer informatie kost meer tijd om te verwerken en hierdoor zal het proces van afweging dat een persoon nodig heeft om tot een beslissing te komen langer duren. Het kan zelfs veroorzaken dat de beslissing die uiteindelijk genomen wordt niet accuraat is.

Hieraan kan nog toegevoegd worden dat de gebruiker problemen kan ervaren in het stellen van goede prioriteiten en het moeilijker kan vinden om zich eerder opgeslagen informatie te herinneren (Schick, Gordon & Haka, 1990:199-220).

Hoewel informatie overload al bestond voor de komst van internet, lijkt het digitaliseren van informatie en het verspreiden van content via internet dit probleem te hebben versterkt. Op internet ontbreken bijvoorbeeld de eerder genoemde gatekeepers. Waar gebruikers voor de komst van internet redelijkerwijs aan konden nemen dat de informatie, die in een studieboek gevonden kon worden waar was, is dit bij de online variant nu lang niet altijd het geval.

De enige beperking die de producent van online content wordt opgelegd, is die van de grenzen van het medium zelf. Hierdoor is er vrijwel geen enkele waarborg over de vorm, inhoud, intentie en daadwerkelijke aanwezigheid van de geplaatste content.

1.2.2 Positieve effecten

Niet iedereen ziet internet als een bedreiging voor het menselijke vind- en selectieproces. Internet is niet alleen een veroorzaker van een enorme stroom aan informatie, het kan ook een grote bijdrage leveren aan het beheersen van deze informatiestroom. Door de huidige toegankelijkheid van internet kunnen gebruikers een veel actievere rol spelen in de verwerking en stroomlijning van informatie. Waar de gebruikers van internet in eerste instantie voornamelijk een consumerende rol hadden en later steeds meer een producerende rol innamen, is er nu in toenemende mate een *selecterende* rol voor de internetgebruikers weggelegd. Deze selecterende rol is bijvoorbeeld duidelijk terug te vinden bij websites als Delicious, waar gebruikers zelf moeten aangeven welke informatie wel en niet in het systeem van de dienst opgenomen moet worden. Ook bij websites als Youtube kan er geselecteerd worden, al is het op een andersoortige wijze. Gebruikers kunnen hier aan de hand van een aantal sterren een waardering toevoegen aan een filmpje. Door deze selectie wordt het voor andere gebruikers zichtbaar welke video's wel of niet de moeite zijn. Het is ook mogelijk voor gebruikers om (andermans) ongepaste video's te laten verwijderen.

Iemand die vertrouwt op de positieve effecten van technologie en de inventiviteit en flexibiliteit van de mens is Yochai Benkler. De titel van zijn boek *The Wealth of Networks* (2006) is een duidelijke verwijzing naar de, voor hem tot het verleden behorende, opvattingen van econoom Adam Smith (1776), uit het vrijwel gelijknamige boek, *The Wealth of Nations*.

Benkler spreekt van een *networked information society*; een maatschappij waarin de gedecentraliseerde productie van informatie en kennis het overneemt van economisch gedreven massaproductie. De toename in technieken en media om informatie te produceren en uit te wisselen zal volgens hem dan ook zorgen voor een situatie waarin op steeds kortere intervallen content gecreëerd kan worden. Hierdoor zal de omvang en diversiteit enorm toenemen.

Gebruikers zullen vervolgens kennis kunnen verwerven en keuzes kunnen maken die dichter aansluiten bij hun behoeften. Een van de meest sprekende voorbeelden van de filosofie van Benkler is Wikipedia. Hier wordt door een internationale en ongestructureerde groep vrijwilligers, zonder enig economisch gewin, een volwaardige encyclopedie geschreven (Giles, 2005:900-901). De informatie die in Wikipedia gevonden kan worden heeft het nadeel dat deze in een constante staat van beweging is en dat de informatie (onjuist) beïnvloed kan worden door iedere gebruiker. Maar dit is nu juist ook de kracht van Wikipedia; de informatie op de site is enorm dynamisch en kan hierdoor snel veranderingen in een kennisgebied opnemen. Ook zorgen de gebruikers dat de informatie vrijwel continue gecontroleerd wordt op inhoud.

Zoals Socrates reeds lange tijd geleden sprak: *“Er is slechts één goed: kennis, en slechts één kwaad: onwetendheid”*. (Laertius, 470 BC-399 BC)

Deze rode draad kan teruggevonden worden in vrijwel alle verschillende opvattingen die er bestaan in de discussie over de enorme vlucht die de creatie en uitwisseling van informatie momenteel neemt. De angsten van de mensen die er negatief tegenover staan zijn niet afkomstig van de (potentiële) waarde van informatie, maar van de kwaliteit en bereikbaarheid ervan (Keen, 2008), de beperkingen in het verwerken ervan (Eppler et al. 2004; Rosen et al. 1997) of de constante blootstelling eraan (McLuhan, 1967). Positivisten als

Benkler (2006) baseren hun utopische beelden op de vanzelfsprekendheid van de openbaarheid van informatie en kennis en de flexibiliteit van de mens.

De grote vraag is dus of de mens in staat is om een systeem of methodiek te creëren waarin deze vrije toegang heeft tot de voor hem of haar relevante informatie, zonder geremd te worden door negatieve aspecten.

De toename in beschikbare en potentieel zeer waardevolle content heeft een achilleshiel. Content verandert pas in waardevolle informatie als deze bij de juiste persoon terecht komt (Kock, McQueen & Corner, 1997). Het is dus essentieel dat informatie vindbaar is.

Nu zijn er veel en uiteenlopende soorten informatie op internet te vinden; informatie kan bijvoorbeeld van hoge en lage kwaliteit zijn, het kan door amateurs of professionals gemaakt zijn, het kan tekst zijn, maar ook beeld of geluid, enzovoort. Door deze verschillen is het ingewikkeld om vorm te geven aan een systeem waarin al deze soorten informatie passen en terug te vinden zijn.

Er zijn dan ook verschillende diensten op internet die informatie vindbaar proberen te maken.

1.3 Toegangspoorten

“Microsoft bids \$44.6 billion for Yahoo” - U.S. business

“Google Acquires YouTube For \$1.65 Billion” – Internetnews.com

“Yahoo to buy Overture for \$1.63 billion” – news.cnet.com

Zoals de headlines hierboven duidelijk aangeven is het voor gebruikers vindbaar maken van informatie big business.

Toegangspoorten zijn de websites die de gebruikers ondersteunen met het ontsluiten van informatie en het behalen van hun zoekdoelen en behoren tot enkele van de meest bezochte sites op het internet. Doordat vele miljoenen bezoekers gebruik maken van deze toegangspoorten, zijn deze websites zeer gewilde plaatsen voor marketeers om hun producten aan de man te brengen. Er worden hierdoor torenhoge bedragen betaald voor advertentieruimtes op deze sites. Om erachter te komen dat dit een lucratieve activiteit is hoeft men slechts in het financieel katern van een willekeurige krant te kijken. Zo wordt de huidige marktwaarde van zoek-gigant Google geschat op zo'n 100 miljard euro en was Yahoo

afgelopen maanden niet tevreden met een overnamebod van Microsoft van “slechts” 30 miljard euro.

Ter verduidelijking allereerst een begripsbepaling van drie centrale begrippen uit dit onderzoek:

Content ontsluitings dienst: dit zijn de hiervoor genoemde toegangspoorten van het internet, de plekken waar de gebruikers van internet starten om content te zoeken. Dit zijn (zeer) bekende websites waarmee gezocht kan worden naar content. Verder in dit onderzoek genoemd ‘diensten’.

Onsluitingsmethoden: de wijzen waarop door de verschillende diensten de content wordt ontsloten. Diensten maken doorgaans gebruik van één methode om content te ontsluiten. Verder in dit onderzoek genoemd ‘methoden’.

Zoekdoel: De doelen die een internetgebruiker probeert te bereiken met een zoekopdracht, verderop gecategoriseerd in 3 (hoofd)doelen en 9 subdoelen.

De hierboven geïntroduceerde diensten hebben hun eigen, unieke manier van het zoeken en aanbieden van informatie aan hun gebruikers. Om er voor te zorgen dat de diensten die voor het onderzoek gebruikt worden ook daadwerkelijk het ontsluiten van informatie als kerntaak hebben, zijn er een aantal randvoorwaarden opgesteld waaraan de diensten moeten voldoen. Deze exacte voorwaarden zijn terug te vinden in het hoofdstuk *Methodologische verantwoording* (H3).

Wanneer alle typen websites die er op internet zijn (Put, 2008) gefilterd worden aan de hand van deze randvoorwaarden blijven er vijf typen websites over, namelijk zoekmachines, weblogs, webdirectories, social bookmarking sites en social news sites. Deze laatste twee typen diensten vertonen een grote overlap: beiden typen zijn volledig samengesteld door gebruikers met als doelstelling het op een sociale wijze uitwisselen van content, of dit nu links zijn, nieuws of andere zaken. Voor dit onderzoek zijn deze twee typen hierom samengevoegd onder de paraplu *folksonomieën*. Op deze wijze zijn er voor dit onderzoek vier categorieën tot stand gekomen.

De eerste categorie is die van de **zoekmachines**, ook wel *query search* genaamd.

Diensten als Google en Live! Search (Microsoft) doorzoeken door middel van een speciaal computerprogramma het internet, en maken van alle sites die software tegenkomt automatisch een indexering aan de hand van korte doorzoekbare samenvattingen. De relevantie van sites binnen deze indexering wordt grotendeels bepaald door het aantal links op het internet dat naar een site verwijst.

De tweede categorie is die van **folksonomieën**.

Sites als Delicious en Digg pakken het organiseren en vindbaar maken van content op een geheel eigen wijze aan. Diensten die met deze methode werken maken gebruik van de aanbevelingen en/of de bookmarks van internetgebruikers om content in hun indexering op te nemen. Ook wordt de content aan de hand van deze internetgebruikers georganiseerd, bijvoorbeeld door het rangschikken ervan aan de hand van voorkomendheid of te ordenen naar omschrijvingen die gebruikers geven over de content.

De derde categorie zijn **web directories**.

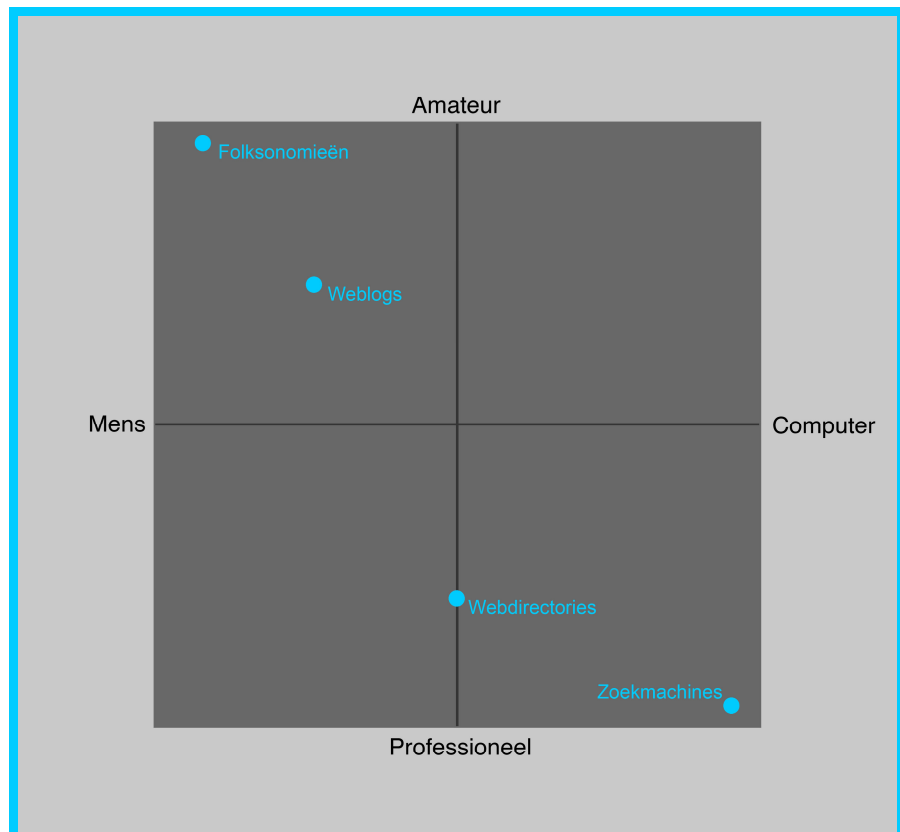
Dit zijn in feite digitale catalogi waar gebruikers door gecategoriseerde links kunnen zoeken, vergelijkbaar met bijvoorbeeld de Gouden Gids. Voorbeelden hiervan zijn het Open Directory Project en Startpagina. De eerste wordt automatisch gegenereerd aan door middel van software, de tweede wordt handmatig samengesteld door personen.

Als vierde categorie zijn er de **weblogs** zoals Nu.nl en Spitsnieuws. De (grotere) diensten van deze categorie hebben een aantal personen in dienst die voor de site content creëren en selecteren de zogenaamde *bloggers*, die de content op meestal chronologische volgorde op de website van de dienst plaatsen.

Soms zijn dit betaalde medewerkers, vaak ook zijn dit mensen die dit vrijwillig doen.

Deze diensten en de methoden waarmee ze werken zijn in het hoofdstuk Dienstenanalyse (hoofdstuk 4) verder uitgediept.

Er wordt constant nieuwe technologie ontwikkeld voor toepassingen en content op internet. Nieuwe manieren worden gebruikt om de omvang van data te verkleinen, de weergave ervan te verbeteren, downloadsnelheden te optimaliseren, nieuwe programmeermethoden



Figuur 1: Schematische weergave werkwijze zoekmethoden

voor websites te ontwikkelen, enzovoorts. Ook worden er continu nieuwe mogelijkheden ontwikkeld waarmee geprobeerd wordt dat de online content vindbaar is, wordt of blijft. Hierdoor verkeert internet in een constante staat van ontwikkeling. Dit wordt veelal aangeduid met de term *bèta* of *bètaversie*. Dit is niet vreemd als er rekening gehouden wordt met het feit dat Yahoo! met een bestaansduur van slechts 12 jaar één van de oudste internetbedrijven is. Het is voor diensten en hun werkwijze noodzakelijk om in te spelen en mee te groeien met de veranderingen die er plaats vinden. Het tempo van deze ontwikkelingen is echter zo hoog dat de verschillende diensten problemen hebben om deze bij te houden.

Het is voor deze diensten vrijwel niet haalbaar meer om met één en dezelfde methode al deze nieuwe (soorten) content te aggregeren. Aan de hand van enkele voorbeelden zal duidelijk worden dat het vaak niet haalbaar is om met alle veranderingen mee te groeien en de diensten hierom zich vaak (gedwongen) richten op één methode, die zij dan proberen zo ver mogelijk uit te diepen.

1.4 Problematiek

De verschillende methoden van informatieontsluiting zijn niet waterdicht.

Eén van de grote problemen die bijvoorbeeld regelmatig naar voren komt, is dat de methode waarop een dienst informatie verwerkt verbonden is met het soort content dat men ermee kan vinden. Sommige problemen zijn hierdoor gekoppeld aan een specifieke methode, maar andere problemen gaan voor alle methoden op. Ter illustratie een aantal (willekeurige) voorbeelden:

Een zoekmachine kan een gebruiker slechts aan de hand van steekwoorden en indexeringen vertellen wat en waar op internet bepaalde onderwerpen te vinden zijn. Deze zoekmachines kunnen echter geen inschatting maken van het waarheidsgehalte van deze tags, de (artistieke) kwaliteit van een kunstwerk of de emotionele lading van afbeeldingen of video's. Gebruikers van de dienst Delicious kunnen dit soort informatie wel verwerken. Zij kunnen dit door middel van deze dienst aanbevelen aan hun medegebruikers. Maar deze gebruikers zijn op hun beurt weer niet opgewassen tegen de snelheid en nauwkeurigheid waarmee de servers van Google databases of bestanden doorzoeken en indexeren.

Een ander probleem waar een groot deel van de diensten tegenaan loopt, is polysemie. Dit is het verschijnsel dat woorden verschillende dingen kunnen betekenen, afhankelijk van de context. Op het moment dat deze context ontbreekt, iets wat veelvuldig voorkomt bij tags, ontstaan er onduidelijkheden. Wordt er met de zoekterm *bank* nu een financieel instituut bedoeld, een zitplaats of een ondiepte van zand in het water gezocht?

De groep die de gevolgen van deze problematiek ondervindt zijn de internetgebruikers. Als de diensten, met al hun middelen, de ontwikkelingen al niet bij kunnen houden, hoe is het dan gesteld met de gebruikers?

De eerder in deze inleiding genoemde methoden, de zoekmachines, folksonomieën, weblogs en webdirectories, zijn manieren die zijn ontwikkeld om toegang tot de schat aan online informatie te verschaffen.

Hoe succesvol zijn de webdiensten in het toepassen van de door hun gebruikte methoden en hoe wordt deze door gebruikers ervaren? Bieden de diensten voldoende ondersteuning aan gebruikers of bevindt de gebruiker zich op een hellend vlak waardoor de gebruikers ten onder zullen gaan aan een information overload?

Deze thesis is een exploratief onderzoek naar de verschillende diensten die er momenteel voor internetgebruikers beschikbaar zijn om informatie op internet te ontsluiten en de methoden die deze diensten hiervoor gebruiken.

Terwijl het grootste deel van de Nederlanders (bijna 94 procent (Wierenga, 2008)) Google als dé manier beschouwt om op zoek te gaan naar content, zijn er veel meer diensten die de internetgebruiker proberen te ondersteunen in het behalen van hun zoekdoelen op internet.

Dit onderzoek zal zich verdiepen in welke diensten op dit gebied toonaangevend zijn, op welke manier deze te categoriseren zijn en hoe deze diensten achter de schermen werken. Ook zal het onderzoek inzicht verschaffen in de werkwijze en beleving van gebruikers met betrekking tot het zoeken en vinden van informatiebronnen op internet.

Welke diensten kennen gebruikers, welke gebruiken ze nog meer en gebruiken ze een andere dienst afhankelijk van het soort doel dat ze proberen te bereiken? En hoe zit het met de tevredenheid van gebruikers over de werking van de diensten?

1.5 Doelstelling

Het eerste doel van dit onderzoek is het verschaffen van een overzicht van de toonaangevende diensten en de methoden die deze diensten gebruiken om informatie op internet te ontsluiten. Daarbij zal er een beschouwing gegeven worden over de voor- en nadelen van de verschillende methoden die door deze diensten gehanteerd worden.

Het tweede doel is inzicht geven in welke van deze diensten en methoden gebruikers hanteren om zo hun (zoek)doelen op internet te verwezenlijken.

Als derde doel zal er met dit onderzoek antwoord gegeven worden op de vraag hoe het zoeken naar informatie op internet door gebruikers wordt ervaren.

Ook is er voor de internetgebruiker nog een andere rol weggelegd dan die van onderzoeksobject, namelijk als doelgroep. Deze masterthesis is namelijk ook geschreven met als insteek om internetgebruikers transparantie en ondersteuning te verschaffen over welke methoden er voor hun beschikbaar zijn op internet om informatie te ontsluiten.

1.6 Hoofd en Deelvragen

1.6.1 Hoofdvraag:

Welke methoden worden door de toonaangevende diensten online aangeboden aan gebruikers om content op internet te ontsluiten, wat zijn de voor- en nadelen van deze methoden en diensten en met welk zoekdoel worden de verschillende methoden door gebruikers gehanteerd en ervaren?

1.6.2 Deelvragen

- Welke type methoden bieden toonaangevende diensten aan op internet voor het ontsluiten van content?
- Wat zijn de voor- en nadelen van deze methoden en met welk zoekdoel worden verschillende methoden door gebruikers gehanteerd?
- Hoe wordt het zoeken naar informatie op internet door gebruikers ondervonden?

1.7 Relevantie

Zoals in de inleiding naar voren is gebracht, voeren verschillende schrijvers, experts en onderzoekers een discussie over de voor en nadelen, maar ook over de kansen en bedreigingen die internet met zich meebrengt. Een deel van dit debat bestaat echter uit speculaties en aannames. Dit onderzoek tracht daarom bij te dragen aan de nuancering en feitelijke ondersteuning van deze discussie.

Vanuit een wetenschappelijk oogpunt levert dit onderzoek een bijdrage aan de kennis over wijze waarop internetgebruikers diensten en methoden van informatieontsluiting toepassen. Ook verschaft het inzichten in de opvattingen en meningen van internetgebruikers over de ontsluiting van informatie op internet. Dit doet het aan de hand van een gebruikersperspectief.

1.8 Leeswijzer

In hoofdstuk 1 (*Inleiding*) is de centrale problematiek rondom dit onderzoek geschetst. Ook zijn hierin de hoofd en deelvragen waarop dit onderzoek poogt antwoord te geven opgesteld.

Hoofdstuk 2 (*Conceptueel raamwerk*) van deze thesis bevat de theoretische en praktische kennis welke het raamwerk vormt voor het onderzoek.

In hoofdstuk 3 (*Methodologische verantwoording*) is een verklaring en verantwoording van de methodologische aanpak van het onderzoek. Hierin staat de werkwijze beschreven waarop de dienstenanalyse (H4), het gebruikersonderzoek (H5) en de verdieping van de onderzoeksresultaten (H6) tot stand zijn gekomen.

Hoofdstuk 4 (*Dienstenanalyse*) bevat de gevonden resultaten van de dienstenanalyse. In dit hoofdstuk zijn de eerste twee deelvragen beantwoord.

Hoofdstuk 5 (*Gebruikersanalyse*) van deze thesis bevat de resultaten van het verrichte gebruikersonderzoek. Dit hoofdstuk bevat ook het antwoord op de laatste deelvraag.

In hoofdstuk 6 (*Verdieping onderzoeksresultaten*) zijn de resultaten van het de dienstenanalyse en het gebruikersonderzoek met elkaar vergeleken om zo een verdiepende analyse te geven over de antwoorden op de deelvragen.

Als afsluiting bevat hoofdstuk 7 (*Conclusie*) een samenvatting en conclusie over de bevindingen, een terugblik op het verloop van het onderzoek en enkele aanbevelingen voor vervolgonderzoek.

2. Conceptueel raamwerk

In dit hoofdstuk worden verschillende theorieën behandeld met internet als uitgangspunt en wordt de benodigde beschikbare achtergrondkennis toegelicht en uitgediept.

In het hoofdstuk worden de oorsprong en de bouwstenen van het internet beschreven.

Daarnaast zullen enkele van de veranderingen die het internet teweegbrengt worden besproken. In het hoofdstuk zullen ook de verschillende wijzen waarop de informatie op internet gestructureerd kan worden uitgediept. Als laatste zal er een beschouwing gemaakt worden van de zoekdoelen die internetgebruikers kunnen hebben.

2.1 Het internet

Het internet is oorspronkelijk afkomstig uit een militair project aan het einde van de jaren zestig. Het werd ontwikkeld onder de naam ARPANET en was in eerste instantie bedoeld als methode om vertrouwelijke informatie uit te wisselen zonder dat deze afgetapt kon worden.

De potentie van dit

netwerk als

communicatiemiddel

werd al snel door velen

gezien en zowel

onderzoekers als

particulieren begonnen

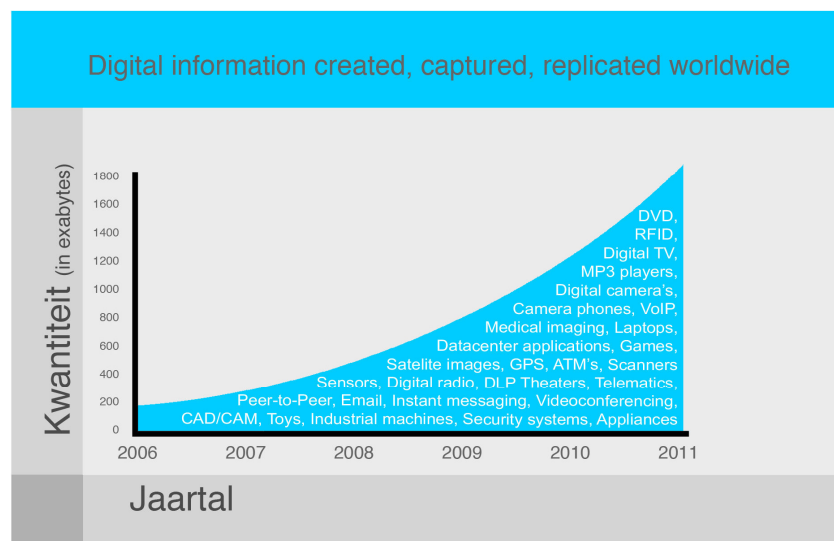
te experimenteren met

vergelijkbare

netwerken. Pas in 1991

ontstond internet in zijn

huidige vorm door de



Figuur 2: Hoeveelheid geproduceerde digitale informatie in exabytes¹

¹ Figuur gebaseerd op Gantz, J. F., Chute, C., Manfrediz, A., Minton, S., Reinsel, D., Schlichting, W., et al. (2008). The Diverse and Exploding Digital Universe. EMC2, P3. <http://www.emc.com/collateral/analyst-reports/diverse-exploding-digital-universe.pdf>

ontwikkeling van HTML: een computertaal die het mogelijk maakt om internetpagina's te ontwerpen en te schrijven en om informatie over het net te transporteren. (Bellis, 1997) In 1993 waren er nog maar 130 websites op heel het internet te vinden (Gray, 1995). Door de technologische ontwikkelingen, zoals HTML, nam de groei vanaf dit punt een vlucht en verdubbelde de omvang van internet elke drie maanden. Slechts 2,5 jaar later waren er 100.000 websites te bezoeken en anno 2008 waren dit er al vele miljarden². De hoeveelheden informatie waar over gesproken wordt zijn haast niet voor te stellen, maar om een indicatie te geven van de omvang van deze informatie die geproduceerd wordt: in 2002 steeg de productie van *nieuwe* informatie naar 5 exabytes³, waarvan 92 procent digitaal opgeslagen werd. Deze hoeveelheid staat gelijk aan de optelsom van *alle* woorden die ooit door de mens gesproken zijn. In 2007 was de toename al 281 exabytes, ofwel gemiddeld 45 gigabyte voor elke persoon op aarde. Geschat wordt dat de hoeveelheid informatie die toegevoegd wordt aan de digitale wereld tussen 2006 en 2010 zal verzesvoudigen, van 161 naar 988 exabytes. (Gantz et al., 2008)

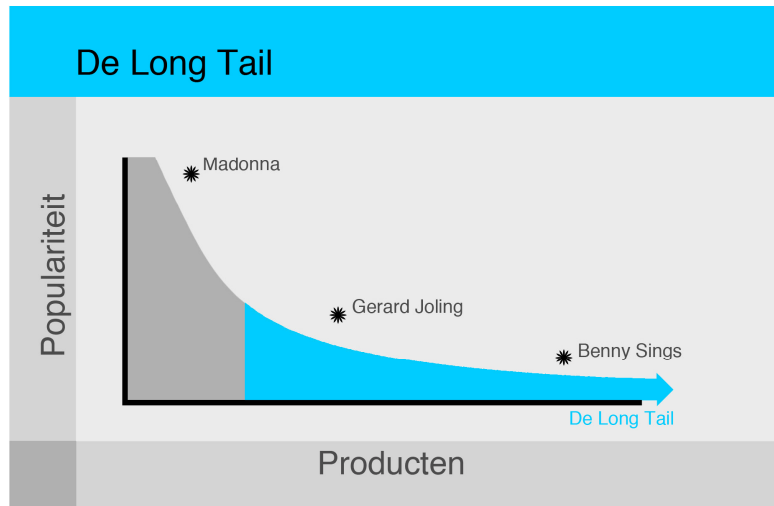
2.2 De Long Tail

Deze toename aan informatie op internet heeft de zogeheten Long Tail tot gevolg. De Long Tail is een concept dat door Chris Anderson gepopulariseerd werd in 2004. In zijn boek (2006) beschrijft hij een fenomeen waarin een grote groep producten waar weinig vraag naar is collectief net zo veel, of zelfs meer, marktaandeel hebben dan een relatief kleine groep producten waar veel vraag naar is. Met de komst van internet, en het wegvallen van de benodigde en kostbare schapruimte, is het mogelijk om deze markt van producten met lage verkoopaantallen optimaal te benutten. Hierdoor is het erg interessant geworden te produceren voor de "markt van de lage verkoopaantallen".

² Schatting n.a.v. Worldwebsitesize.com, geraadpleegd 1 december 2008.

³ Een exabyte is 1,000,000,000,000,000,000 bytes, ofwel een miljard gigabytes

Een voorbeeld hiervan is de iTunes store; een online muziekwinkel van Apple waar muziek alleen digitaal aangeboden wordt. De iTunes store biedt gebruikers momenteel een enorme hoeveelheid titels aan



om te downloaden, een **Figuur 3: The Long Tail⁴** kwantiteit die voor reguliere muziekwinkels niet haalbaar is vanwege de enorme hoeveelheid ruimte die dit in zou nemen. Dit zou erg kostbaar zijn. Ter indicatie: er zijn volgens de laatst bekende cijfers uit 2008 al meer dan vijf miljard nummers verkocht via de iTunes Store⁵.

Het mes snijdt aan twee kanten. Het uitbrengen van muziek door artiesten is een kostbare aangelegenheid door onder andere de perskosten en de distributie van cd's. Deze hoge investeringen hadden als gevolg dat alleen die muziek werd uitgebracht waaraan deze investeringen het waarschijnlijkst konden worden terugverdiend. Door muziek via internet aan te bieden, valt een groot deel van deze kosten weg. Hierdoor kan er nu een veel breder aanbod aan muziek via internet (rendabel) worden uitgebracht.

De Long Tail geldt echter niet alleen voor producten. Het fenomeen heeft ook betrekking op dit onderzoek. Vanwege het wegvallen van de fysieke barrière van content is het mogelijk om internetgebruikers een veel uitgebreider spectrum aan content aan te bieden tegen een fractie van de kosten.

Door de toename én specialisatie van content kunnen gebruikers uit een veel breder aanbod hun selectie maken, en hierdoor het totaalpakket aan content veel strakker laten aansluiten bij hun individuele wensen.

⁴ Figuur gebaseerd op http://www.novelfr.com/wp-content/uploads/2008/02/long_tail_graph.gif

⁵ iTunes Store tops over 5 billion songs sold. <http://www.apple.com>. Geraadpleegd 10 februari 2009

De diversificatie in content, die voor een deel door de Long Tail veroorzaakt wordt, heeft ook een ander gevolg. Er ontstaat ook een steeds breder assortiment aan contenttypen, metadata en locaties waar content te vinden is. Hierdoor wordt het organiseren en vindbaar maken van content ook steeds complexer.

Om een indicatie te geven van de omvang van de Long Tail: in theorie zou dit betekenen dat wanneer er bij alle zoekmachines gezocht wordt met de duizend meest populaire zoektermen en de resultaten hiervan opgeteld zou worden, er nog steeds 89,4 procent van de antwoorden niet gevonden zou worden. (Tancer, 2008) Dit geeft aan dat in deze staart klaarblijkelijk een enorme hoeveelheid, potentieel waardevolle, content verborgen kan zitten.

Niet alleen de hoeveelheid informatie op internet neemt snel toe. Ook het aantal gebruikers van internet blijft vooralsnog sterk groeien, de snelheden om informatie te up- en downloaden worden steeds hoger en de intensiviteit van het gebruik van internet neemt toe. (CBS, 2007)

In het gebruik van internet behoort Nederland tot de absolute wereldtop: 88 procent van de Nederlanders heeft een pc, 86 procent heeft een internetaansluiting waarvan 74 procent een breedbandverbinding. (CBS, 2008)

Zoals in de inleiding reeds beschreven is, worden veel van deze gebruikers in toenemende mate in staat gesteld een bijdrage te leveren aan het internet, bijvoorbeeld in zogenaamde Web2.0 diensten. (O'Reilly, 2005) Dit wordt ook wel Social Computing⁶ genoemd.

Web2.0 kan gezien worden als de tweede ontwikkelingsfase van het Web, de opvolger van de eenzijdige en statische eerste fase van het internet. Deze eerste fase van het internet kwam in 2001 metaforisch op haar einde door het knappen van de zogenaamde "dotcom-bubble". (Slot & Frissen, 2007:201)

Het belangrijkste uitgangspunt van de term web 2.0 is dat (groepen) gebruikers een actieve bijdrage leveren aan het creëren van toegevoegde waarde door bijvoorbeeld zelf content te

⁶ Met de term social computing wordt er gerefereerd aan de sociale wijze waarop het web en haar content tot stand komen.

creëren. Door deze toegenomen activiteiten, is ook het aanbod van content gemaakt door gebruikers – user created content – sterk gestegen.

Voordat er verder gegaan wordt met het uiteenzetten van de verschillende opvattingen, theorieën en onderzoeken over het onderwerp, zal er eerst een stap terug gedaan worden naar de basis, namelijk het definiëren van wat nu precies datgene is wat er op internet gevonden kan worden.

2.3 Wat zijn data, informatie en content?

In dit onderzoek worden de termen *data*, *informatie* en *content* veel gebruikt. Op het eerste gezicht lijken deze begrippen veel op elkaar, toch is er een duidelijk onderscheid dat gemaakt moet worden tussen deze begrippen.

2.3.1 Data

De eerste in deze rij is het begrip *data*.

De definitie van dit woord is erg omstreden, vanwege de verschillende gebruiksvormen.

Wetenschappelijk gezien wordt data beschouwd als opgeslagen gegevens in hun ruwste vorm. Deze gegevens kunnen veel verschijningsvormen aannemen en er is geen enkele relatie tussen de data en andere zaken. (Bellinger, Castro & Mills, 2004) Door het ontbreken van een context heeft data geen significantie buiten dat van het eigen bestaan. (Ackoff 1989:3-9) Deze vrij abstracte vorm van het begrip data heeft een zeer omvattende inhoud.

In de praktijk wordt de term data meestal gebruikt wanneer er gesproken wordt van vastgelegde gegevens, hoewel dit niet volledig nauwkeurig is.

Wanneer over data gesproken wordt in technologische context heeft men het over digitaal opgeslagen gegevens, waarmee dus alle content op internet onder deze noemer valt.

2.3.2 Informatie

Van *informatie* wordt gesproken op het moment dat er een betekenisvol verband gelegd kan worden tussen data en andere zaken.

Dit betekent meestal dat de data verwerkt is in een verschijningsvorm die voor mensen begrijpbaar is in combinatie met een ander significant verband. Als voorbeeld kunnen we een grote lijst met een willekeurige selectie van Nederlandse woorden nemen. Volgens de definitie is deze lijst informatie, aangezien de vorm waarin de data is vastgelegd een taal is en er dus een verband bestaat. In de praktijk heeft deze lijst van woorden weinig functie en wordt deze pas interessant als deze voor de ontvanger in context wordt geplaatst. Dit gebeurt dan door nog een verband, zoals bijvoorbeeld de kennis dat de woorden die in de lijst staan oorspronkelijk afgeleid zijn uit een andere taal.

In dit onderzoek zal het begrip informatie op een praktische wijze worden gebruikt, namelijk als gegevens die begrijpbaar zijn en potentiële waarde voor de ontvanger kunnen bevatten.

2.3.3 Content

Het derde begrip is *content*.

Content is simpel gezegd de technische inhoud van een webpagina. (Microsoft, 2002) Dit kan zijn in de vorm van tekst, beeld, video, geluid en software. (IANA, 2007) Het format, vormgeving en structuur van de website worden niet meegerekend.

Data, informatie en content komen op internet erg dicht bij elkaar. Content op internet is altijd data, aangezien het opgeslagen ligt op een (digitaal) medium, maar het hoeft dus niet altijd informatie te zijn.

Aangezien dit onderzoek zich voornamelijk richt op het gebruik en de perceptie van internetgebruikers, zal er vanuit worden gegaan dat deze op zoek is naar (digitale) data op internet, oftewel content, die voor hem of haar begrijpbaar is en een bepaalde informationele waarde heeft.

Content op internet kan, zoals hierboven al kort beschreven werd, verdeeld worden in verschillende verschijningsvormen. De mogelijkheden zijn tekst, afbeeldingen, video, audio

en software. Deze content kan door verschillende partijen op internet beschikbaar worden gemaakt, zoals door overheden, bedrijven en gebruikers.

Bij overheden en bedrijven wordt de content vanuit een professioneel kader gecreëerd. Naast deze professionele content is er ook nog de content die door internetgebruikers is gemaakt, ofwel User Created Content (UCC). Deze wordt ook wel User Generated Content (UGC) genoemd.

Het is moeilijk om te definiëren wat User Created Content nu precies is en aan welke eisen het voldoet. Er zijn door de Organisatie voor Economische Samenwerking en Ontwikkeling (OECD)⁷ een aantal karakteristieken opgesteld om UCC te duiden (Vickery, Wunsch-Vincent, 2007:18).

Als eerste karakteristiek stelt het OECD dat de content op één of andere wijze online gepubliceerd moet worden. Hier wordt onder verstaan dat het werk online toegankelijk is voor derden.

Als tweede karakteristieke eigenschap stelt het OECD dat er een bepaalde creatieve inzet moet zijn geweest om het werk te creëren of bestaande werken aan te passen. Hieronder moet creatieve inzet erg breed verstaan worden; bijvoorbeeld toevoegen van tekst, bewerken van een afbeelding, omzetten naar een andere format etc. Gesteld kan worden dat op enige wijze door de gebruiker waarde toegevoegd moet worden aan de content.

De derde karakteristiek is dat de geproduceerde content voornamelijk buiten het professionele circuit wordt geproduceerd. Over het algemeen zit er geen commerciële of institutionele context aan verbonden en wordt het gecreëerd met één of meer van de volgende doelstellingen: het contact leggen met soortgelijken (*peers*), het verkrijgen van bekendheid of prestige en het zichzelf willen uiten. (Frissen, 2008:26)

Nu is voor dit onderzoek begrip over de herkomst en productie van de content op eerste gezicht niet essentieel. Het feit dat deze content nu eenmaal aanwezig is leek in eerste instantie dan ook voldoende als fundament voor het onderzoek. Gedurende het onderzoek kwam echter naar voren dat de unieke problematiek van het vinden en vindbaar maken van content deels het gevolg is van de verschillende bronnen waaruit deze content afkomstig is.

⁷ Het OECD is een samenwerking van 30 landen opgericht ter ondersteuning van het Marshallplan, welke statistische, economische en sociale data verzameld. Het doel hiervan het is het monitoren, voorspellen en onderzoeken van patronen in handel, milieu, technologie, economie en meer. (OECD.COM)

2.4 Metadata

Om orde te scheppen in de grote mengelmoe van content wordt op internet meestal gewerkt met *metadata*.

Metadata is gestructureerde informatie die het terugvinden, beheren of gebruiken van een informatiebron (content) vergemakkelijkt aan de hand van een beschrijving, toelichting of plaatsbepaling. (Hodge, 2001:3) Vaak wordt metadata wordt ook wel data over data genoemd. (Mathes, 2004:1)

Metadata beperkt zich zeker niet tot het digitale domein; de indexering van boeken in een bibliotheek is ook metadata. De digitale variant vindt hierin zelfs zijn oorsprong.

Metadata kan op twee verschillende wijze toegepast worden.

Enerzijds kan metadata informatie bevatten over de bron zelf, bijvoorbeeld het bestandstype of de omvang ervan. Denk hierbij bijvoorbeeld aan de informatie over een bestand zoals deze in Windows onder de rechtermuisknop/ eigenschappen gevonden kan worden.

Anderzijds kan metadata gebruikt worden om de inhoud van een bestand te beschrijven, bijvoorbeeld dat het humoristisch is of gaat over geschiedenis. Deze laatste toepassing van metadata wordt voornamelijk gedaan aan de hand van zogeheten *tags* of *keywords*. Het komt er in de praktijk meestal op neer dat er aan de content enkele steekwoorden worden verbonden, die de hoofdzaken van de content samenvatten.

Data en metadata kunnen soms door elkaar heen lopen. Zo kan de titel van een document de inhoud beschrijven (metadata), maar tegelijkertijd is het ook onderdeel van het document (data).

2.5 De methoden

2.5.1 Introductie

De basis van het classificeren, organiseren en doorzoeken van informatie is afkomstig uit de wetenschap van de taxonomie.

Deze wetenschap legt zich er op toe manieren te ontwikkelen voor het ordenen van allerlei zaken. Voorbeelden hiervan zijn bijvoorbeeld diersoorten of literatuur. Van oorsprong gebeurde deze indelingen veelal aan de hand van typen en subtypen, die worden opgesteld aan de hand van specifieke voorwaarden. Op deze wijze wordt er een (boom)structuur gevormd die algemeen begint en stap voor stap specifiek wordt.

Een voorbeeld hiervan is: Dieren > Zoogdieren > Katachtigen > Panthera > Luipaard.

Het inrichten van een dergelijk oorspronkelijk systeem vereist veel (voor)kennis en is hierdoor dan ook vaak voorbehouden aan professionals. Dit soort systemen worden hierdoor dan ook vaak centraal gecontroleerd door een selecte groep betrokkenen.

Wanneer een dergelijk systeem goed is opgezet, heeft dit als grote voordelen dat het zeer betrouwbaar, overzichtelijk en accuraat is. Ook is het door de eenduidigheid erg voorspelbaar waar de gezochte informatie te vinden is. Toch kleven er ook een aantal nadelen aan het op deze wijze organiseren van informatie. Het systeem wordt over het algemeen beheerd door professionals en het toevoegen van nieuwe content is, door de nauwkeurigheid waarmee het moet gebeuren, relatief arbeidsintensief. Dit maakt een dergelijk systeem erg kostbaar. Ook vraagt het enige kennis bij de gebruiker om iets te kunnen vinden, kennis die niet altijd aanwezig is.

Een probleem wat zich vooral op internet erg duidelijk laat zien is de rigiditeit van deze oorspronkelijke vorm van informatie organiseren. De constante veranderingen, ontwikkelingen en vernieuwing van onderwerpen en content op internet zorgen ervoor dat bestaande randvoorwaarden en criteria van het systeem snel achterhaald raken.

Er zijn veel partijen, waaronder de diensten die in dit onderzoek behandeld worden, die zich bezig houden met de taxonomische problematiek.

De verschillende partijen gebruiken daarvoor een verschillende aanpak.

Hierdoor zijn er binnen de wetenschap van de taxonomie, en nog specifiek, taxonomie van content op internet, nieuwe en aangepaste werkwijzen ontstaan om de informatie te ordenen en classificeren.

Binnen dit onderzoek zijn de volgende vier methoden van informatieontsluiting op internet onderzocht/gebruikt: de zoekmachine-, folksonomische-, de webdirectory- en de weblogmethode.

2.5.2 Zoekmachines

De bekendste manier om op internet informatie te zoeken is door middel van een zoekmachine zoals Google, Yahoo of Live!Search.

Deze zoekmachines hebben speciale software, web crawlers of web spiders geheten, die continu het web doorzoeken en een korte samenvatting maken van alle websites die ze tegenkomen. Deze worden opgeslagen in een index. Hoe deze manier van indexeren en relevantiebepaling precies wordt samengesteld, is een zeer complex systeem en werkt met een groot aantal variabelen. Voor dit onderzoek volstaat kennis van de basis van dit proces.

De manier van doorzoeken verschilt per zoekmachine, maar in grote lijnen werkt dit systeem als volgt: de software van de zoekdienst scant internet af, maakt een samenvatting van de site en creëert hierover metadata die wordt opgeslagen in een index van de zoekmachine. Deze metadata wordt voornamelijk samengesteld uit het onderwerp en de kopteksten van de website, uit combinaties van woorden binnen de tekst en uit door de makers zelf meegegeven metadata.

Ook wordt er een weging meegegeven aan de verwachte kwaliteit van de content. Dit wordt meestal bepaald aan de hand van de rangschikking van een site. Wanneer een site veel bezoekers trekt en er op veel andere (veelbezochte) sites links worden geplaatst naar de site, stijgt deze in rang. Andere zaken die meewegen in de rangschikking van een site zijn onder andere de bestaansduur van de site en de regelmaat waarmee deze aangepast wordt. Wanneer er gezocht wordt met een zoekmachine krijgt de gebruiker een respons die volgens de dienst het meest passend is voor zijn zoekopdracht gebaseerd op de index in combinatie met de relevantie-weging.

2.5.3 Folksonomieën

Met de komst van Web2.0 is er ook een nieuwe vorm voor het organiseren van content ontstaan, namelijk de *folksonomie*. (Vanderwal, 2005)

Dit is ook wel bekend onder de naam *collaborative tagging*. Waar bij een oorspronkelijke taxonomie het toekennen van tags of andere metadata oorspronkelijk top-down gedaan wordt door (veelal) professionals, wordt dit bij een folksonomie juist van onderaf (bottom-up) gedaan door de makers van de content of de internetgebruikers. In de praktijk is de basis van een folksonomie vaak een online gemeenschap van internetgebruikers die vanuit een sociaal oogpunt links en content met elkaar delen.

Dit systeem functioneert optimaal wanneer het relatief eenvoudig is voor de gebruikers om tags toe te voegen aan content. Meestal houdt dit in dat er weinig of zelfs geen inhoudelijk eisen worden gesteld aan de metadata die toegevoegd kunnen worden aan content. Er ontstaat hierdoor een grote, hiërarchieloze verzameling steekwoorden. Deze verzameling kan vervolgens doorzocht kan worden. Dit kan bijvoorbeeld door middel van een zoekvenster of een wordcloud. Relevantie van bepaalde content wordt veelal bepaald aan de hand van de voorkomendheid van bepaalde content of steekwoorden in de database.

2.5.4 Webdirectories

Een webdirectory is een andere zoekmethode variant waar internetgebruikers hun zoektocht naar de gewenste content kunnen starten.

Webdirectories zijn gespecialiseerd in het verzamelen van en organiseren van links naar verschillende internetsites en/of content aan de hand van de oorspronkelijke taxonomische wijze, namelijk per categorie en eventueel verder in een subcategorie etcetera.

Deze categorische indeling, in plaats van bijvoorbeeld aan de hand van steekwoorden, is een essentiële randvoorwaarde voor een webdirectory.

De verschijning van een webdirectory is in praxis meestal een webpagina met een (geclusterde) lijst van links.

De links kunnen direct naar externe content verwijzen, maar veelal verwijzen ze naar een andere (sub)categorie met links. Op deze wijze ontstaat een boomstructuur.

Bij een webdirectory worden de links in sommige gevallen door middel van een beschrijving weergegeven in een lijst.

Deze webdirectories kunnen op twee verschillende manieren worden samengesteld, namelijk door mensen of door computers. Ook komt het regelmatig voor dat het een combinatie van de twee is, waarin de mens dan voornamelijk een controlerende rol vervuld. Veelgebruikte webdirectories zijn het Open Directory Project, Yahoo! Directory en Startpagina.

2.5.5 Weblogs

Een weblog is een dienst die bestaat uit een periodieke lijst met artikelen die over het algemeen in omgekeerde chronologische volgorde te zien zijn. (Technische Universiteit Delft, 2008) Dit houdt in dat de laatst toegevoegde informatie bovenaan aan de lijst met artikelen staat.

De lijst met artikelen wordt over het algemeen regelmatig bijgewerkt en combineert meestal tekst, afbeeldingen en verwijzingen naar andere websites en content. (WGBH, 2007)

Weblogs, waaronder degenen die in dit onderzoek zullen worden geanalyseerd, worden in de praktijk meestal door een groep mensen aangevuld en onderhouden.

Weblogs kunnen herkend worden aan de specifieke programmeerwijze van de site. Deze programmeerwijze scheidt de tekst en de opmaak van de tekst in twee verschillende delen, waardoor de content die op de site geplaatst wordt automatisch zichtbaar wordt binnen eerder opgestelde kaders van de site. Doorgaans gebeurt dit door een lange, dynamische rij artikelen in het midden van de pagina waar de nieuwste items bovenaan verschijnen.

Toonaangevende diensten in deze categorie zijn Nu.nl en Spitsnieuws.

2.5.6 Andere methoden

Door de constante dynamiek van internet en de ontwikkeling van nieuwe technologie en inzichten binnen de wetenschap van taxonomie worden er constant nieuwe concepten ontwikkeld voor het ontsluiten van informatie op internet. De vier methoden die hierboven zijn beschreven zijn hiervan de meest succesvolle.

Naast deze vier methoden bestaan nog een groot aantal andere methoden, waarvan sommige nog vroeg in het ontwikkelingsproces staan en andere niet succesvol bleken. Deze methoden zijn over het algemeen gericht op het doorzoekbaar en/of vindbaar maken van de inhoud van mediabestanden zoals audio, video en afbeeldingen, iets waar de meeste methoden nog geen sluitende oplossing voor hebben. In de praktijk zijn deze methode vaak varianten op de vier meest succesvolle methoden, of een hybride combinatie van enkele ervan.

Een interessante ontwikkeling op het gebied van het doorzoeken van audio en video zijn de mogelijkheden van spraakherkeningssoftware.

Een voorbeeld hiervan is Podscope⁸, een dienst die door middel van speciale software woorden uit audio- en videobestanden kan herkennen en deze omzet naar geschreven tekst. Deze tekst is vervolgens aan de hand van een zoekmachine te doorzoeken en af te luisteren. Op deze wijze is de dienst niet afhankelijk van metadata van om media te indexeren. Er wordt op deze wijze dus gezocht binnen de inhoud van de content, in plaats van in de omschrijving van de content.

Een dienst die momenteel hoge ogen gooit wanneer het om het inventariseren en doorzoekbaar maken van content gaat is Blinx⁹. Deze dienst heeft als hoofddoel video's inhoudelijk doorzoekbaar te maken. Dit doen ze aan de hand van een breed arsenaal aan software. De software is in staat om, onder andere, audio om te zetten in tekst, muziek te herkennen, afbeeldingen en gezichten uit video's te identificeren enzovoorts.

De dienst Tastekid¹⁰ probeert gebruikers te helpen met zoeken aan de hand van de smaak die ze hebben. Door hun favoriete band, film of boek in te typen krijgt de gebruiker door de dienst een tal aan aanbevelingen te zien over andere dingen die passen bij zijn of smaak. Deze aanbevelingen zijn een mix tussen waarderingen die op andere sites zijn gegeven, informatie van Wikipedia, filmpjes van Youtube enzovoorts.

Helaas zijn de bovengenoemde methoden nog verre van waterdicht. De complexiteit en diversiteit van de content zorgen nog voor veel struikelblokken. Problemen die deze diensten tegenkomen zijn bijvoorbeeld dat de kwaliteit van opnames lang niet altijd

⁸ <http://www.Podscope.com>

⁹ <http://www.Blinx.com>

¹⁰ <http://www.Tastekid.com>

optimaal is, er dialecten gesproken worden die (nog) niet herkend kunnen worden door computers of dat er nog een tekort is aan referentiedata.

Er zijn meerdere redenen waarom dit soort diensten en de door deze diensten gebruikte methoden niet zijn meegenomen in dit onderzoek.

De methoden hebben nog relatief lage gebruikersaantallen en zijn vaak nog in een te vroeg stadium van ontwikkeling om significant invloed te hebben op het zoeken en vinden van informatie op internet. Ook is, met Blix als uitzondering, de omvang van hun databases aan geïndexeerde content vaak verre van toereikend.

De voornaamste reden waarom deze diensten niet zijn meegenomen in het onderzoek is dat de methoden zich vrijwel allemaal toespitsen op slechts één of enkele contenttypen. De zoekdoelen die internetgebruikers proberen te behalen, wanneer ze op zoek zijn naar content, zijn niet gekoppeld aan contenttypen, maar aan de inhoud van de content.

Hierdoor is er voor deze diensten en methode binnen dit onderzoek (vooralsnog) geen plaats weggelegd.

2.6 Zoekdoelen

Ieder persoon die op internet op zoek gaat naar een bepaald soort informatie doet dit vanuit een uniek, persoonlijk motief.

De inhoud van de gezochte informatie is net zo divers is als de inhoud van internet zelf, maar het *doel* van deze zoekopdrachten is wel te categoriseren.

De doelen die een internetgebruiker probeert te bereiken met een zoekopdracht zijn onder te verdelen in drie verschillende categorieën. Deze zijn ontleend aan de onderzoeken van Broder en van Rose en Levinson. Deze zoekdoelen zijn het *navigationale*, *informationele* en het *transactionele* zoekdoel. (Broder, 2002:5) Deze laatste wordt door Rose en Levinson ook wel het *resource* zoekdoel genoemd. (Rose, Levinson, 2004:14)

2.6.1 Navigationaleel

Als eerste is er het navigationale zoekdoel. Dit zoekdoel wordt gedefinieerd door de wens van de gebruiker om een bepaald webadres of site (terug) te vinden en wordt ook wel *homepage finding task* genoemd. Belangrijk bij dit zoekdoel is dat de gebruiker de gezochte website van te voren reeds in gedachte heeft. Er wordt gezocht naar een concreet webadres, ook wel een URL genoemd (Uniform Resource Locator), omdat de gebruiker in de veronderstelling is dat de gezochte informatie op dit webadres te vinden is. Dit komt in de praktijk vaak neer op de websites van bedrijven, instellingen, etcetera, daar waar er redelijkerwijs van aangenomen kan worden dat er een internetsite van bestaat. Geschat wordt dat tussen de 10 en 15 procent van wat gebruikers zoeken, binnen deze categorie valt. (Rose, Levinson 2004:18) Voorbeelden ervan zijn het zoeken naar de website van bedrijf Coca-Cola of de website om belastingaangifte te doen.

Aan de theorie over dit zoekdoel hoort echter nog een onderdeel te worden toegevoegd, namelijk het *terugvinden* van een eerder gevonden site. Het doel in deze is hetzelfde, namelijk het localiseren van een website, dit ondanks dat het feit dat deze reeds eerder bezocht is. In dit onderzoek is het navigationale zoekdoel hiervoor gesplitst in twee subdoelen, waarvan het oorspronkelijke navigationale zoekdoel (het vinden van een webadres) als subdoel *navigate* wordt aangeduid en het toegevoegde subdoel (het terugvinden van een webadres) als *retrieval*.

2.6.2 Informatieel

Het tweede zoekdoel dat een internetgebruiker kan hebben wanneer hij of zij internet bezoekt is het *informatieele* zoekdoel, ook wel de *topic relevance task*.

Dit zoekdoel kan omschreven worden als het gericht zijn op het verkrijgen van informatie over een gezocht onderwerp. Dit zoekdoel is erg breed en er is dan ook een onderverdeling in aan te brengen. Zo kan iemand gericht op zoek zijn naar het antwoord (directed) op een vraag of ongericht zoeken en meer willen weten over een onderwerp in zijn algemeenheid (undirected). Ook kan iemand op zoek zijn naar advies over een onderwerp (advice) of proberen iets te vinden in de echte wereld (locate). Als laatste kan de persoon op zoek zijn naar de resultaten van zijn zoekopdracht op zich (list). Aan dit laatste abstracte zoekdoel wordt in dit onderzoek geen aandacht geschonken.

Geschat wordt dat ongeveer 60 á 65 procent van de zoekopdrachten op internet gedaan worden met het informatiele zoekdoel voor ogen. (Rose, Levinson 2004:18) Voorbeelden van dit zoekdoel zijn het zoeken naar de historie van Coca-Cola, het adres van de belastingdienst in Rotterdam of advies over opvoeding.

2.6.3 Transactioneel

Het derde en laatste zoekdoel dat een gebruiker zich op internet kan stellen is het transactionele zoekdoel.

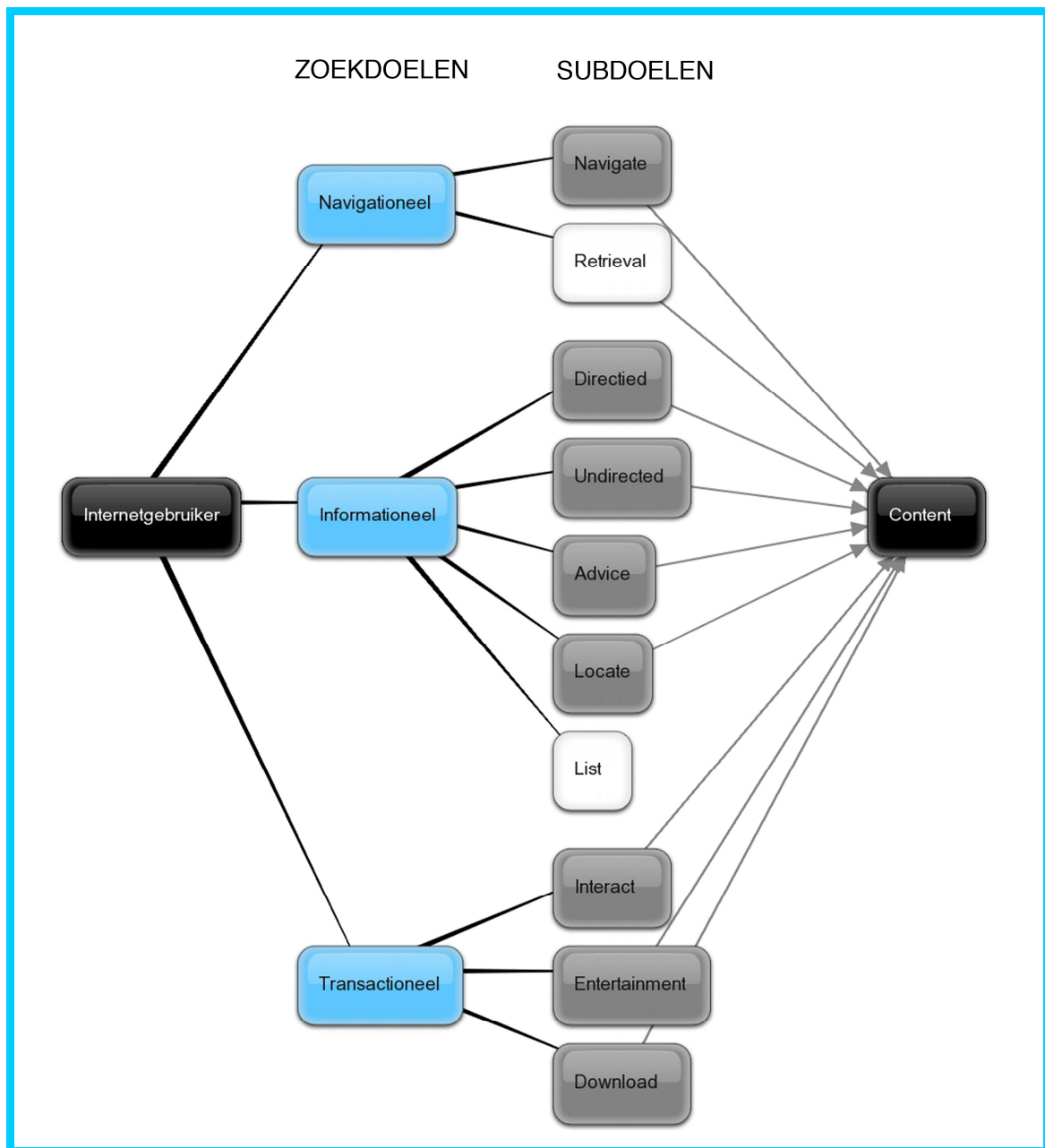
Als een gebruiker dit zoekdoel voor ogen heeft is hij naar iets anders op zoek dan informatie. Zo kan hij of zij op zoek zijn naar content om te downloaden naar de computer (download) zoals mp3-bestanden of software.

Ook beslaat het transactionele zoekdoel het zoeken naar zaken waarvan de online beleving het einddoel is (entertainment). Er kan hier altijd nog een andere handeling op volgen, nadat de content is bekeken of beluisterd, maar de insteek op het moment van zoeken was puur het ondervinden van de content.

Als laatste kan iemand met dit zoekdoel op zoek zijn naar een interactieve handeling waarvoor internet gebruikt wordt als platform (interact).

Voorbeelden van dit zoekdoel zijn bijvoorbeeld het zoeken naar een nieuwe screensaver van Coca-Cola om op de pc te installeren, erotische websites of online spellen. Ongeveer 20 á 25

procent van de zoekopdrachten op internet valt onder deze noemer. (Rose, Levinson 2004:18)



Figuur 4: Overzicht van de potentiële zoekdoelen en subdoelen van een zoekopdracht

Deze beschreven categorieën zullen in dit onderzoek gebruikt worden om te onderzoeken welke zoekdoelen gebruikers voor ogen hebben, wanneer zij een bepaalde methode selecteren om online content te vinden.

2.7 Conclusie

Samenvattend zijn er verschillende opvattingen en theorieën over de gevolgen die de ongekend snelle toename van content op internet teweeg zouden kunnen brengen. Het relatief korte bestaan van het internet heeft echter tot gevolg dat deze theorieën nog relatief weinig of zelfs geen praktische en/of feitelijke onderbouwing hebben op het onderzoeksgebied van deze masterthesis.

Een belangrijk concept in deze beeldvorming is het principe van de Long Tail, welke inzicht verschaft in de gevolgen van de toename in diversificatie en omvang van content op internet. De toename in diversiteit van content stelt de gebruiker in staat om een toenemend gespecialiseerd pakket aan content samen stellen om aan te sluiten bij zijn of haar persoonlijke doelen. Enerzijds bevestigt de Long Tail hiermee dus de (potentiële) meerwaarde van de diversificatie en specialisatie van content. Anderzijds laat het concept ook zien dat door de toename in specialistische content, en het daarmee groeiend aantal contenttypen en metadata, het doorgronden van deze hoeveelheid content steeds complexer wordt.

Ook heeft de literatuurstudie uitgewezen dat er een hiaat zit in de bestaande literatuur over zoekdoelen. Het terugvinden van een eerder gevonden website was nog niet opgenomen in de mogelijke doelstelling van de gebruiker. Dit terugvinden is daarom als extra subdoel opgenomen in de categorie van het navigationeel zoekdoel. Het navigatiele zoekdoel bestaat daarom uit twee subdoelen; het oorspronkelijke navigatiele zoekdoel als het subdoel *navigate* en het subdoel van het terugvinden van een eerder gevonden website *retrieval*.

Door destilatie van de zoekdiensten, die op internet aanwezig zijn, bleven er vier categorieën over waarin de diensten kunnen worden ingedeeld.

Deze categorieën zijn gevormd en benoemd aan de hand van de methode die ze gebruiken om gebruikers te ondersteunen om webcontent te ontsluiten, respectievelijk de zoekmachines, folksonomieën, webdirectories en weblogs. Door deze categorieën te combineren met de drie mogelijke zoekdoelen van gebruikers, het navigatiele,

informatie en transactie doelen en met de bijbehorende subdoelen, ontstaat er een raamwerk waartegen de resultaten van de dienstenanalyse en het gebruikersonderzoek afgezet kunnen worden.

3. Methodologische verantwoording

Het onderzoek bestaat uit twee delen: een dienstenanalyse en een gebruikersonderzoek.

Het eerste deel is een analyse van de diensten en hun methoden die op internet gebruikers ondersteunen in het behalen van hun zoekdoelen. Hierin is een inventarisatie gemaakt van de verschillende mogelijkheden die er op internet geboden worden aan gebruikers om content te zoeken en/of te vinden. Ook komt in deze analyse naar voren wat de voor- en nadelen zijn van de verschillende zoekmethoden.

In het tweede deel is aan de hand van een survey onderzocht hoe gebruikers tegenover deze mogelijkheden staan, op welke wijze en met welk doel ze deze diensten gebruiken.

3.1 Verantwoording dienstenanalyse

Het eerste deel van het onderzoek is een analyse van verschillende diensten die op internet gebruikers ondersteunen in het zoeken naar informatie.

Voor dit onderzoek is gezocht naar de meest gebruikte en toonaangevende diensten op het gebied van informatie ontsluiting. Ter indicatie zijn hiervoor de vele top 10, top 20 en top 100 lijsten van de meest bezochte websites en diensten als uitgangspunt genomen. Er zijn een groot aantal bedrijven die dit meten, zij het op zeer uiteenlopende wijze. Sommige van deze bedrijven, zoals bijvoorbeeld Alexa¹¹, meten dit aan de hand van software die op de pc wordt geïnstalleerd. Dit is enigszins vergelijkbaar met de kijkcijferboxen die gebruikt worden door Stichting Kijkonderzoek om het televisiepubliek te meten. Andere bedrijven, zoals Technorati¹² bepalen deze hiërarchie aan de hand van het aantal onderlinge links tussen websites die in hun indexering voorkomen.

Helaas zijn deze lijsten zijn niet altijd een goede indicatie; de lijst van Technorati is bijvoorbeeld voornamelijk gericht op weblogs en de samenstelling van de groep mensen die meewerkt aan de onderzoeken van Alexa is onduidelijk. Daarnaast worden bezoekcijfers op

¹¹ <http://www.Alexa.com>

¹² <http://www.Technorati.com>

verschillende manieren berekend, bijvoorbeeld door het wel of niet meerekenen van terugkerend bezoek, de tijdsduur van sitebezoek, het aantal binnen de site bezochte pagina's etc. Deze verschillende manieren kunnen afwijkingen in de lijsten veroorzaken. De lijsten zijn veelal internationaal of Amerikaans georiënteerd en hierdoor ontbreekt er vaak een waardering van Nederlandse, of door Nederlanders bezochte internetsites. Bedrijven als Multiscope¹³ proberen aan de hand van wegingen meerdere van deze rankings samen te voegen tot één lijst, die daardoor betrouwbaarder zou moeten zijn.

Ook al zou er een eenduidig beeld bestaan over welke diensten het meest gebruikt worden, dan nog is dit geen garantie: een hoog bezoekersaantal betekent niet per definitie dat een dienst toonaangevend is. Zo kunnen sites zeer toonaangevend zijn binnen een niche, een bepaalde groep gebruikers, of ze kunnen veel invloed uitoefenen op die individuen die vooraan in een innovatieproces staan.

De uiteindelijke selectie van de cases voor dit onderzoek is daarom wel gebaseerd op deze lijsten, maar verder naar eigen inzicht van de onderzoeker gemaakt. In de afweging zijn onder anderen meegenomen: hoeveelheid on- en offline aandacht die aan verschillende diensten wordt besteed, de verschillende statistieken over bezoek en gebruik en snelheid in toename van het aantal bezoekers van sites.

Ook is er voor dit onderzoek een kader opgesteld van praktische eisen waaraan de diensten moesten voldoen.

Zo moesten de gekozen diensten voor *alle* internetgebruikers *kosteloos* toegankelijk zijn.

Daarnaast moet de diensten links, verwijzingen etcetera faciliteren naar content *buiten* het kader van hun eigen dienst en mochten ze geen beperkingen hebben wat betreft het bestandsformaat van de content of het onderwerp waarnaar ze doorverwijzen.

Als laatste moest de zoekmethode die de dienst biedt aan gebruikers zonder in te loggen, registreren, of een soortgelijke handeling, volwaardig te gebruiken zijn.

Deze eisen zijn opgesteld met twee concrete doelen voor ogen. Het eerste doel is om te garanderen dat de geselecteerde diensten ook daadwerkelijk een specialisatie hebben op het gebied van informatieontsluiting. Het tweede doel is ervoor te zorgen dat er een zo ongekleurd mogelijk beeld wordt gegeven van het zoekgedrag van internetgebruikers, doordat er geen enkele groep gebruikers of content wordt uitgesloten in het onderzoek.

¹³ <http://www.Multiscope.nl>

Nu zijn er enkele uitzonderingen op deze randvoorwaarden.

Zo wordt in sommige landen het internetgebruik gecensureerd en is een dienst hierdoor niet of slechts gedeeltelijk door gebruikers toegankelijk.

Ook pornografie wordt als content door veel diensten (deels) geweerd vanwege vervuiling en sommige diensten beheren het faciliteren van een zoekmachine ook een grote hoeveelheid content als nevenactiviteit (Yahoo).

Sommige voegen zelfs content toe aan een bestaande content (zoals veel weblogs).

Deze, en meer, grijze gebieden zijn kritisch meegenomen in de afweging welke diensten wel en niet onder de paraplu van dit onderzoek vallen.

Uit de lijst diensten zijn uiteindelijk zeventien diensten geselecteerd die in de analyse meegenomen zijn.

Tabel 1: Lijst van geanalyseerde diensten.

<u>Web Directories:</u> <i>Startpagina (startpagina.nl)</i> <i>Craigslist (craigslist.com)</i> <i>DMoz/ODP/Google Directory (google.com/dirhp)</i> <i>Yahoo Directory (dir.yahoo.com)</i>	<u>Zoekmachines:</u> <i>Google (google.com)</i> <i>Lycos (lycos.nl)</i> <i>Ilse (ilse.nl)</i> <i>Altavista (altavista.com)</i> <i>Ask (ask.com)</i> <i>Microsoft Live Search (live.com)</i> <i>Yahoo! (search.yahoo.com)</i>
<u>Folksonomieën</u> <i>Nujij (nujij.nl)</i> <i>Delicious (delicious.com)</i> <i>Digg (digg.com)</i> <i>Stumbleupon (stumbleupon.com)</i>	<u>Weblogs:</u> <i>Nu (nu.nl)</i> <i>Spitsnieuw (spitsnieuws.nl)</i>

3.1.1 Operationalisatie

In hoofdstuk 4 worden de verschillende diensten beschreven aan de hand van de topiclijst (te vinden in bijlage B2) en ook worden de vier methoden beschreven vanuit een gebruikersperspectief. Daarnaast wordt een vergelijkingsanalyse gemaakt. De geselecteerde diensten zijn zeer divers. Om toch tot een vergelijkingsanalyse te komen is een lijst opgesteld met eigenschappen die de onsluitingseffectiviteit bepalen. De diensten worden aan de hand van deze lijst vergeleken. Dit is een eerste exploratieve lijst van eigenschappen, die als doel heeft een eerste inzicht te genereren in de voor- en nadelen van de verschillende

methoden. In vervolg onderzoek zou deze lijst vervolledigd en verder getoetst kunnen worden.

De gedefinieerde eigenschappen zijn

- Omvangrijkheid van de ontsloten content (grootte database).
- Snelheid van verwerken van nieuwe content.
- Kosten voor dienst om content te ontsluiten.
- Mate van indexering op inhoud van content.
 - tekstuele content (e.g. Google)
 - audio of visuele content (e.g. menselijke beoordelaars van een folksonomie)
- Mate waarin relevantie wordt bepaald door actualiteit.
- Mate van flexibiliteit van organiseren (e.g. webdirectory kan niet gemakkelijk anders worden opgezet).
- Mate van verstoring (dalende relevantie) van zoekresultaten (e.g. beïnvloeding zoekresultaten door commerciële belangen, opzettelijk verkeerde tags toekennen, onnauwkeurige metadata of te generieke metadata zoals de tag “good” en polysemy).
- Aanwezig zijn van waardeoordeel over de content.
- Mogelijkheid om zoekresultaat terug te vinden.
- Actief versus passief (e.g. Spitsnieuws.nl biedt actief content aan, Startpagina.nl niet).
- Gebruiksvriendelijkheid
 - intuïtief zoekkarakter (toegankelijkheid)
 - Hoeveelheid vereiste input om tot zoekresultaat te komen

Tenslotte worden in H4 de diensten gerelateerd aan de zoekdoelen en subdoelen zoals vastgesteld in paragraaf 2.6. Een eerste inschatting is gemaakt hoe de diensten presteren qua informatieontsluiting in relatie tot de verschillende zoekdoelen. De inschatting wordt gevalideerd in het gebruikersonderzoek (H5).

3.2 Verantwoording gebruikersonderzoek

3.2.1 Onderzoekstype

Het tweede deel van het onderzoek bestaat uit een *survey-onderzoek*. Door middel van deze empirische methode is het mogelijk aan de hand van een kwantitatief onderzoek inzicht te krijgen in de opvattingen, meningen en gedragingen van een groep of individu. (Hansen et al., 1998)

Het gebruikersonderzoek richt zich op de opinie van de internetgebruiker. Het gaat hier om een *explorerend* onderzoek. Dit houdt in dat het onderzoek gericht is op het zoeken naar verbanden of verschillen in kenmerken binnen een specifieke groep. (Baarde & de Goede, 1997) Dit kan gedaan worden aan de hand van verschillende methoden: een survey-onderzoek, een focus-groep onderzoek of diepte-interviews.

Aangezien de onderzoeksgroep (internetgebruikers) omvangrijk en divers van karakter is, is het van belang dat de onderzochte groep zoveel mogelijk generaliseerbaar is voor de totale populatie. Vanwege het relatief lage aantal respondenten bij diepte-interviews en het feit dat een focusgroep onderzoek niet generaliseerbaar is naar de totale groep internetgebruikers is gekozen voor een survey-onderzoek. Bij het onderzoeken van een groep geldt bij dit type onderzoek dat hoe hoger de responsaantallen zijn, hoe representatiever de resultaten.

De survey kan op vier verschillende manieren aan respondenten worden voorgelegd. Dit kan gebeuren per telefoon, persoonlijk, schriftelijk of digitaal. (Hair, Bush, Ortinau 2003)

Aangezien de onderzoekspopulatie bestaat uit internetgebruikers, is gekozen voor een digitaal survey-onderzoek. Hier kleven enkele nadelen aan, zowel vanuit het oogpunt van de onderzoeker als de respondent. Het grootste nadeel aan een survey-onderzoek via internet is dat de onderzoeker weinig invloed kan uitoefenen op het type respondenten. (Fowler, 2002) Deze verkleuring kan deels weggenomen worden door rekening te houden met zaken die hier invloed op kunnen hebben. Bij dit onderzoek zijn de twee belangrijkste factoren die meegewogen moeten worden de plaats van initiële verspreiding en de primaire groep die gevraagd is mee te doen. Hier wordt later in dit hoofdstuk op teruggekomen.

Verder is het voor de onderzoeker niet makkelijk om te controleren of surveys meerdere keren worden ingevuld door dezelfde respondent. Wel is er gecontroleerd hoe vaak een survey is ingevuld vanaf een bepaalde internetaansluiting. Dit is gedaan door op het moment van deelname het ip-adres van de deelnemers op te slaan. Deze methode is niet volledig waterdicht; meerdere mensen kunnen het op dezelfde computer invullen, of er kunnen meerdere computer aangesloten staan op één internetaansluiting. Toch biedt deze methode de mogelijkheid om grote afwijkingen te signaleren en deze uit het databestand te verwijderen. Het verwijderen van data om deze reden is niet nodig gebleken.

Een nadeel voor de respondent is dat deze geen vragen kan stellen aan de persoon of organisatie die de survey afneemt. Onduidelijkheden kunnen leiden tot verkeerd ingevulde of onafgeronde surveys, hierom is het van belang dat de respondent goed geïnformeerd wordt over de wijze waarop de survey ingevuld moet worden. Dit is gebeurd door bij de vragen een korte toelichting te plaatsen. Ook is de survey, voordat deze is uitgezet, voorgelegd aan proefpersonen om zo inzicht te krijgen in eventuele onduidelijkheden en aan de hand hiervan verder te verbeteren.

Deze nadelen wegen echter niet op tegen de voordelen die een internetsurvey met zich meebrengt. Een groot voordeel aan dit type onderzoek is dat er ondanks een beperkte hoeveelheid middelen, zoals geld en manuren, een kwalitatief volwaardig onderzoek gehouden kan worden onder een grote populatie. Met speciale survey-software kan deze omvangrijke hoeveelheid feedback direct georganiseerd worden in werkbare databestanden. Ook voor de respondenten zijn er voordelen aan deze manier van ondervraging. Het eerste voordeel hiervan is dat de respondent zelf kan beslissen op welk moment hij of zij de enquête invult.

Het tweede voordeel is dat de respondent zelf kan bepalen met welke snelheid hij of zij de enquête invult.

Nog een belangrijk voordeel van deze manier van ondervragen is dat het anoniem is, waardoor een respondent minder snel een antwoord zal geven dat als sociaal wenselijk wordt gezien. (Fowler, 2002)

Voor het maken van de survey is de dienst Survey Monkey¹⁴ gebruikt.

¹⁴ <http://www.Surveymonkey.com>

3.2.2 Verspreidingmethode

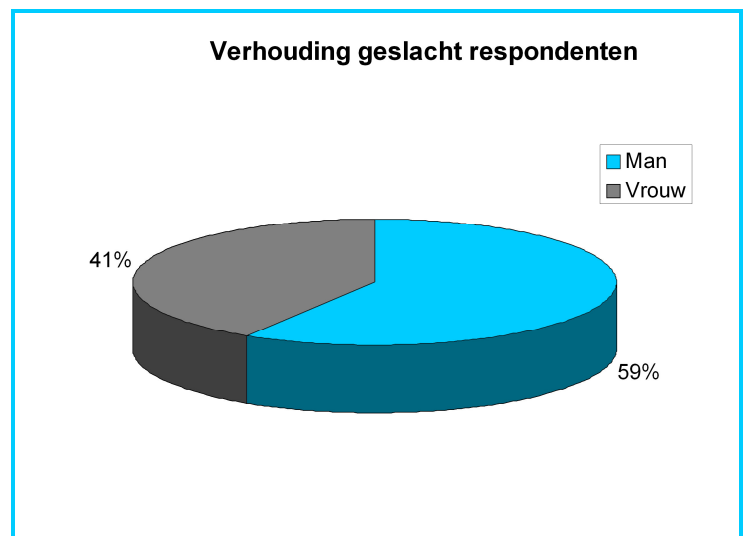
De survey is verspreid aan de hand van een aangepaste sneeuwbalmethode. De sneeuwbalmethode wordt normaal gebruikt om literatuur te doorzoeken en informatie te vergaren. Dit gebeurt door de (literatuur)verwijzingen in een artikel te volgen, om zo nieuwe literatuur te vinden. Op dit punt herhaalt het proces zich en hierdoor zal het aantal artikelen zich snel uitbreiden. Vandaar ook de metafoor van een sneeuwbal die in omvang toeneemt als hij van de berg afrolt. In respondentenonderzoek wordt de sneeuwbalmethode ook gebruikt om respondenten te vinden. De onderzoeker interviewt een respondent, die binnen de onderzoeksdoelgroep valt. Aan deze respondent vraagt de onderzoeker vervolgens de contactgegevens van andere mogelijke respondenten voor het onderzoek.

Voor dit onderzoek is gebruik gemaakt van de sociale netwerken van gebruikers. In de email die naar de eerste groep respondenten is gestuurd stond, naast het verzoek om mee te werken aan het onderzoek, ook de vraag de email door te sturen naar internetgebruikers in zijn of haar omgeving. De link naar de survey is ook op internet op het forum van Thesistools¹⁵ geplaatst met hetzelfde verzoek. Als de benaderden hieraan meewerken en de link aan één of meer personen doorsturen, kan een exponentiële groei ontstaan in de verspreiding van de survey. Over deze methode van informatieverbreiding op internet wordt vaak gesproken onder de naam virale verspreiding. Er wordt hiermee een overeenkomst getrokken tussen deze vorm van verspreiding en de infectiewijze van een virus. Misschien zou de term *virale surveying* hier dan ook niet misstaan.

¹⁵ <http://www.Thesistools.com>

Per email is de survey uitgezet onder ongeveer 700 studenten van de Erasmus Universiteit en 4000 medewerkers van TNO in Nederland. Verder is de survey verstuurd naar 350 leden van tennisvereniging De Hoge Limiet Rotterdam, 60 personeelsleden van de Reconvalescentenschool in Rotterdam en het volledige mailadressenbestand van Snijders Hi-Fi Stereo, ook te Rotterdam (circa 400 contacten). Ook is de survey op internet gepost in het forum op de website thesistools.com (500 views¹⁶). Over het verdere verloop van verspreiding is geen informatie beschikbaar. De respondenten hadden drie weken de tijd om mee te doen aan het onderzoek, van 16 juni tot 7 juli 2008.

Het totaal aantal mensen dat uiteindelijk binnen deze periode heeft meegedaan aan het onderzoek is 1066. Hiervan hebben er 896 (84.1%) de survey (succesvol) afgerond. De leeftijden van de deelnemer zijn normaal verdeeld tussen de 16 en 72 jaar, met een gemiddelde leeftijd van 35. De verdeling tussen het aantal mannen en vrouwen ligt afgerond op 59 procent mannen versus 41 procent vrouwen.



Figuur 5: Verhouding tussen het geslacht van de respondenten

3.2.3 Representativiteit

Met het generaliseren van de uitkomsten moet rekening gehouden worden met een aantal factoren.

Ten eerste is het onderzoek in het Nederlands gehouden, dus de resultaten gelden alleen voor Nederlandssprekende internetgebruikers.

Daarnaast moet er rekening gehouden worden met de leeftijd van de respondenten. Van de 882 respondenten waren er slechts vijf minderjarig (16 of 17 jaar). Er kan dus gesteld worden dat de resultaten van het onderzoek alleen gelden voor volwassen internetgebruikers.

¹⁶ Het topic over deze survey had 500+ views op het forum, hoeveel hiervan omgezet zijn in een respons is niet te zeggen. Dit getal is slechts ter indicatie.

Nog een factor waar rekening mee gehouden moet worden, is dat een relatief groot deel van de eerste lichting uitgezette surveys terecht gekomen is bij hoger opgeleiden, namelijk TNO-medewerkers, studenten van de Erasmus Universiteit en bezoekers van de site Thesistools (voormalig Studentonderzoek.com). Hoe de verspreiding van het onderzoek zich hierna heeft voltrokken is niet te controleren, toch kan er met enige zekerheid vanuit worden gegaan dat hierdoor het gemiddelde opleidingsniveau van de uiteindelijke groep respondenten niet vergelijkbaar is met die van een dwarsdoorsnede van de Nederlandse internetgebruikers. Informatie om dit te controleren is echter niet beschikbaar.

Dit neemt niet weg dat het onderzoek, oa. vanwege de grote respons, een zeer goede indicatie kan geven over gebruikers en hun omgang met online content en diensten die betrekking hebben op de ontsluiting ervan.

3.2.4 Surveyconstructie

De survey bestaat uit vijf delen.

In het eerste gedeelte is er gevraagd naar de bekendheid van diensten onder gebruikers. Hierin hadden ze de mogelijkheid om uit een lijst van diensten een selectie te maken welke ze wel of niet kenden, en zo ja, hoe vaak ze deze diensten gebruiken. De lijst van diensten is dezelfde als die van de diensten analyse aan het begin van dit hoofdstuk. Deze lijst bestaat uit zeventien verschillende diensten, die ingedeeld kunnen worden in de categorie zoekmachine, folksonomie, weblog of webdirectory. Ook was er de mogelijkheid voor de respondenten om aan deze lijst hun favoriete diensten toe te voegen als deze er nog niet in voorkwamen. Alle toegevoegde diensten (zie Figuur 40) uit vraag 5, 6 en 7 van de survey zijn eerst individueel getoetst of ze binnen de scope van het onderzoek vallen. Dit is gedaan aan de hand van dezelfde randvoorwaarden waaraan de eerder geselecteerde diensten reeds aan moesten voldoen (zie paragraaf 3.1).

De toegevoegde diensten die binnen de scope vielen zijn ingedeeld naar de vier typen methoden. Dit is gedaan door individuele observatie en analyse van de websites aan de hand van de kenmerken zoals beschreven in paragraaf 2.5 *De methoden*. Door deze categorisering van de toegevoegde diensten was het mogelijk deze mee te nemen in het onderzoek om vervolgens een nauwkeuriger beeld te schetsen van het gebruik van de verschillende zoekmethoden.

Het tweede deel van de survey is opgezet om inzicht te krijgen welke diensten het meest gebruikt worden en met welke doelstelling deze gebruikt worden. De doelstellingen waarvoor een dienst wordt gebruikt zijn ontleend aan de onderzoeken van Rose en Levinson (2004) en Broder (2002). In hun onderzoeken hebben zij aan de hand van verschillende zoekopdrachten een classificatie opgesteld met welke doelen een gebruiker een zoekopdracht uitvoert. Er worden drie zoekdoelen gedefinieerd: het *navigatieve*, het *informatieve* en het *transactieve* (ook wel resource (Rose 2004:18) geheten) doel. (Broder 2002:5) Deze zoekdoelen zijn in het conceptueel raamwerk uitgebreid toegelicht. De stellingen waar in de survey gevraagd wordt op te reageren zijn een aangepaste en ingekorte versie van de stellingen waarmee Rose (2004:15) de verschillende zoekdoelen uiteenzet. De gebruikte stellingen kunnen terug gevonden in de survey (Bijlage 3).

Het derde deel van de survey bestaat uit een aantal stellingen waarin gebruikers hun algemene mening over informatie zoeken op internet kunnen geven.

Het vierde gedeelte van de survey is een open vraag over de werkwijze die gebruikers hanteren om informatie boven water halen op internet. Deze vraag is open gelaten om de respondenten zo de mogelijkheid te geven om een antwoord te geven, wat misschien niet mogelijk was binnen de relatief rigide kaders van de multiple choice vraagstelling van de survey.

Het vijfde gedeelte van de survey bestaat uit de persoonsgegevens van de respondenten om inzicht te krijgen in de karakteristieken van de populatie.

3.2.5 Dataverwerking

De data uit de survey is aan de hand van een aantal statistische computerprogramma's, waaronder Microsoft Excel en SPSS.

Van het eerste deel van de survey, over de naamsbekendheid en het gebruik van de diensten en bijbehorende methoden, is het antwoordschema gekoppeld aan numerieke waarden. Dit hield in de praktijk in dat de antwoorden *nooit van gehoord, gebruik ik zelden tot nooit, gebruik ik soms, gebruik ik regelmatig, gebruik ik vaak en gebruik ik altijd*, respectievelijk gekoppeld zijn aan de waarden 0, 2, 4, 6, 8 en 10.

Aan de hand hiervan is een gemiddelde waarde tussen 0 en 10 berekend. Dit getal geeft een indicatie van de frequentie in het gebruik van een dienst en/of de methode die bij de dienst hoort.

Voor het tweede en derde deel van de survey zijn de berekeningen gemaakt aan de hand van de frequentie waarop antwoorden zijn gegeven ten opzichte van andere antwoordmogelijkheden. Hierdoor zijn de antwoorden dus weergegeven als de relatieve frequentie van het antwoord

De antwoorden op de open vraag, het vierde deel van de survey, zijn door middel van software zichtbaar gemaakt aan de hand van een wordcloud. Deze is gebaseerd op de frequentie waarmee woorden voorkomen in de antwoorden van de gebruikers. Hierin zijn veelgebruikte Nederlandse woorden weggelaten¹⁷.

Het vijfde deel van de survey bevat enkel statistische gegevens, waardoor het mogelijk was de andere vragen te segmenteren aan de hand van bijvoorbeeld leeftijd en geslacht.

De survey zoals deze is afgenomen kan teruggevonden worden in de bijlage (B3).

¹⁷ <http://www.Wordle.net>. Wordle is een online tagcloud-generator.

3.3 Verantwoording verdieping onderzoeksresultaten

Na de diensten-analyse (hoofdstuk 4) en het gebruikersonderzoek (hoofdstuk 5) volgt het hoofdstuk (hoofdstuk 6) waarin de resultaten van deze onderzoeken met elkaar in verband worden gebracht.

In het hoofdstuk is, onderverdeeld per zoekmethode, een analyse gemaakt tussen de werkwijze van de methode en het gebruik door de respondenten.

Hieruit zal naar voren komen in hoeverre er een verband bestaat tussen de door internetgebruikers gekozen methode om hun zoekdoel te behalen en de capaciteiten van deze methode op het gebied van dit zoekdoel.

4. Resultaten Dienstenanalyse

In dit hoofdstuk zullen de resultaten van de dienstenanalyse en het gebruikersonderzoek besproken. Er zal hier ook antwoord gegeven worden op de desbetreffende deelvragen.

4.1 Zoekmachines

4.1.1 Diensten

In de categorie diensten die aan hun gebruikers een zoekmachine bieden zijn voor dit onderzoek de diensten Google, Yahoo!, Live!search, Ask, Lycos, Altavista, en Ilse gekozen.

Google

De eerste in dit rijtje is Google.

Google heeft in Nederland veruit het grootste marktaandeel. Met een aandeel van rond de 93 procent is de dienst heer en meester in de Nederlandse zoekmachine markt. (Checkit, 2008) Ter vergelijking, de andere diensten in de hierboven genoemde opsomming hebben een marktaandeel lager dan drie procent. (Checkit, 2008)

Google biedt twee mogelijkheden van zoeken. Wanneer er in het zoekvenster een of meerdere zoektermen zijn ingevoerd kan de gebruiker kiezen tussen de optie *Google*



Figuur 6: Screenshot Google.com. (januari 2009)

zoeken of *Ik doe een gok*. Wanneer er gekozen wordt voor de optie *Google zoeken* krijgt men een lijst van alle relevante resultaten. In deze lijst staat het, volgens Google, meest relevante resultaat bovenaan gevolgd door toenemend minder relevante resultaten.

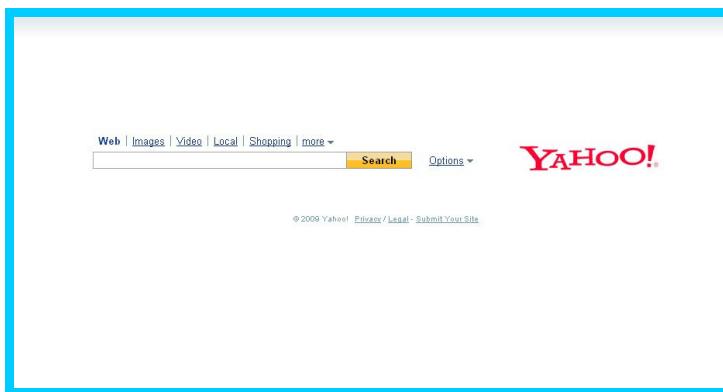
Wanneer de keuze wordt gemaakt voor de mogelijkheid *Ik doe een gok* verwijst Google de gebruiker direct naar de meest relevante content.

De relevantie van content wordt door Google bepaald aan de hand van *Pagerank*. Dit is software die aan de hand van een complex algoritme een numerieke waarde toekent aan content. Dit gebeurt aan de inhoud van de pagina, de hoeveelheid en weging van verwijzingen, populariteit, datering enzovoort.

Naast deze zoekopties kan er ook gekozen worden binnen een bepaald bereik of contenttype te zoeken. De voornaamste opties hierin zijn zoeken op het internet algemeen, binnen een taalgebied, naar afbeeldingen of naar video's. Er zijn verder op de pagina nog andere opties en links te zien. Deze zullen buiten beschouwing gelaten worden, vanwege het gebrek aan relevantie ten opzichte van het onderzoek.

Yahoo!

Yahoo! is één van de oudste zoekdiensten op internet. De dienst is gestart in 1994. Eind



Figuur 7: Screenshot search.yahoo.com. (januari 2009)

jaren negentig begon Yahoo! met het aanbieden van een zoekmachine aan haar gebruikers, maar pas in 2004 kwam Yahoo! met een zelfstandig werkende, spidergebaseerde zoekdienst. Een saillant detail is dat voor de verzelfstandiging van de

zoekmachine deze tussen 2001 en 2003 deze werd aangestuurd door middel van Google.

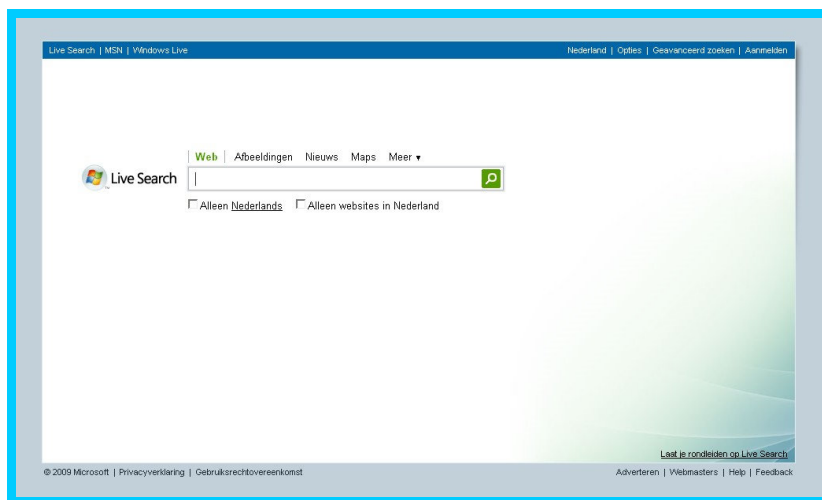
De werkwijze van de zoekmachine van Yahoo! is redelijk conform de standaard werkwijze van zoekmachines; speciale spidersoftware (door)zoekt de content en metadata op internet, indexeert deze en slaat deze op. Wanneer een internetgebruiker een zoekopdracht opgeeft wordt de relevantievolgorde van de teruggekeerde resultaten bepaald aan de hand van inhoud, populariteit en datering.

Yahoo! biedt websites ook de mogelijkheid om tegen betaling van ongeveer 250 euro per jaar opgenomen worden in de resultaten van de zoekmachine. Dit is geen garantie voor de bedrijven om hoog in de relevantielijst te komen, maar verhoogt de kans hierop sterk. Door deze commerciële aanvulling op de oorspronkelijke werkwijze bestaat er binnen deze dienst het gevaar van een belangenverstremming. Commerciële doelstellingen komen over het algemeen niet overeen met de doelstellingen van gebruikers.

Live! Search.

Live! Search is de zoekmachine van softwaregigant Microsoft.

De zoekmachine is relatief nieuw en heeft al vele namen en gedaantes aangenomen, waaronder Windows live search en MSN search. De dienst bestaat uit een zelfstandige



Figuur 8: Screenshot Live.com. (januari 2009)

zoekmachine, een index en een crawler, zoals vergelijkbaar bij veel andere zoekdiensten. Volgens tests zijn de zoekresultaten van Live! Search relatief commercieel gekleurd in vergelijking met andere zoekmachines. (Wall, 2006)

Dit wordt deels veroorzaakt

door het Live Cashback Program. In het kort komt dit er op neer dat partijen die diensten of producten verkopen via internet, korting kunnen geven aan de klanten die doorverwezen zijn via Live! Search. Dit kunnen ze vanwege kortingen die Microsoft ze hiervoor geeft op advertentie-uitgaven.

Ask

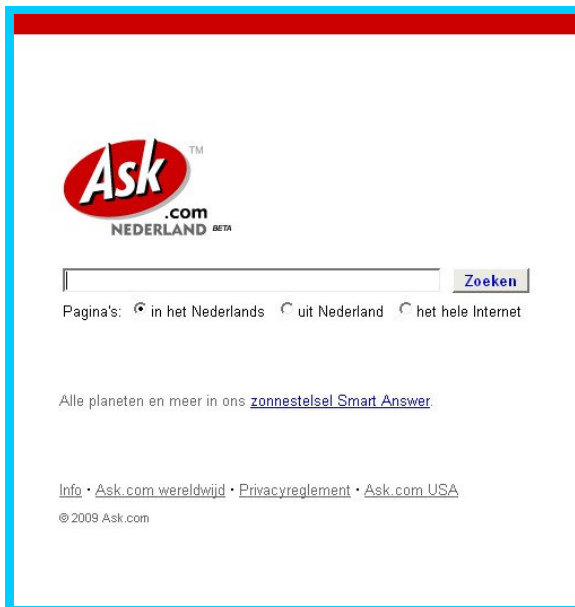
De zoekmachine Ask is het resultaat van een langdurige transformatie. De dienst heeft in het verleden een aantal betrekkelijk experimentele technieken gebruikt om de concurrentie aan de te gaan met de grote zoekmachines. Deze hebben echter geen van alle succes gehad. Het gaat op het moment niet erg goed met het bedrijf achter Ask.

Dit is voor een deel te merken in de kwaliteit van de zoekmachine en de grote hoeveelheid gesponsorde links die site toont.

Er bestaat een hardnekkige speculatie dat Ask het plan heeft zich voornamelijk op vrouwen

te richten, en hiermee een niche te creëren, in plaats van de directe concurrentie aan te gaan met Google. (Buresh, 2008)

De huidige werking van de dienst is betrekkelijk standaard. Zo kan er op het web gezocht worden, binnen Nederlandse pagina's en afbeeldingen. Iets wat wel een redelijk uniek zoekelement is, is de mogelijkheid om alleen te zoeken binnen weblogs. Hierdoor is het gemakkelijk te zoeken naar het laatste nieuws omtrent een onderwerp. Jammer genoeg als er gezocht wordt op de naam van



Figuur 9: Screenshot Ask.com. (januari 2009)

een weblog zelf als onderwerp, komen er alleen vermeldingen van alle items op die specifieke dienst in plaats van nieuws over de dienst.

De zoekmachine van de dienst Lycos verkrijgt zijn zoekresultaten uit drie verschillende



Figuur 10: Screenshot Lycos.com. (januari 2009)

bronnen. De eerste bron is Overture, een bedrijf dat gespecialiseerd is in het aanleveren van commerciële links aan zoekmachines. De tweede bron is een team van werknemers van Lycos zelf. Dit team indexeert en filtert websites zelf. Als laatste heeft ook Lycos spidersoftware die het web afspeurt naar alle content en deze indexeert en opslaat. Naast de meest voorkomende zoekmogelijkheden, zoals zoeken op

het web, naar afbeeldingen en video biedt Lycos ook enkele gespecialiseerde zoekmogelijkheden. Als eerste biedt Lycos de mogelijkheid om te zoeken naar personen. Dit is in grote lijnen vergelijkbaar met een digitale telefoongids. Ook biedt Lycos de mogelijkheid te zoeken naar bedrijven in een bepaalde regio. Helaas werken deze opties alleen voor Amerikaanse personen en bedrijven. Een interessante optie is de mogelijkheid te zoeken naar nieuws. Er kan gekozen worden wat voor nieuws er gezocht wordt, zoals entertainment, sport, economisch en gezondheid. Helaas vind de dienst alleen Engelstalig nieuws. Er wordt ook aan de zijkant van de pagina automatisch een overzicht gegeven van de populairste news-items.

Altavista

De dienst Altavista kent in vergelijking met groot aantal van de andere diensten een relatief lange historie. De dienst opende in 1995 haar online zoekmachine voor het publiek. Altavista was hiermee een van de eerste zoekmachines die op basis van spidersoftware het internet doorzocht. Eind jaren negentig behoorde Altavista tot de top van meest gebruikte

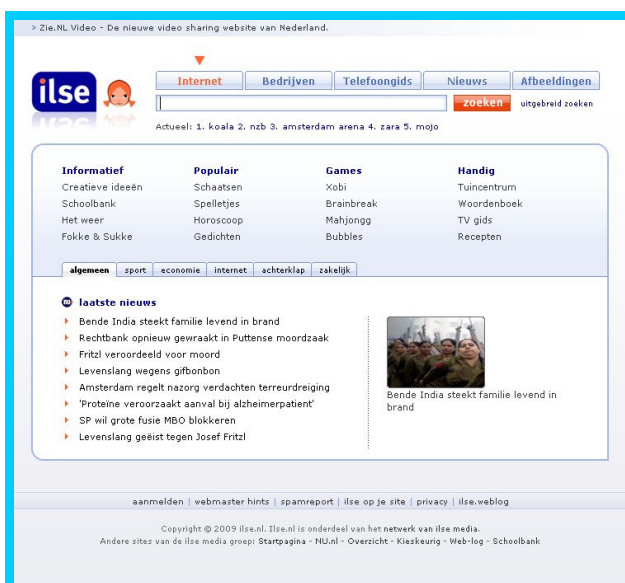


Figuur 11: Screenshot Altavista.com. (januari 2009)

Overture, wat op zijn beurt weer overgenomen werd door Yahoo!. Altavista is momenteel geen autonome zoekmachine meer. De huidige zoekresultaten van die dienst zijn direct afkomstig van de zoekmachine van Yahoo!.

Ilse

De enige Nederlandse dienst die zich, alhoewel erg beperkt, aan de meest populaire



Figuur 12: Screenshot Ilse.nl. (januari 2009)

(Nu.nl, Startpagina.nl). Dit is duidelijk terug te zien in de resultaten van zoekopdrachten.

zoekmachines. De wet van de remmende voorsprong begon voor de dienst eind jaren negentig grote gevolgen te hebben. Met name de populariteit van Google kostte de dienst veel gebruikers. De dienst werd overgenomen door

zoekmachines kan meten is Ilse.

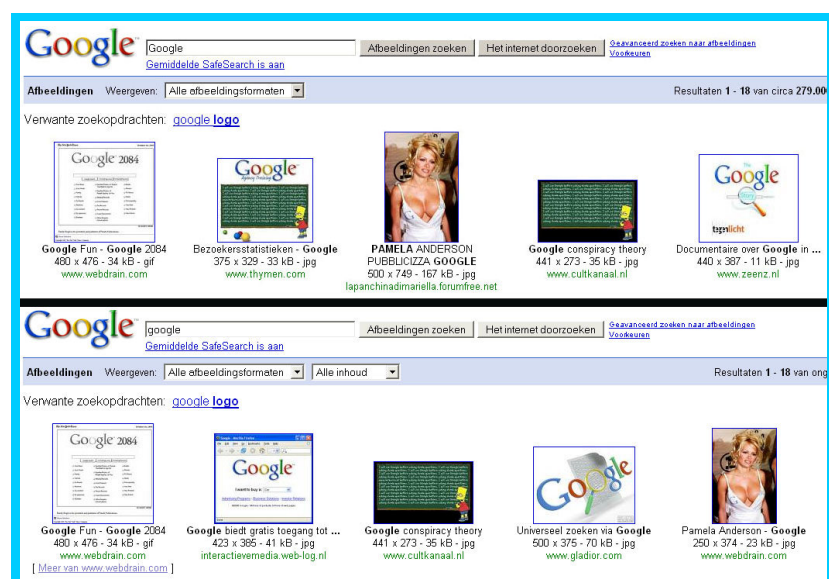
De zoekmachine van Ilse biedt een aantal zoekmogelijkheden: er kan gezocht worden naar tekst, afbeeldingen, personen, bedrijven en nieuws. De nadruk bij Ilse ligt op het zoeken naar content binnen Nederlandse websites. Het moederbedrijf van Ilse is ook eigenaar van enkele populaire nieuwsdiensten en webdirectories

4.1.2 Methode

De zoekmethode, die door zoekmachines gebruikt wordt, heeft voor- en nadelen. Een duidelijk voordeel van deze methode is dat computers het werk doen. Deze kunnen tegen zeer lage kosten met enorme snelheden door een ongekend aantal websites en de bijbehorende content crawlen om deze te indexeren en rangschikken. Hierdoor behoren de databases van content van deze zoekmachines tot de allergrootsten op het internet. Aan de screenshots, er is hier gezocht op het woord *thesis*, wordt duidelijk over welke enorme aantallen er dan gesproken wordt.



Vanwege de methode waarmee sites beschouwd worden, werkt deze methodiek voornamelijk goed voor tekstuele content of content die voldoende metadata bevat. De crawlers komen in de problemen wanneer ze content tegenkomen waaraan ze geen informatie kunnen ontleneren. Dit geldt onder andere voor veel afbeelding-, video- en geluidsformats. Ook sommige programmeertypen voor internetsites worden door hun tekstueel ongestructureerde opbouw en afbeelding-gebaseerde vormgeving vaak niet goed geïndexeerd. Om deze content te indexeren is een zoekmachine dus volledig afhankelijk van eventuele bestaande metadata over het object. De zoekmachine heeft echter geen manier om de juistheid van de metadata te controleren. Een voorbeeld dat zoekmachines niet “in” afbeeldingen kunnen kijken blijkt uit dat er tweemaal,



Figuur 13: Resultaten Google.com, zoekterm: *Google*, resp. 12 december 2008 en 12 maart 2009

¹⁸ Resultaten zoekopdracht: *Thesis*. Respectievelijk Google, Yahoo! en Live!.

met een interval van vier maanden, bij de dienst Google naar afbeeldingen onder de zoekterm het woord Google is gezocht. Deze zoekopdracht zou voor een internetgebruiker een informatieel of een transactioneel zoekdoel kunnen bevredigen, zoals deze zoekopdracht ook voor dit onderzoek informatiele waarde bevat. Tot twee keer toe kwam in de gevonden resultaten een foto van de Hollywood ster Pamela Anderson naar voren. Verder onderzoek naar deze afbeelding wees uit, dat er geen enkel verband bestond tussen de term Google en de afbeeldingen, anders dan de toevoeging van de term aan de metadata. Op deze wijze kan Google dus duidelijk gemanipuleerd worden.

Ook de andere diensten hebben dit probleem. Wanneer met de zoekmachine van Yahoo! gezocht werd naar afbeeldingen met de zoekterm Yahoo, kwamen er op het moment van zoeken op de zevende plaats een vakantiefoto van vier dames in bikini en als negende een uitdagende modelfoto van een mevrouw Anitaakuffo¹⁹. Deze dame werd gevonden als relevant onderwerp vanwege haar emailadres anitaakuffo@yahoo.com, wat als metadata was vermeld bij de foto.

Bij Ask.com geeft eenzelfde zoekopdracht als meeste relevante item een afbeelding van Britney Spears²⁰.

Dat content aan de hand van, onder andere, metadata vindbaar gemaakt wordt voor zoekmachines zorgt ervoor dat er voor zoekmachines geen enkele beperking ligt het gebied van contenttypen. Dit maakt zoekmachines erg flexibel ten opzichte van de ontwikkeling van content en contenttypen.

De zoekmachines worden steeds geavanceerder. Er worden steeds meer factoren meegenomen om te zorgen dat de gebruiker de meest relevante informatie krijgt geleverd. Hiertegenover staan helaas bedrijven die gespecialiseerd zijn in het zogeheten Search Engine Optimisation (SEO). Vanwege commerciële belangen is het belangrijk voor bedrijven of instanties om gevonden te worden door internetgebruikers en deze seo-bedrijven ondersteunen commerciële bedrijven met sluwe methoden om zo hoog mogelijk in de relevantielijst te eindigen. Door dit soort gedrag raken de indexen van zoekmachines

¹⁹ Zoekopdracht uitgevoerd op 24 juni 2008

²⁰ Zoekopdracht uitgevoerd op 24 juni 2008

vervuild en is de respons van de zoekmachine niet altijd bepaald door middel van de relevantie voor de gebruiker, maar door commerciële belangen.

Zoekmachines voegen inhoudelijk geen waardeoordeel toe aan de content en bieden gebruikers geen mogelijkheid om eerder gevonden content terug te vinden. Ondanks de afwezigheid van deze mogelijkheden zijn zoekmachines relatief gebruiksvriendelijk te noemen en kost vereist het weinig input om tot een zoekresultaat te komen. De actieve houding van de gebruiker is echter wel vereist, zoekmachines zijn uit zichzelf passief.

4.1.3 Conclusie

Deze methode waarmee deze diensten informatie ontsluiten voor gebruikers is zeer sterk in grote hoeveelheden tekstuele content en metadata te (door)zoeken en te indexeren. Vanwege het tekstuele karakter van webadressen, en doordat veel informatie op internet in de vorm van tekst wordt aangeboden leent deze methode zich hierdoor goed voor het ondersteunen van gebruikers bij het behalen van navigatieve en informatiele zoekdoelen.

Op het gebied van het transactionele zoekdoel is deze methode minder krachtig. De computers en algoritmen waarmee websites worden geïndexeerd en samengevat zijn niet in staat om een waardeoordeel te geven over de inhoud van een groot aantal contenttypen. Juist deze soorten content, zoals afbeeldingen, video en animaties zijn vaak verbonden met dit navigatieve zoekdoel.

Hieronder zijn de voor- en nadelen van de zoekmethode die zoekmachines gebruikt nogmaals naast elkaar gezet. Ook is er een overzicht gemaakt met een waardering van de (potentiële) competenties van zoekmachines ten opzichte van zoekdoelen die gebruikers kunnen hebben.

Tabel 2: Voor- en nadelen van werkwijze zoekmachines

Voordelen	Nadelen
Vaak omvangrijke database aan geïndexeerde content	Niet in staat de inhoud van audio of visuele content te analyseren
Snelle verwerking nieuwe content	De negatieve invloed van professionele bedrijven op relevantiebepaling van content voor commerciële doeleinden
Relatief lage kosten indexering nieuwe content.	Ontbreken van waardeoordeel over content
Over het algemeen sterk op het gebied van het indexeren van de inhoud tekstuele content	Relatief veel vervuiling en/of overbodige resultaten.
Oneindige flexibiliteit in categorieën van content dmv. metadata.	Eerdergevonden content niet op andere wijze terug te vinden dan door hetzelfde proces te herhalen.
Nier erg veel input vereist om tot een resultaat te komen.	Dienst is passief, actieve houding van gebruiker vereist.
	Niet erg intuïtief.

Tabel 3: Waardering van de potentiële competentie van zoekmachines voor het behalen van zoekdoelen

Zoekdoel	Verwachte prestatie	Uitleg van verwachte prestatie
Navigationeel zoekdoel	++	Bij alle zoekmachines volstaat ingeven van zoekterm 'Erasmus Universiteit' om direct de juiste URL (www.eur.nl) te vinden.
<i>Navigate</i>	<i>++</i>	
<i>Retrieval</i>	<i>+/-</i>	
Informationeel zoekdoel	++	Zoekmachines zijn zeer goed in het inhoudelijk doorzoeken van tekstuele content en kunnen effectief inhoudelijke informatie vinden. Bijvoorbeeld zoektermen 'geboortejaar Napoleon' geeft direct het gewenste resultaat. Wat advies betreft is een zoekmachines iets minder effectief omdat
<i>Directed</i>	<i>++</i>	
<i>Undirected</i>	<i>++</i>	
<i>Advice</i>	<i>+</i>	
<i>Locate</i>	<i>++</i>	

			het zoekresultaat op deze subjectieve vraag bepaalt wordt aan de hand van relevantie, maar niet noodzakelijkheid op basis van de beste interpretatie voor de gebruiker. Het gevonden antwoord kan bijvoorbeeld een populair, maar onjuist antwoord geven.
	Transactioneel zoekdoel	+/-	Zoekmachines zijn minder sterk in het behalen van een transactioneel zoekdoel aangezien de contenttypen die hier veelal mee gepaard gaan, bijvoorbeeld audiovisuele content, niet inhoudelijk kunnen worden geïndexeerd. De zoekmachine moet daarvoor vertrouwen op metadata zoals tags en verwijzingen voor de inhoud. Gevolg hiervan is dat deze zoekopdrachten sterk vervuild kunnen zijn. Bij zoeken naar 'download Mozart' via een zoekmachines zijn de meeste sterk verstoord.
	<i>Download</i>	+/-	
	<i>Entertainment</i>	-	
	<i>Interact</i>	+/-	

4.2 Folksonomieën

4.2.1 Diensten

Uit de methoden die werken met een folksonomie om structuur aan te brengen in online content zijn er, zoals in de methodologie al naar voren is gekomen, vier toonaangevende diensten gekozen. Dit zijn de diensten Delicious, Digg, Stumbleupon en Nujij.

Delicious

De eerste, Delicious, is een van oorsprong Amerikaanse dienst opgericht in 2003 waar momenteel naar schatting meer dan 5 miljoen mensen gebruik van maken.

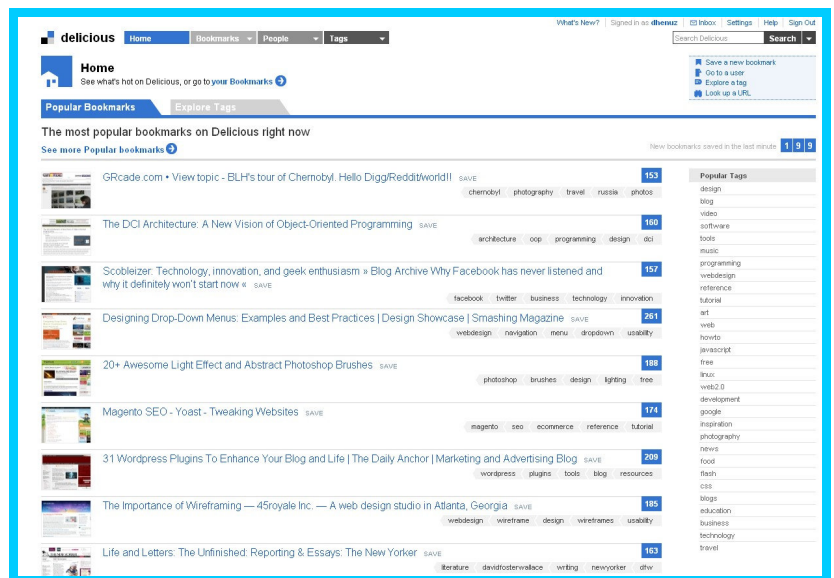
Delicious biedt gebruikers de mogelijkheid om hun favoriete websites, ofwel bookmarks, online op te slaan en er metadata

aan toe te voegen in de vorm van een korte beschrijving en tags.

Ook kunnen gebruikers hun bookmarks onderling uitwisselen en doorzoeken. Er kan gezocht worden op tags, maar ook op alle bookmarks die een specifieke webadres bevatten, bijvoorbeeld alle bookmarks met

microsoft.com erin. Door het aantal keer dat een webadres is

opgeslagen ontstaat er een hiërarchie van de meest populaire bookmarks op de site. Ook is er een lijst met de meest gebruikte tags te vinden. De site is opvallend simpel en maakt het geheel voor de gebruiker erg overzichtelijk.



Figuur 14: Screenshot Delicious.com (januari 2009)

Digg

Digg is een dienst waarmee gebruikers de door hun gemaakte of gevonden content kunnen adviseren aan medegebruikers.

Momenteel zijn er rond de 3 miljoen geregistreerde Digg gebruikers die bijdragen aan de site. Dit doen ze door content te *diggen*. Dit kunnen ze doen door op een speciaal daarvoor bestemde knop te klikken. Wanneer ze dit doen verschijnt de link naar de content die ze *gedigged* hebben op de website van Digg. Deze wordt gerangschikt naar het aantal Diggs, oftewel de populariteit ervan. Als een onderwerp genoeg Diggs heeft verschijnt het op de voorpagina van de site.

De items op de site kunnen op een aantal wijzen doorzocht worden.

Categorieën kunnen doorzocht worden op onderwerp, contentsoort (nieuws, video of afbeelding)

of de combinatie van de twee. Op

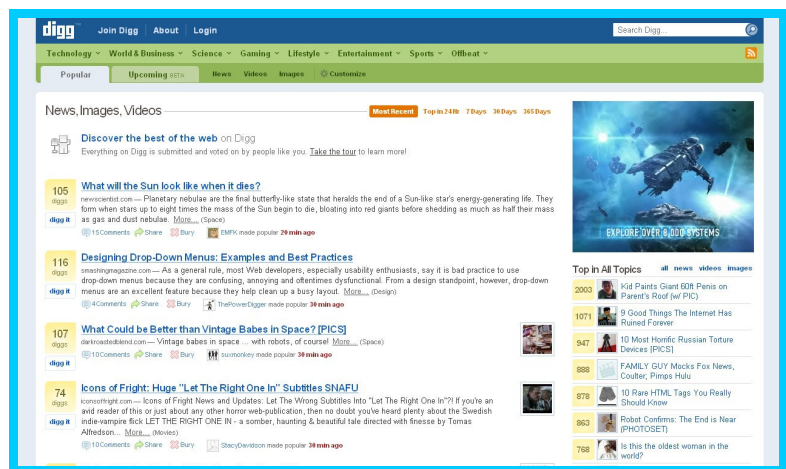
deze wijze kan er bijvoorbeeld gezocht worden naar de

populairste afbeeldingen binnen

de categorie honkbal, of naar nieuws over gaming.

Ook het doorzoeken van de Diggs van andere gebruikers is mogelijk.

Als laatste kunnen de gebruikers door middel van een zoekopdracht de items doorzoeken aan de hand van sleutelwoorden in de titel of het webadres van het item.



Figuur 15: Screenshot Digg.com (januari 2009)

Stumbleupon

De derde dienst, Stumbleupon, is sterk vergelijkbaar met de dienst Digg.

Deze dienst, die rond de vijf miljoen actieve gebruikers heeft, biedt de gebruiker de mogelijkheid om verwijzingen naar content toe te voegen aan de dienst. Ook deze items worden gerangschikt naar onderwerp en naar soort content. De content is handmatig te doorzoeken per



Figuur 16: Screenshot Stumbleupon.com (januari 2009)

categorie of per zoekopdracht. Een opvallende mogelijkheid extra die Stumbleupon biedt boven de andere diensten is dat het voor de gebruiker mogelijk is om de dienst te “leren” wat hij of zij interessant vindt. Het is mogelijk om te specificeren welke onderwerpen de gebruiker interessant vindt en welke niet. Ook kan de gebruiker items als positief of negatief waarderen en zo nog nauwkeuriger aangeven waar zijn interesse ligt. Vervolgens kan de dienst hierdoor de gebruiker items laten zien die, waarschijnlijk, overeenkomen met zijn of haar voorkeur. Ook kan de gebruiker kiezen om een willekeurig item geadviseerd te krijgen door de dienst, gebaseerd op de aangegeven voorkeur. Deze redelijk ongestuurde wijze om naar content op zoek te gaan doet denken aan het zappen door televisiekanalen en lijkt dan ook voornamelijk gebruikt te worden voor een transactioneel motief, ondanks dat de gevonden content niet altijd deze transactionele waarde deelt.

Nulij

De laatste dienst binnen deze methode is Nulij. Deze Nederlandse dienst met meer dan 50.000 leden biedt

gebruikers, net als de vorige twee diensten, de mogelijkheid om gevonden content aan te raden aan andere gebruikers. Ook hier is het mogelijk om de hoogst gewaardeerde items te zien of de hoogst gewaardeerde per categorie. De categorieën zijn redelijk vergelijkbaar met die



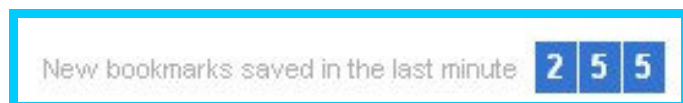
Figuur 17: Screenshot Nulij.nl (januari 2009)

van de andere diensten: algemeen, economie, tech, sport, showbizz etcetera. Toch heeft ook deze dienst zich, naast het feit dat deze Nederlands is, een karakter waarmee het zich distantieert van de andere diensten. De dienst richt zich voornamelijk op nieuws en de geplaatste items zijn vrijwel allemaal afkomstig uit sites met een weblog-format. Hierdoor lijkt de dienst nog het meest op een door gebruikers samengestelde online krant, waar alleen de meest gelezen artikelen de voorpagina halen.

4.2.2 Methode

Wat al deze folksonomieën gemeen hebben is dat de gebruikers uiteindelijk beslissen wat er te vinden is met de dienst en op welke manier. De grote getallen gebruikers zorgen voor een voor een grote dynamiek en flexibiliteit van de diensten. Deze eigenschappen zorgen voor een sterke actualiteit en kunnen zorgen voor een grote betrokkenheid van de gebruiker bij de dienst. (Poortman & Bierens 2005) Hoe groter het aantal gebruikers tags toevoegt aan items, hoe specifiek en descriptiever de uiteindelijke verzameling tags zal zijn. Ook zal logischerwijs de hoeveelheid bookmarks of geadviseerde items stijgen, waardoor er meer content (terug)vindbaar georganiseerd wordt binnen de dienst. De toonaangevende en populaire folksonomieën hebben hierdoor reeds succesvol omvangrijke databases van content gerealiseerd. Het vereist van de gebruiker relatief weinig input om uit deze databases een resultaat te verkrijgen.

Doordat gebruikers de content eerst consumeren wordt alle content inhoudelijk beoordeeld. Deze wordt hierdoor vervolgens door diezelfde gebruikers, inclusief een beschrijving en waarde-oordeel opgeslagen in de database van de folksonomie.



Figuur 18: Bookmarkteller website Delicious

Hierdoor kunnen dus zowel tekst, als audio-visuele content worden geïndexeerd. Door de benodigde menselijke tussenkomst ontstaat er echte wel enige vertraging tussen de publicatie van content en de opname in het bestand van de folksonomische diensten.

Omdat de tags toegevoegd worden door gebruikers zelf, liggen deze vaak een stuk dichterbij de belevingswereld van de gebruiker dan de woorden die in een professioneel georganiseerde taxonomie gebruikt worden. Zoals Weinberger het zo passend formuleert; *“lezers, en niet slechts de auteurs, krijgen de mogelijkheid om objecten te taggen. Een auteur is een autoriteit wanneer het komt tot het verklaren wat de bedoeling is achter zijn of haar werk, maar niet wanneer het gaat over wat dit werk voor anderen betekent. Wanneer het om zoeken gaat is wat een*



Figuur 19: Voorbeeld van een tagcloud

werk betekent voor de zoeker veel belangrijker dan de bedoeling van de auteur. (Weinberger, 2005)

Doordat de tags dichter aansluiten bij de belevingswereld van de gebruikers krijgt het zoeken naar content ook een intuïtiever karakter.

Dit wordt nog eens versterkt door zogeheten tagclouds. Dit zijn wolkvormige visualisaties van de meest gebruikte tags bij een dienst. In een tagcloud worden de populairste termen meestal benadrukt door de orde van grootte of intensiteit van de kleur van de tag.

Nog een groot voordeel van dit sociale systeem zijn de lage kosten ervan. Gebruikers organiseren met grote aantallen tegelijk gratis de content. Bij de enorme stromen content, neem bijvoorbeeld sites als Youtube (filmpjes) en Flickr (foto's), zou het onhaalbaar en daarbij onbetaalbaar zijn om dit door (dure) professionals te laten doen.

Een laatste voordeel van deze methode is dat het niet alleen de verschillende content samenbrengt aan de hand van de

tags, maar hierdoor ook de gebruikers die de metadata

toevoegen verbindt. Op deze wijze



Figuur 20: De mogelijkheid tot het zoeken in andermans favorieten

kunnen gebruikers op zoek naar content bijvoorbeeld via hun peers, mensen met eenzelfde interesse of mensen met veel kennis op een bepaald gebied.

Aan een folksonomie kleven ook een aantal nadelen.

Een van de grootste problemen waar een folksonomie mee te kampen heeft, is het gebrek aan controle over de toegevoegde metadata. Gebruikers zijn niet georganiseerd en niet getraind om informatie (terug)vindbaar op internet beschikbaar te maken. Gevolg hiervan is dat er fouten

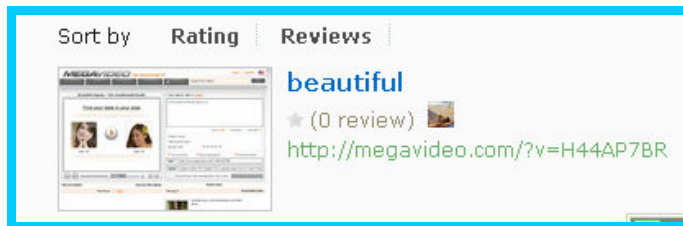
in het systeem sluipen. Het gaat hier om fouten zoals



Figuur 21: Spelfout in titel Digg.com

spellingsfouten (advertent), onnauwkeurigheden (inaccurate) of irrelevante tags (irrelevant). Een voorbeeld hiervan is hiernaast te zien. Hier is gezocht op het woord desing, in plaats van design.

Deze onduidelijkheden in metadata worden toepasselijk meta-ruis ofwel *metanoise* genoemd. (Bateman, Brooks & McCalla, 2006)

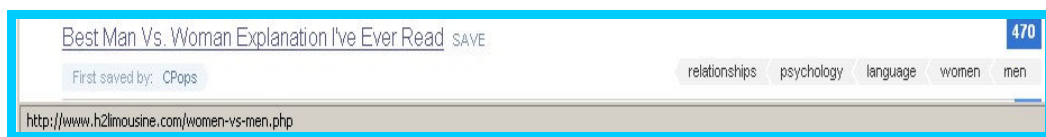


Figuur 22: Een voorbeeld van generieke metadata

ontcijferen wat de inhoud van de content is. In dit geval bleek het om een erotiek te gaan.

Hiernaast staat nog een van de vele voorbeelden waaraan er door gebruikers onduidelijke tag(s) zijn toegevoegd. Andere gebruikers kunnen aan de hand hiervan niet

Nog een risico waar een folksonomie gevoelig voor is, is het opzettelijk aanbrenen van onjuiste metadata, bijvoorbeeld voor commerciële doeleinden. Het bijgevoegde voorbeeld illustreert dit; een uitleg over mannen en vrouwen, met bijbehorende tags, die uiteindelijk verwijst (grijze balk onderaan) naar een website waar een commerciële partij limousines verhuurt in de Verenigde Staten.



Figuur 23: Voorbeeld discrepantie tussen verwijzing en beschrijving

Gebruikers van folksonomieën moeten zich actief opstellen om de door de dienst ontsloten content te vinden.

4.2.3 Conclusie

In de analyse wordt duidelijk dat de kracht van het op een folksonomische wijze indexeren van content ligt bij het waardeoordeel wat de gebruikers, wel of niet bewust, meegeven aan de content. Ook de extra dimensie die gegeven wordt aan het zoeken van content door middel van het bekijken van de voorkeuren en gevonden items van soortgelijke gebruikers, draagt hier sterk aan bij. Door deze kracht lijkt deze manier van het vindbaar maken van content met name geschikt voor het ondersteunen van internetgebruikers in het behalen van transactionele zoekdoelen. De voor- en nadelen van deze zoekmethode zijn in Tabel 4 nogmaals tegenover elkaar gezet. Ook is er in

Tabel 5 een inventarisatie gemaakt van de potentiële competenties van deze zoekmethode.

Tabel 4: Voor- en nadelen van werkwijze folksonomieën

Voordelen	Nadelen
Over het algemeen een relatief omvangrijke hoeveelheid ontsloten content.	Content moet eerst door gebruikers gevonden worden, wil het in de dienst opgenomen kunnen worden. Hierdoor ontstaat een vertraging in de indexering.
Goedkope manier van organiseren van content, dienst levert slechts een platform.	
	Onduidelijkheden en/of fouten in metadata, generieke metadata/polysemie.
Alle soorten inhoud van content kunnen geïndexeerd worden.	Misbruik mogelijk door opzettelijk verkeerde tags en titels toe te kennen aan content.
Flexibel in categorieën.	Dienst is passief.
Impliciet waardeoordeel internetgebruikers.	
Terugvinden/registreren van eerder content mogelijk.	
Intuitief zoekkarakter en beschrijvingen sluiten relatief dicht bij de belevingswereld gebruiker.	
Niet erg veel input vereist om tot een zoekresultaat te komen.	

Tabel 5: Waardering van de potentiële competentie van folksonomieën voor het behalen van zoekdoelen

Zoekdoel		Verwachte prestatie	Uitleg van verwachte prestatie
Navigationeel zoekdoel		-	De meeste folksonomieën geven geen URL als zoek resultaat, slechts links naar content met coca cola als onderwerp. Zoekopdracht naar Coca-Cola bij Delicious.com geeft links naar allerlei content, maar niet naar de website van Coca-Cola.
	<i>Navigate</i>	-	
	<i>Retrieval</i>	++	
Informationeel zoekdoel		+	Folksonomieën zijn hierin sterk omdat de gevonden content is beoordeeld door vaak gelijkgestemde gebruikers en hiermee in staat is om de content inhoudelijk te indexeren. Vooral op het aspect advice komt deze eigenschap goed naar voren. Een voorbeeld hiervan zijn de zeer relevante resultaten van de zoekopdracht 'photoshop tutorials'.
	<i>Directed</i>	+	
	<i>Undirected</i>	+	
	<i>Advice</i>	++	
	<i>Locate</i>	--	
Transactioneel zoekdoel		++	Bij transactionele content gaat het vaak om een waardeoordeel van de gebruikers. Folksonomieën zijn bij uitstek geschikt om deze waarde-oordelen te doorzoeken. De gebruiker hoeft maar naar stumbleupon.com te gaan en wordt direct ge-entertained met een relevante selectie content.
	<i>Download</i>	+	
	<i>Entertainment</i>	++	
	<i>Interact</i>	+	

4.3 Webdirectories

4.3.1 Diensten

Voor de analyse over de werking van toonaangevende webdirectories zijn de diensten Startpagina, Craigslist, Open directory project en Yahoo! Directory bekeken.

Startpagina

De eerste webdirectory die onder de loep genomen is, is Startpagina.



Figuur 24: Screenshot Startpagina.nl (Januari 2009)

De dienst is onderdeel van de Ilse media groep en trekt meer dan 4,5 miljoen bezoekers naar de hoofdpagina van de site. De webpagina van Startpagina bestaat uit een lijst van tientallen verschillende onderwerpen.

Onder de onderwerpen volgt een rij met links naar websites, die door de dienst wordt aanbevolen. Ook staat er bij elk

onderwerp een verwijzing naar een gespecialiseerde (start)pagina op dat gebied. Deze hebben vrijwel dezelfde lay-out en worden dochterpagina's genoemd. Zo kan iemand die iets zoekt over scripties terecht op scriptie.startpagina.nl. Er zijn meer dan 5000 van deze dochterpagina's²¹. Deze pagina's worden door vrijwilligers samengesteld en onderhouden. Het is dan ook aan deze gebruikers om de rangorde en selectie van de pagina te bepalen. Deze worden op hun beurt wel gecontroleerd door de dienst Startpagina zelf, om ervoor te zorgen dat er geen misbruik wordt gemaakt van de site en of de dochterpagina's wel goed onderhouden worden. Inhoudelijk gezien oefent de dienst niet of nauwelijks invloed uit op de inhoud van de pagina. Om de gebruikers tegemoet te komen voor hun inspanningen

²¹ N.a.v. telefonische contact Public Relations-afdeling Ilse media.

mogen deze maximaal vijftien commerciële links exploiteren op de pagina, deze worden typerend sponsoren genoemd. Hier komt dan meestal in rode letters het woord Tip! achter te staan, om de aandacht er op te vestigen. De gebruiker heeft ook de mogelijkheid om de pagina naar eigen voorkeur in te delen. Er kunnen onderwerpen verwijderd worden van de pagina, of verschoven worden. Op deze manier kan de gebruiker de startpagina personaliseren naar eigen gebruik. Opvallend detail is het verschil in commercie tussen de startpagina en de dochterpagina's. Waar de dochterpagina's slechts enkele commerciële links hebben, bestaat de voorpagina uit vrijwel uitsluitend commerciële links. De site is erg monotoon. Alle kaders hebben hetzelfde uiterlijk. Er wordt wel verschil gemaakt in blauwe kopjes boven aan de pagina, en rode onderaan. Waarom dit is, is onduidelijk.

Craigslist

De dienst Craigslist is een vreemde eend in de bijt.

Deze dienst biedt gebruikers de mogelijkheid om simpele en korte tekstberichten te plaatsen over een grote diversiteit aan onderwerpen, eventueel in combinatie met een afbeelding. De site trekt hiermee mee vijftig miljoen gebruikers per dag en is goed voor twaalf miljard pageviews per maand²². De site staat hiermee nummer acht in de



Figuur 25: Screenshot Craigslist.com (Januari 2009)

top vijftig van meest bezochte Engelstalige sites ter wereld. De site is erg minimaal in zijn voorkomen. De vormgeving beperkt zich tot kale tekst met dunne omkadering per onderwerpgroep.

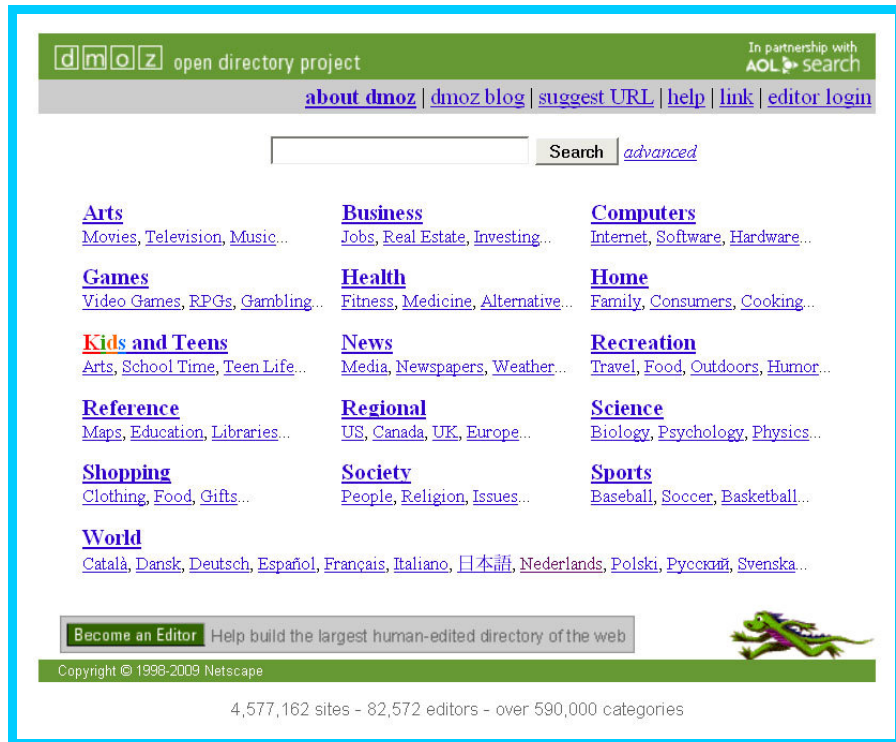
²² Factsheet Craigslist. Geraadpleegd 8 Juli 2008 <http://www.craigslist.org/about/factsheet>

DMOZ – Open directory project.

Het Open Directory Project is de meest omvangrijke, door gebruikers samengestelde, directory op internet.

De dienst heeft momenteel meer dan 4,5 miljoen sites geïndexeerd verdeeld over zo'n 590 duizend categorieën.

De index is samengesteld door bijna 90 duizend gebruikers. Deze site is qua voorkomen vergelijkbaar met de andere directories. De site is erg eenvoudig en de vormgeving bestaat slechts uit blauwe tekst. Toch heeft ook deze dienst een karakteristieke



Figuur 26: Screenshot DMOZ (Januari 2009)

eigenschap. Men kan kiezen om binnen alle talen te zoeken, maar er kunnen ook een of meerdere talen geselecteerd worden waarbinnen er gezocht moet worden naar content. De selectie van talen is opvallend uitgebreid; er kan door gebruikers uit meer dan honderd talen gekozen worden. Bij andere vergelijkbare diensten beperkt zich dit veelal slechts tot één enkele taal.

De categorieën van de dienst staan op alfabetische wijze georganiseerd en zijn ook te doorzoeken aan de hand van een zoekopdracht. Achter de naam van de categorieën en subcategorieën staat ook het aantal items dat binnen deze categorie te vinden is.

Yahoo! Directory

De laatste dienst van deze methode die onder de loep is genomen is de Yahoo! Directory.

Deze directory is de eerste grote directory die op internet samengesteld werd. De omvang van het aantal items benadert die van the open directory project.

In eerste instantie werden de diensten door gebruikers aangeraden en door

medewerkers gescreend. Hierin werden ze ondersteund door speciale software. Yahoo! Maakte na enkele jaren de overstap naar een zoekmachine die met een crawler werkt.

Hiermee kreeg de directory een nieuwe functie, namelijk het vindbaar maken van websites op de zoekpagina. Naar aanleiding van deze nieuwe functie kreeg de directory een, voornamelijk, commercieel karakter. Het gevolg hiervan is dat het tegenwoordig geld kost om opgenomen te worden in de directory. Er moet rond de 300 dollar per jaar betaald worden voor een listing in de index. Voor een site met 18+ content moet er 600 dollar per jaar betaald worden.

De onderwerpen van ook deze directory kunnen handmatig en aan de hand van een zoekmachine doorzocht worden. Wanneer er handmatig door de directory gezocht wordt, kan er gesorteerd worden op onderwerp en op locatie. Onderaan de webpagina worden dan gedurende het zoeken de meest populaire links binnen de categorie waarin gezocht wordt getoond.

De site is qua voorkomen een stuk drukker dan de eerderbehandelde directories. Dit komt voornamelijk door het grote artikel dat het grootste deel van de pagina vult. Het artikel is afkomstig van *the spark of Yahoo*, en wisselt per dag. Er is geen informatie beschikbaar over het motief van het plaatsen van het artikel en/of de onderwerpkeuze ervan.



Figuur 27: Screenshot Yahoo.com (Januari 2009)

4.3.2 Methode

Wanneer een webdirectory degelijk en specialistisch samengesteld is, kan dit een zeer efficiënte manier zijn voor gebruikers om binnen een korte tijd een hoop content te vinden. Dit specialisme is echter relatief gezien erg kostbaar. De manier waarop de gebruikers categorieën kunnen zoeken is een methode die lang voor de komst van internet al bestond. Denk bijvoorbeeld aan de indeling van een bibliotheek of de Gouden Gids. Dit zal het zeker voor een minder ervaren internetgebruiker gemakkelijker maken om deze werkwijze te doorgronden.

Ook zijn dit soort directories een goede startplaats voor gebruikers om te beginnen met de sneeuwbalmethode (zie H4). Doordat (meer) ervaren gebruikers al aan de hand van een kwaliteits-of relevantieoordeel een selectie hebben gemaakt, zijn in dit soort directories hoofdzakelijk de links naar toonaangevende websites en content over een bepaald onderwerp geselecteerd en geplaatst. In directories is in de meeste gevallen zowel tekstuele content als audio-visuele content geïndexeerd. Wederom is dit het resultaat van menselijke tussenkomst in de selectie ervan.

Toch zijn er een aantal kanttekeningen te plaatsen bij het op deze wijze zichtbaar maken van beschikbare webcontent. Een van de kritieken, die eerder reeds besproken werd bij een taxonomieën in het algemeen, is de rigiditeit van dit systeem. Wanneer een nieuw, ongebruikelijk of zelfs uniek onderwerp ontstaat kan het erg ingewikkeld zijn te categoriseren op een wijze dat het voor een gebruiker terug te vinden is.

Het is ook niet ongebruikelijk dat bij dit soort diensten met enige regelmaat onenigheid ontstaat tussen de categoriseerders over de wijze waarop de links geordend dienen te worden.

Een ander punt van kritiek heeft te maken met de vertraging, de tijd die ligt tussen het verzoek tot plaatsing van content en de daadwerkelijk plaatsing ervan in de directory. Dit heeft te maken met de onmisbare tussenkomst van mensen om de ingediende content te screenen, te selecteren en te plaatsen. De

reviewed and the Yahoo! editorial team will respond to you within 7 business days, for a review fee of only \$299 (\$600 for adult

content moet gescreend worden om de directories schoon te houden van spam en andere

ongewenste content. De meerwaarde die hiertegenover staat is dat hierdoor de directories over het algemeen vrijwel vrij van verstoringen, spam etcetera zijn. Door deze manier van werken ontstaat een vertraging tussen de aanbieder en de daadwerkelijke plaatsing. Zeker als er gedacht wordt aan bijvoorbeeld nieuws-gerelateerde content is dit wachten op plaatsing relatief lang.

Zoals eerder beschreven werken veel directories voor een gedeelte soms ook met enkele commerciële links in de directory. Deze betaalde toevoegingen zullen niet altijd optimaal zijn voor de wensen van de gebruiker. De motivatie achter de plaatsing ervan wordt veelal niet gedeeld door de zoekende gebruiker. Wel wordt er meestal aan de gebruikers kenbaar gemaakt wanneer het een gesponsorde link betreft.

Webdirectories kunnen ook onoverzichtelijk worden als er heel veel content geplaatst moet worden en er een grote diversiteit binnen deze content blijkt te bestaan.

Als we als voorbeeld startpagina nemen wordt dit duidelijk. De hoofdpagina van deze dienst bevat honderden links. Dat zijn op een computermonitor²³ vijf schermen vol met links. Deze staan wel per categorie verdeeld, alleen de structuur van de categorieën is niet eenduidig. De categorieën zijn niet gealfabetiseerd en ook kan er geen logica ontdekt worden in samenhangende categorieën. Zo staan *banken* bijvoorbeeld links aan het begin van de pagina, *hypotheken* rechts in het midden en *zorgverzekeraars* helemaal midden onderaan de pagina. De categorie *Sinterklaas* moet het doen met een plaats tussen *kredietcrisis* en *woningen*.

Ook kan de uitgebreidheid van een directory tot gevolg hebben dat er veel handelingen verricht moeten worden om tot de gewenste content te komen. Wanneer er bij de Open Directory Project gezocht wordt naar the Open Directory Project moet men, enigszins ironisch, meer dan zeven categorieën doorklikken om deze te vinden. Tenminste, als men direct de goede categorieën weet te selecteren. Een actieve opstelling van de gebruiker is bij een webdirectory vereist om tot een resultaat te komen.

4.3.3 Conclusie

Uit de analyse komt naar voren dat directories voornamelijk de meest toonaangevende content zichtbaar maken voor gebruikers. Het biedt veelal een solide startpunt om te

²³ Er is hier uitgegaan van een monitor van 19 inch.

beginnen wanneer een internetgebruiker, met nog weinig kennis over een onderwerp, zich wil verdiepen in een bepaald onderwerp.

Voor het vinden van webadressen is deze methode minder geschikt, het zou relatief veel tijd en handelingen kosten om, mogelijk, een webadres te vinden bij dit type dienst.

Een deel van de webdirectories op internet wordt aan de hand van menselijk inzicht samengesteld. Deze directories hebben door deze menselijke tussenkomst dan ook een verhoogde waarde wanneer het gaat om het toevoegen van een oordeel over de geplaatste content.

Door de besproken voor en nadelen van de methode moet geconcludeerd worden dat deze methode niet erg geschikt is voor het navigatieve zoekdoel. Het karakter van de methode leent zeer voornamelijk voor het bereiken van het informatiele zoekdoel en, in iets mindere mate, het transactionele zoekdoel. In

Tabel 6 en Tabel 7 komen voor- en nadelen en de waardering van hun werkwijze nogmaals overzichtelijk naar voren.

Tabel 6: Voor- en nadelen van werkwijze webdirectories

Voordelen	Nadelen
Redelijke omvangrijkheid ontsloten content.	Relatief lang en ingewikkeld proces tussen indexering van content en de plaatsing ervan met vertraging als gevolg.
Alle soorten content en inhoud kunnen worden ontsloten.	Hoge kosten voor contentontsluiting, dwz arbeidsintensief
Webdirectories hebben over het algemeen weinig last van verstoringen of beïnvloeding van geïndexeerde content door gatekeeperfunctie samenstellers.	Het meest rigide systeem van de vier behandelde systemen. Nieuwe, ongebruikelijke of unieke content kan er voor zorgen dat heel het systeem aangepast moet worden.
Expliciet waarde-oordeel geplaatste content door gemaakte (menselijke) selectie.	Geen specifieke wijze voor het terugvinden van eerder gevonden content.
Selectie van content door samenstellers leidt over het algemeen tot	De dienst is passief in het aanbieden van de geïndexeerde content.

relevante/toonaangevende content.	
Intuïtieve wijze van het ordenen en zoeken naar content. Werkwijze reeds lange tijd door mensen gebruikt, zowel off- als online.	Relatief veel handelingen nodig om tot de gewenste content te komen. (Voornamelijk bij omvangrijk systeem)

Tabel 7: Waardering van de potentiële competentie van webdirectories voor het behalen van zoekdoelen

Zoekdoel		Verwachte prestatie	Uitleg van verwachte prestatie
Navigatieve zoekdoel		+/-	Webdirectories zijn erop gericht om richting juiste onderwerpen content te verwijzen, en minder naar een vooraf gedefinieerde locatie. De website van TNO is niet zo gemakkelijk te vinden via Startpagina.nl, maar informatie over ICT wel.
	<i>Navigate</i>	+/-	
	<i>Retrieval</i>	--	
Informatieve zoekdoel		+	Webdirectories zijn goed in verwijzen naar bepaalde categorieën content. Wanneer het informatieve zoekdoel daarbij aansluit, scoort deze methode goed (e.g. overzicht van alle kranten in Nederland via Startpagina.nl). Echter wanneer de zoekopdracht specifiek is, kan het lastiger zijn om de juiste content te vinden (e.g. geboortjaar van Napoleon).
	<i>Directed</i>	+	
	<i>Undirected</i>	+	
	<i>Advice</i>	+	
	<i>Locate</i>	+/-	
Transactieve zoekdoel		+/-	De webdirectories kunnen helpen bij het vinden van de plek waar de content wordt aangeboden, maar bieden dit (meestal) niet zelf aan. Voorbeeld hiervan is dat wel de betere locaties van online games gevonden kunnen worden via webdirectories, maar niet de games zelf.
	<i>Download</i>	+	
	<i>Entertainment</i>	++	
	<i>Interact</i>	+	

4.4 Weblogs

4.4.1 Diensten

In deze categorie zijn de diensten Nu.nl en Spitsnieuws.nl geanalyseerd.

Nu.nl

Nu.nl, eigendom van Ilse Media, is een dienst die zich voornamelijk richt op het aggregeren van nieuws-gerelateerde content.

Op de pagina van de dienst is er een overzicht van de laatste items te zien, verdeeld per categorie. De categorieën variëren van de meer reguliere onderwerpen zoals algemeen, economie, sport, wetenschap en beurs tot categorieën als lifehacking, games, plugged en lifestyle.

De items worden chronologisch toegevoegd aan de categorie, zoals met de meeste weblogs het geval is.

De items en de artikelen die bij de geplaatste content worden geschreven door een vaste redactie, bestaande uit een vijftigtal betaalde medewerkers. In de meeste gevallen is er bij

de artikelen een afbeelding te zien over het onderwerp en is ook een link beschikbaar naar de originele bron van de content. In sommige gevallen is er ook een korte video beschikbaar waarin er een toelichting wordt gegeven over het artikel.

De items kunnen per categorie doorzocht worden. Hier zijn de meest recente vijftig items te zien. Oudere items, beginnend vanaf januari 2004, kunnen worden teruggezocht door middel van een zoekfunctie.



Figuur 28: Screenshot Nu.nl (januari 2009)

Spitsnieuws

Spitsnieuws, een dienst van NewsMedia, is een site die in grote lijnen vergelijkbaar is met die van Nu.nl.

Ook bij deze dienst zijn de laatste nieuws-gerelateerde items op chronologische wijze geordend en per categorie verdeeld. Op de voorpagina van de dienst kunnen de laatste en belangrijkste headlines van het moment gelezen worden. Wanneer er doorgedrukt wordt op een bepaalde categorie verschijnen de laatste items betreffende het gekozen onderwerp. Aan de rechterzijde van de startpagina is onder andere een lijst te zien met links naar die artikelen die het meest gelezen zijn.

Een verschil met Nu.nl is dat bij deze dienst de gebruikers informatie toe kunnen voegen aan items, door middel van commentaar onder het artikel. In deze reacties kunnen hoofdzakelijk meningen over het betreffende item worden gevonden, maar ook toevoegingen, correcties en links naar andere content die op een of andere wijze verband houden met het onderwerp worden hierin geplaatst. Ook is te zien hoeveel reacties er reeds zijn op het artikel aan het getal tussen haakjes achter de titel, bijvoorbeeld (134), oftewel 134 reacties.

Het aantal categorieën waar de content in is onderverdeeld is aanzienlijk kleiner dan bij Nu.nl en er ligt een grotere nadruk op wat luchtiger items zoals bijvoorbeeld met de categorie Raar. Opvallend is dat het taalgebruik binnen de dienst een stuk populairder en minder formeel is dan bij de andere diensten van dit onderzoek. Onduidelijk is of de artikelen geplaatst zijn door een professionele- of amateur-redactie of een combinatie van de twee.



Figuur 29: Screenshot Spitsnieuws.nl (januari 2009)

4.4.2 Methode

De besproken weblogs hebben veel zaken gemeen.

De methode die de diensten hanteren om gebruikers in contact te brengen met gezochte content heeft een aantal duidelijke voordelen. Zo krijgen gebruikers in een oogopslag de meeste recente berichten over een breed scala aan onderwerpen te zien. Op deze manier kunnen ze gemakkelijk op de hoogte blijven van ontwikkelingen op een bepaald gebied. Ook krijgen ze vaak een multimediaal aanbod over het onderwerp te zien, vaak bestaande uit tekst met een beschrijving of oordeel in combinatie met een afbeelding of andere audiovisueel element zoals een filmpje. Doordat de content door de een selectieve groep betrokken bloggers wordt geplaatst, is de mate van verstoring ervan erg laag.

Een extra optie die deze diensten aan hun gebruikers verschaffen is het gebruik van RSS. Deze manier van programmeren stelt gebruikers in staat om software te installeren waar automatisch de laatst geplaatste items van de dienst worden gedownload. Op deze wijze hoeven de gebruikers niet eens de site van de dienst te bezoeken, maar krijgen ze deze automatisch te zien met speciale software. In deze zogeheten RSS-reader kunnen zijn van een groot aantal diensten de laatste berichten ontvangen.

Door de chronologische volgorde zijn gebruikers altijd in contact met de meest recente informatie. Nu staat de meeste recente informatie lang niet altijd gelijk aan de meest relevante. Dit kan als gevolg hebben dat voor een gebruiker relevante content ondergesneeuwd kan raken in de grote hoeveelheid nieuwe items. Hierdoor is het vaak lastig om eerdergevonden content terug te vinden. Nu kan een gebruiker natuurlijk wel de geplaatste content doorzoeken op onderwerp, maar dit is lang niet altijd toereikend. Een fictief voorbeeld ter illustratie: een gebruiker zal niet middels de zoekfunctie op zoek gaan naar informatie over een grote brand op de Erasmus Universiteit in Rotterdam, als deze niet op de hoogte is van de gebeurtenis. Hij of zij zal deze informatie alleen meekrijgen via de weblogs wanneer items over deze gebeurtenis voorkomen in de meest recente onderwerpen. Hoe lang dit duurt is afhankelijk van de omloopsnelheid van items binnen de dienst.

Weblogs zijn vaak gespecialiseerd op een nieuws of een thema. Hierdoor wordt niet alle beschikbare internetcontent geïndexeerd, omdat deze niet aansluit bij de nieuwswaardigheid of thema van de weblog. Dit plaatsen van content gebeurt door

personen, dit maakt het onderhouden van een weblog een relatief intensief en kostbaar proces.

4.4.3 Conclusie

De weblogs zijn over het algemeen sterk georiënteerd op de nieuw(st)e content over een specifiek onderwerp. Dit kan in sommige situaties gelijk zijn aan de meest relevante content, maar dit is niet altijd het geval.

Het zoeken naar een webadres is met deze methode vrijwel onmogelijk; er kan door gebruikers niet geanticipeerd worden de aanwezigheid van hun zeer specifieke zoekdoel. De marginale kans dat een weblog exact dit adres heeft opgenomen is klein.

De beheerders van een weblog voegen over het algemeen vaak zelf nog een extra toelichting of informatie toe aan de geïndexeerde of geplaatste content.

De gebruiker van de dienst kan op twee manieren content vinden bij deze methode.

Er kan gekozen worden aan de hand van een zoekvenster te zoeken binnen content die de dienst aanbiedt, maar er kan ook gekozen worden de content volgens het chronologische principe van de plaatsing te consumeren.

Dit maakt het bij deze diensten mogelijk om de artikelen op een bijna recreatieve wijze te bekijken.

De combinatie van de verschillende karakteristieken van de methoden maakt diensten die deze methode hanteren relatief sterk op het gebied van informatiele zoekdoelen en het gebied van transactionele zoekdoelen. Voor navigationele zoekdoelen is deze methode ongeschikt. Onderstaand hiervan een overzicht.

Tabel 8: Voor- en nadelen werkwijze weblogs

Voordelen	Nadelen
Weblogs kunnen alle soorten content indexeren.	Relatief weinig content geïndexeerd
Meestal gefocussed op recente ontwikkelingen en nieuws, hierdoor vaak snel met plaatsen nieuwe content.	Kosten relatief hoog, een kleine groep mensen moet veel doen.

Vrijwel geen verstoring vanwege de menselijke selectie van content en omschrijvingen rond content.	Niet erg flexibel vanwege vaste chronologische werkwijze.
Vaak is er een omschrijving met daarin een waardeoordeel over de content aanwezig.	Het terugvinden van eerdergevonden content is soms lastig vanwege de grote stroom nieuwe content die deze bedelft.
De content wordt op een actieve wijze aangeboden aan de gebruiker.	
De chronologische werkwijze is intuïtief gezien erg makkelijk te begrijpen en gebruiken.	
De gebruiker hoeft geen vrijwel geen input te geven om tot een zoekresultaat te komen.	

Tabel 9: Waardering van de potentiële competentie van weblogs voor het behalen van zoekdoelen

Zoekdoel	Verwachte prestatie	Uitleg van verwachte prestatie
Navigationeel zoekdoel	--	Weblogs zijn erop gericht om content aan te bieden op hun eigen website, en niet op verwijzen naar een vooraf gedefinieerde locatie.
<i>Navigate</i>	--	
<i>Retrieval</i>	--	
Informationeel zoekdoel	+	Wanneer relevantie van informatie bepaald wordt door actualiteit kan op weblogs effectief en gemakkelijk de juiste content gevonden worden. Als een gebruiker bijvoorbeeld op zoek is naar de laatste economische ontwikkelingen, zijn die direct te vinden onder het kopje economie bij Nu.nl.
<i>Directed</i>	+	
<i>Undirected</i>	++	
<i>Advice</i>	+	
<i>Locate</i>	-	
Transactioneel zoekdoel	++	Weblogs bieden content aan, die zijn voorzien van een subjectieve context welke een meerwaarde geeft voor de gebruiker. Bijvoorbeeld: Spitsnieuws.nl biedt naast bij nieuwsartikelen ook de reacties van gebruikers aan.
<i>Download</i>	++	
<i>Entertainment</i>	++	
<i>Interact</i>	++	

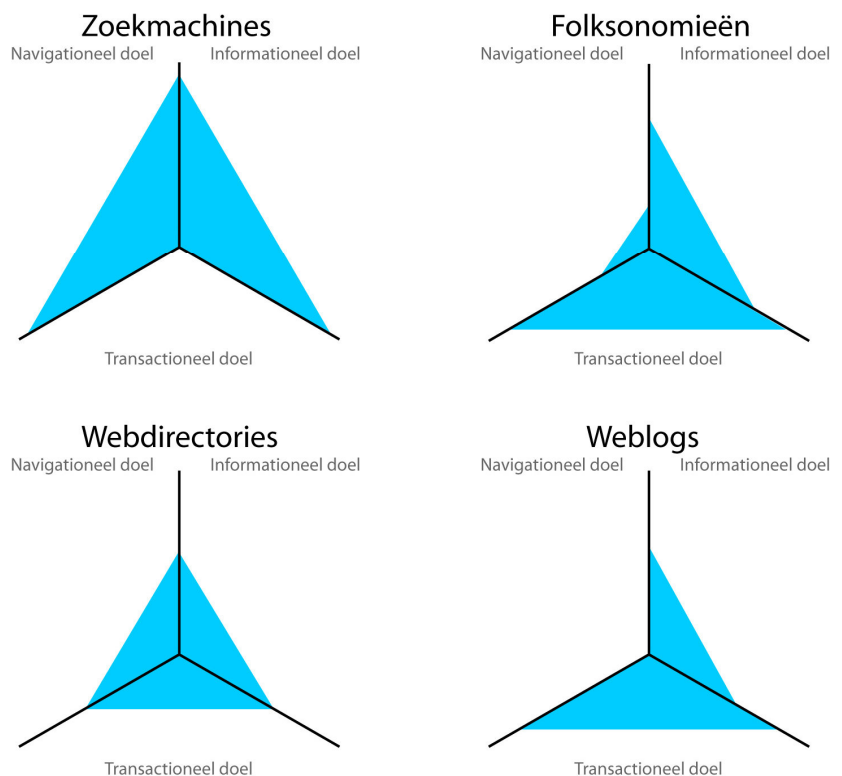
4.5 Conclusie dienstenanalyse

Johan Cruyff sprak ooit in zijn eigen, enigszins onbegrijpelijke, wijsheid: *Elk voordeel hep z'n nadeel!*

Voor de diensten die voor deze analyse beschouwd zijn blijkt dit ook op te gaan. Wanneer de methoden onder de loep genomen worden, blijkt dat ze allen hun eigen voordelen en sterkten hebben wat betreft het ontsluiten van content voor internetgebruikers.

Enkele voorbeelden hiervan zijn: dat gebruikers bij folksonomieën *alle* soorten en typen content kunnen begrijpen en beschrijven, de ongekende snelheid waarmee context door zoekmachines geïndexeerd kan worden, de natuurlijke wijze waarop de gebruiker door een webdirectory kan zoeken en de aanwezigheid van context rond de content bij weblogs. Deze krachten hebben ook hun keerzijde. Elke methode die besproken is, heeft ook tekortkomingen wat betreft het ondersteunen van gebruikers in het behalen van hun

zoekdoel(en). Zo begrijpen de softwareprogramma's van zoekmachines niet wat er op afbeeldingen of video's te zien is, zo zijn folksonomieën gevoelig voor menselijke fouten, zo worden webdirectories erg omvangrijk zodat het moeilijk is om het overzicht te erin bewaren en de wildgroei van de categorieën te voorkomen en zo weegt bij weblogs de chronologie van de gevonden content zwaarder dan (potentiële) relevantie.



Figuur 30: Overzicht van potentiële krachten van ontsluitingsmethoden

In Figuur 30 zijn deze krachten en zwakten van de verschillende zoekmethoden gevisualiseerd aan de hand van een diagram.

5. Resultaten

Gebruikersonderzoek

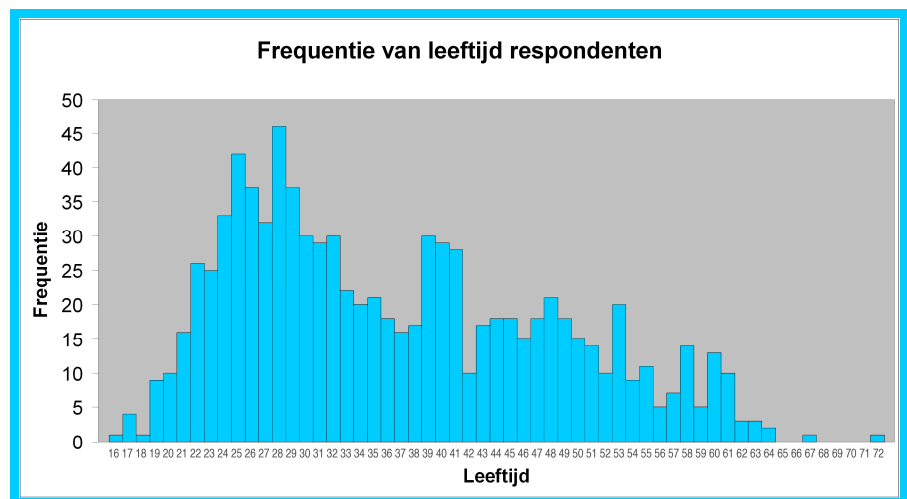
5.1 Populatie

In dit hoofdstuk worden de resultaten van het gebruikersonderzoek behandeld.

De resultaten die in dit hoofdstuk besproken worden zijn afkomstig uit de online survey.

De populatie van het onderzoek bestaat uit 1066 respondenten, waarvan 896 de volledige enquête hebben afgerond. Dit is een percentage van 84,1 procent.

De leeftijd binnen deze groep internetgebruikers loopt van 16 tot en met 72 jaar. De gemiddelde leeftijd van de respondenten is 36 jaar, de leeftijd 28 was de meest voorkomende (45 keer). De mannen waren gemiddeld iets ouder dan de vrouwen, met 38 versus 34 jaar.

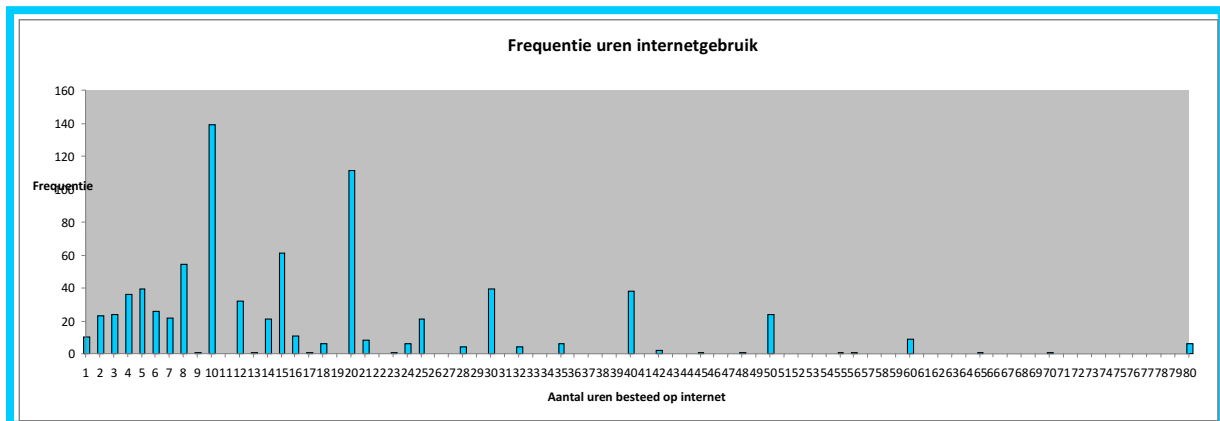


Figuur 31: Frequentie van leeftijd respondenten

Zoals al eerder in dit onderzoek is vermeld, is de verhouding tussen mannen en vrouwen die de survey hebben ingevuld respectievelijk 59 en 41 procent.

Het aantal uren dat de groep respondenten per week op internet besteedt, zowel werk als privé, ligt rond de 16,7 uur. In onderstaande grafiek is duidelijk te zien dat de respondenten de neiging hebben gehad dit antwoord af te ronden op tientallen, de staven van tien en twintig uur tijdsbesteding aan internet steken er dan ook met kop en schouders bovenuit.

De mannen uit het onderzoek besteden gemiddeld iets meer tijd op internet per week, 17,5 ten opzichte van de vrouwen met 15,6 uur.



Figuur 32: Frequentie van gemiddeld aantal uren internetgebruik per week

In de survey is de mening van de respondenten gevraagd over het zoeken en vinden van informatie op internet, om zo een globaal beeld te krijgen over hoe gebruikers hiertegenover staan. Om dit beeld verder uit te diepen is er, aan de hand van een open vraag, aan de respondenten gevraagd om zijn of haar zoekwijze kort te beschrijven.

5.2 Stellingen

De respondenten geven de indruk redelijk positief te zijn over het zoeken en vinden van informatie op internet.

Rond de vijfenveertig procent geeft aan het eens of redelijk mee eens te zijn dat zoeken op internet gemakkelijk is. Meer dan de helft van de respondenten is het er in meerdere mate mee eens dat als ze iets zoeken op internet, ze dit altijd kunnen vinden.

De stellingen waarin problemen met het vinden van informatie (stelling 5 en 8) worden voorgelegd, worden door de internetgebruikers grotendeels als onjuist betiteld.

Ongeveer een zesde van de respondenten geeft aan de manier *waarop* ze zoeken niet of nauwelijks te laten afhangen van *wat* ze zoeken. Ruim zestig procent geeft hiermee aan meestal of altijd hun werkwijze aan te passen aan hetgene wat zij op internet zoeken.

Enigszins naar verwachting scoort Google ook bij deze stellingen hoog. Ruim vijfenzestig procent van de gebruikers geeft aan het redelijk of volledig eens te zijn met de stelling dat ze

met Google alles kunnen vinden wat ze nodig hebben (stelling 6). Nog geen vijf procent is van mening dat zijn of haar omgeving beter weet waar informatie te vinden is dan Google. Hieronder een overzicht van de respons.

Tabel 10: Resultaten reactie respondenten op voorgelegde stellingen

8. In hoeverre ben jij het eens met de volgende stellingen? (-- oneens/ ++ eens						Response Count
--	-	-/+	+	++		
Als ik iets zoek op internet, lukt het me altijd om dit te vinden	2.7% (24)	15.4% (139)	27.6% (250)	45.6% (413)	8.7% (79)	905
Er staat erg veel informatie op internet, maar er in zoeken is gemakkelijk.	2.0% (18)	13.7% (124)	39.0% (353)	38.8% (351)	6.5% (59)	905
Als ik iets zoek op internet dan ga ik altijd hetzelfde te werk.	2.4% (22)	12.4% (112)	24.6% (223)	46.6% (422)	13.9% (126)	905
Ik laat de manier waarop ik zoek afhangen van wat ik zoek.	4.2% (38)	12.7% (115)	23.2% (210)	41.8% (378)	18.1% (164)	905
Er staat te veel informatie op internet, hierdoor kan ik vaak niet meer vinden wat ik zoek.	15.9% (144)	36.0% (326)	31.5% (285)	14.0% (127)	2.5% (23)	905
Ik kan met Google alles vinden wat ik nodig heb.	2.5% (23)	11.0% (100)	20.4% (185)	49.1% (444)	16.9% (153)	905
Mijn omgeving (vrienden/collega's etc.) kunnen mij beter vertellen waar ik iets op internet kan vinden dan google.	26.3% (238)	48.7% (441)	20.4% (185)	4.1% (37)	0.4% (4)	905
Als ik iets zoek op internet, kost het me veel tijd om het te vinden.	21.8% (197)	43.2% (391)	26.0% (235)	8.4% (76)	0.7% (6)	905
answered question						905

5.3 Open vraag

De open vraag “Hoe vind jij je weg door de grote hoeveelheid informatie?” werd in totaal door zeventienhonderdveertig mensen beantwoord.

Wanneer men de respons op deze vraag bekijkt, komen enkele kenmerken meer dan regelmatig naar voren.

Een zeer groot gedeelte van de respondenten laat door middel van deze open vraag weten dat zij hun succes voornamelijk bepalen, of proberen te behalen, door middel van de keuze van de juiste zoektermen.

5.4 Naamsbekendheid en gebruiksfrequentie

5.4.1 Diensten

Er is binnen de survey ook gekeken naar hoe het gesteld was met de naamsbekendheid van de geselecteerde diensten en het gebruik ervan.

Zonder dat de respondenten hiervan op de hoogte zijn gesteld, waren zeventien diensten die in een lijst stonden, onzichtbaar onderverdeeld in vier categorieën, namelijk die van zoekmachines, folksonomieën, webdirectories en weblogs. Dit maakt het mogelijk een vergelijking te trekken tussen het gebruik van de verschillende zoekmethoden. In Tabel 11 is een overzicht te zien van de respons op deze vraag.

Tabel 11: Naamsbekendheid zoekdiensten onder respondenten

Google	99,8%	Spitsnieuws	44,0%
Startpagina	97,8%	Ask	41,5%
Yahoo!	95,1%	Dmoz/ODP/Google directory	29,8%
Ilse	91,9%	Nujij	25,1%
Altavista	88,7%	Digg	22,7%
Nu	88,2%	Delicious	22,3%
Lycos	85,2%	Stumbleupon	16,9%
Live!	78,4%	Craigslist	15,1%
Yahoo! Directory	58,3%		

Niet geheel onverwacht komt Google als bekendste dienst uit de bus.

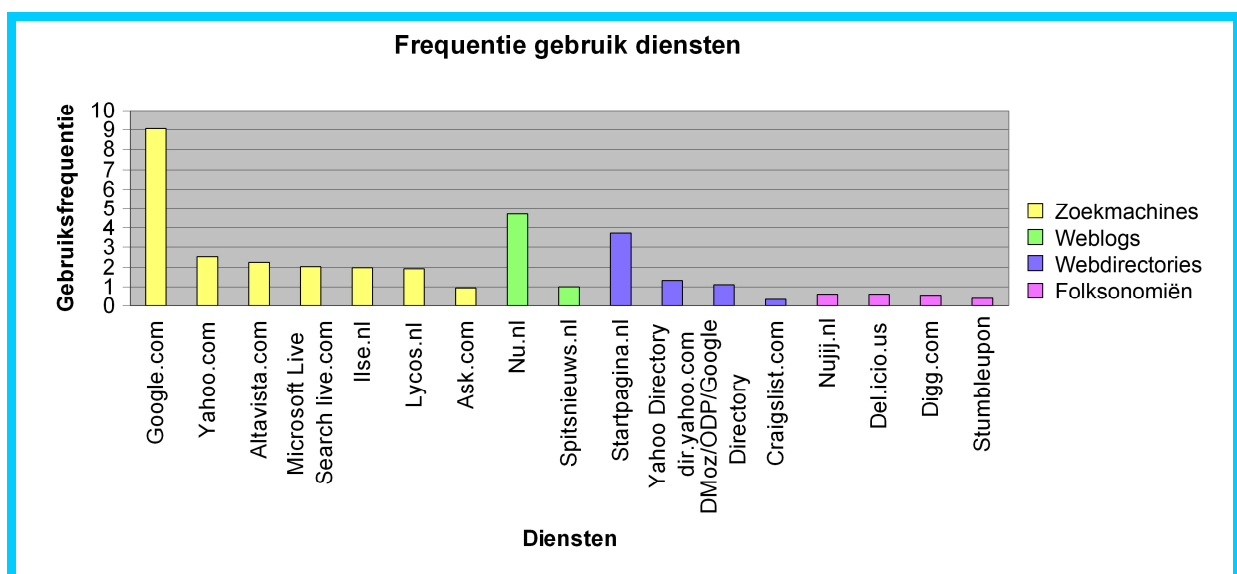
De dienst geniet een naamsbekendheid van 99,8 procent onder de van de respondenten.

Wat betreft naamsbekendheid wordt Google gevolgd door Startpagina met een naamsbekendheid van 97,8 procent en Yahoo! met 95,1 procent. Het minst bekend bij de respondenten zijn de diensten Delicious(22,3 procent), Stumbleupon (16,9 procent) en Craigslist (15,1 procent).

De respondenten is, naast de naamsbekendheid, ook gevraagd naar de regelmaat waarmee zij diensten uit de diensten uit de lijst gebruiken. De schaalverdeling waarmee geantwoord kon worden bestaat uit zes antwoordmogelijkheden, beginnend bij *nooit van gehoord*

oplopend tot *gebruik ik altijd*. Aan deze zes mogelijke antwoorden is een weging (0,2,4, 6,8 of 10 punten) toegekend en deze zijn vervolgens omgezet naar een gemiddelde per *dienst* én per *methode*.

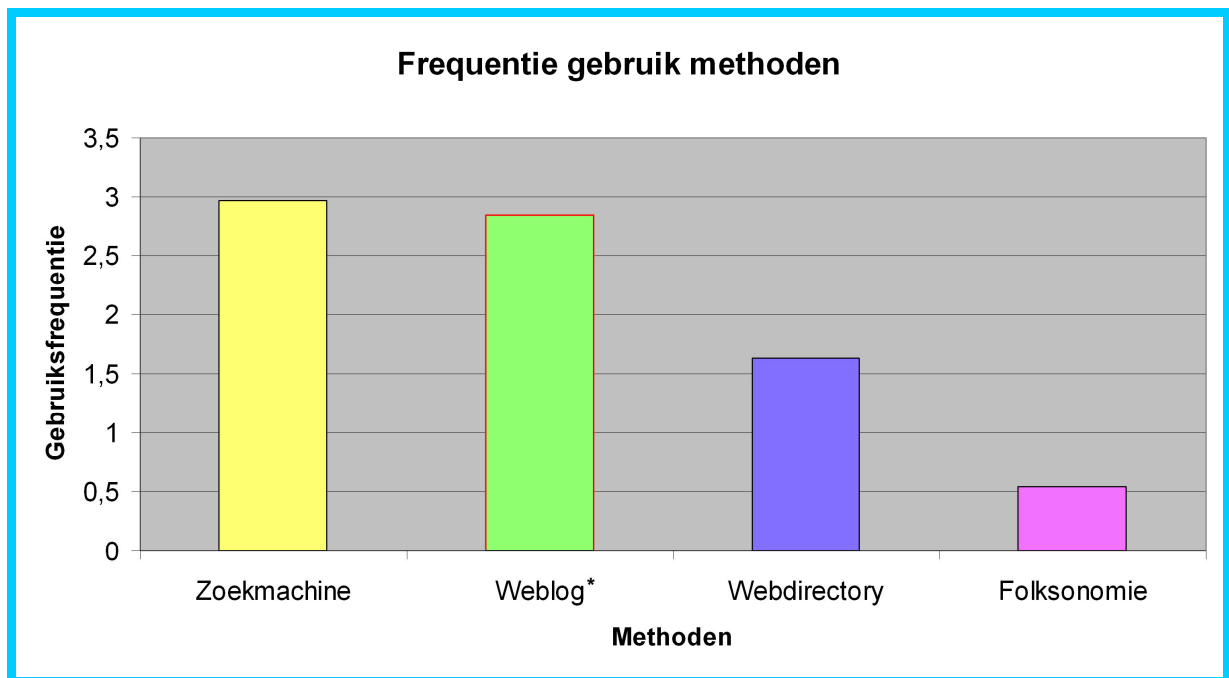
De uitkomsten over de regelmaat in het gebruik van de diensten scheidt een duidelijk beeld: Google laat zien de concurrentie ver achter zich te laten. Nu.nl en Startpagina zijn de enige twee diensten die nog enigszins overtuigend een plaats weten te behouden naast Google. In vergelijking met deze drie diensten scoort de rest significant lager qua regelmaat van gebruik.



Figuur 34: Frequentie van gebruik zoekdiensten

5.4.2 Methoden

Wanneer er gekeken wordt naar de vergelijking tussen verschillende zoekmethoden komt naar voren dat zoekmachines gemiddeld gezien populairder zijn in het gebruik dan de andere zoekmethoden. In Figuur 35 is dit duidelijk zichtbaar. De weblogs komen op het eerste gezicht qua frequentie van gebruik dicht in de buurt van de zoekmachines. Dit beeld moet echter bijgesteld worden vanwege het feit dat er slechts een klein aantal weblogs, die aan de opgestelde vereisten voldeden, toonaangevend genoeg waren om in dit onderzoek opgenomen te worden. Het beeld wordt hierdoor sterk vertekend door de populariteit van Nu.nl. Wanneer gekeken wordt naar de andere uitkomsten die in dit onderzoek naar voren komen is de verwachting dat weblogs gemiddeld gezien qua gebruik tussen de webdirectories en folksonomieën in zullen zitten.

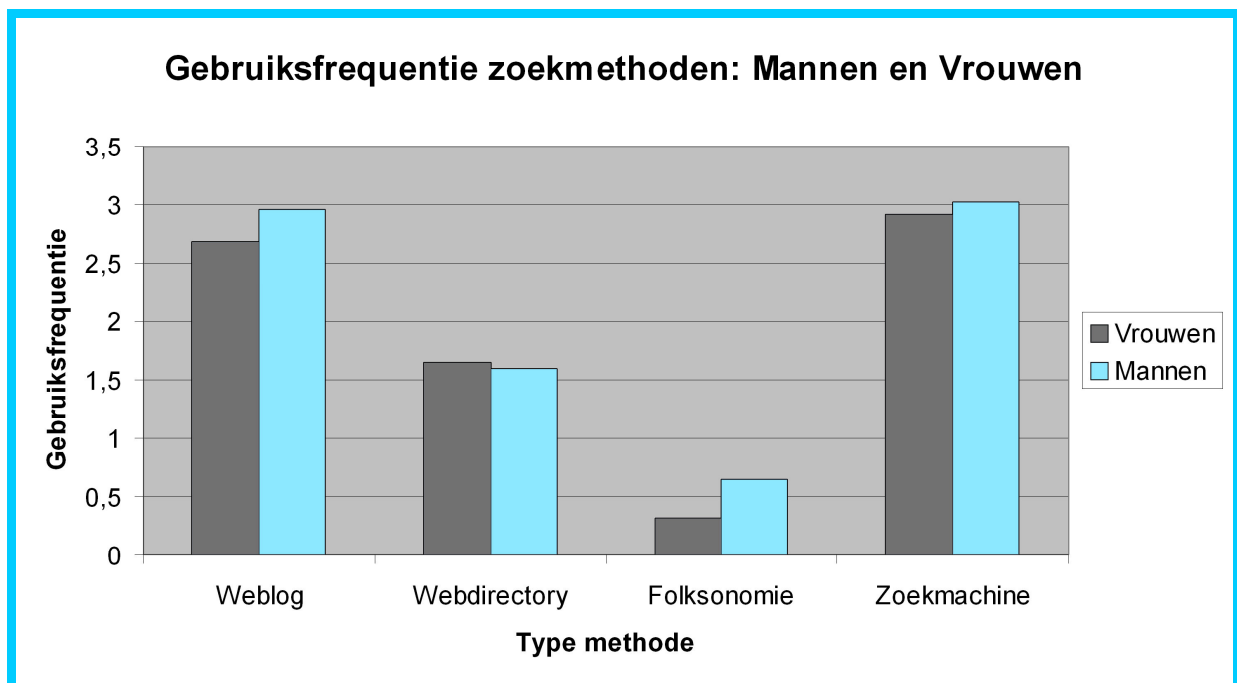


Figuur 35: Gebruiksfrequentie zoekmethoden.

*In verband met het relatief lage aantal diensten dat is meegewogen in de categorie *Weblogs* en de hoge populariteit van Nu.nl is er mogelijk een vervorming opgetreden wat betreft de gebruiksfrequentie van deze zoekmethode.

5.4.3 Mannen versus vrouwen

Wanneer er gekeken worden naar verschillen tussen mannen en vrouwen in de frequentie waarop de zoekmethoden worden gebruikt zijn er slechts kleine verschillen te ontdekken. Het meest opvallende verschil is dat de mannen, op de categorie webdirectories na, aangeven alle methoden vaker te gebruiken dan vrouwen. Dit zou te verklaren kunnen zijn door het gemiddeld hogere aantal uur dat de mannen op internet besteden. Deze extra tijd biedt hun de mogelijkheid een groter aantal diensten te gebruiken. Waarom de webdirectories fractioneel hoger scoren bij de vrouwelijk populatie is niet volledig duidelijk.

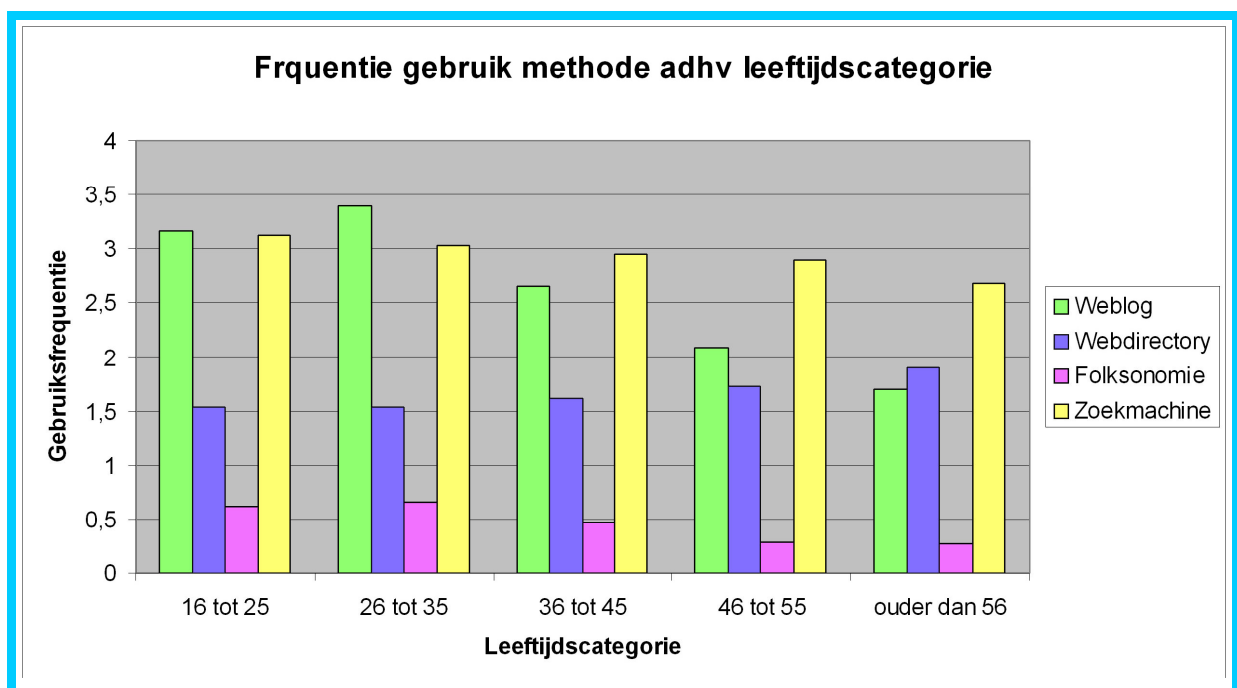


Figuur 36: Verschillen in frequentie gebruik zoekmethoden mannen en vrouwen

5.4.4 Leeftijd

Wanneer er gekeken wordt of er een verband is tussen de leeftijd van de respondent en de zoekmethode die zijn of haar voorkeur heeft komt een opvallende tendens naar voren. Het eerste wat opvalt is dat de relatief jongere respondenten (<35) significant meer gebruik maken van weblogs dan hun oudere tegenhangers (>35). Ditzelfde geldt voor het gebruik van folksonomieën. Ook het gemiddeld gebruik van zoekmachines neemt langzaam af naarmate de gemiddelde leeftijd van de respondenten hoger is.

Bij webdirectories is het omgekeerde zichtbaar, hoe ouder de respondent, hoe groter de kans dat deze gebruik maakt van deze zoekmethode. Een verklaring hiervoor kan zijn dat de eerste drie genoemde zoekmethoden een “nieuwere” vorm van zoeken zijn; jongere gebruikers worden grootgebracht met deze nieuwe methode en hebben hier dan ook een voorkeur voor. Oudere gebruikers lijken in deze dan ook te neigen naar een oudere manier van het doorzoeken van informatie, namelijk door middel van een directory, vergelijkbaar zoals deze bijvoorbeeld in bibliotheken e.d. gevonden kan worden.



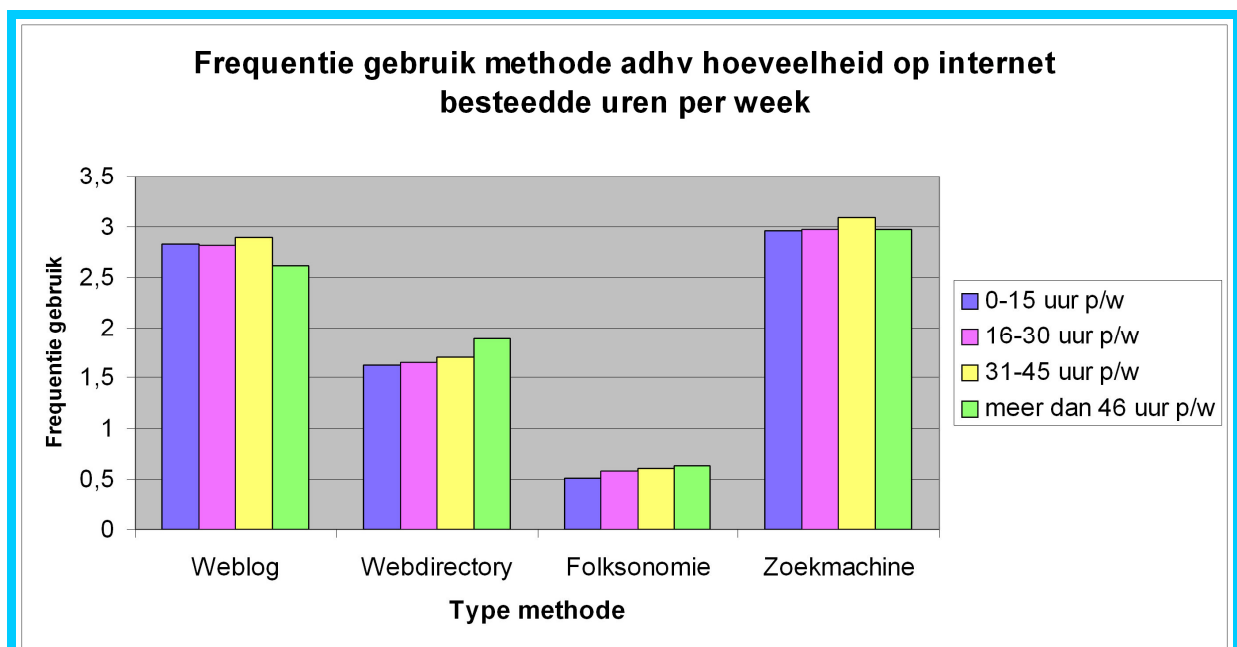
Figuur 37: Frequentie in gebruik zoekmethode gerangschikt aan de hand van leeftijdsklassen

5.4.5 Tijdsbesteding

Een laatste interessante kwestie is of het aantal uren dat de gebruiker op internet besteed invloed heeft op zijn of haar werkwijze.

In Figuur 38 ziet u deze twee zaken tegenover elkaar gezet. Zoals te zien in de grafiek, zijn er enkele verschillen te ontdekken.

Een duidelijk afwijking is zichtbaar in het lagere gebruik van weblogs door internetgebruikers die meer dan 45 uur per week op internet besteden. Bij webdirectories en folksonomieën neemt het gebruik zichtbaar toe naarmate de gebruiker meer tijd op internet doorbrengt. Bij de zoekmachines is er, op de groep 31-45 na, relatief weinig verschil in gebruik. De uitschieter van deze groep kan hoogstwaarschijnlijk worden toegerekend aan de duur van de gemiddelde fulltime werkweek en het leer/werkgebied van een deel van de responsegroep, namelijk onderzoekers en studenten van het wetenschappelijk onderwijs. Er kan aan de hand van deze grafiek de conclusie worden getrokken dat over het algemeen de keuze voor een zoekmethodiek niet verandert wanneer een persoon meer tijd op internet doorbrengt, maar slechts het absolute gebruik ervan toeneemt.



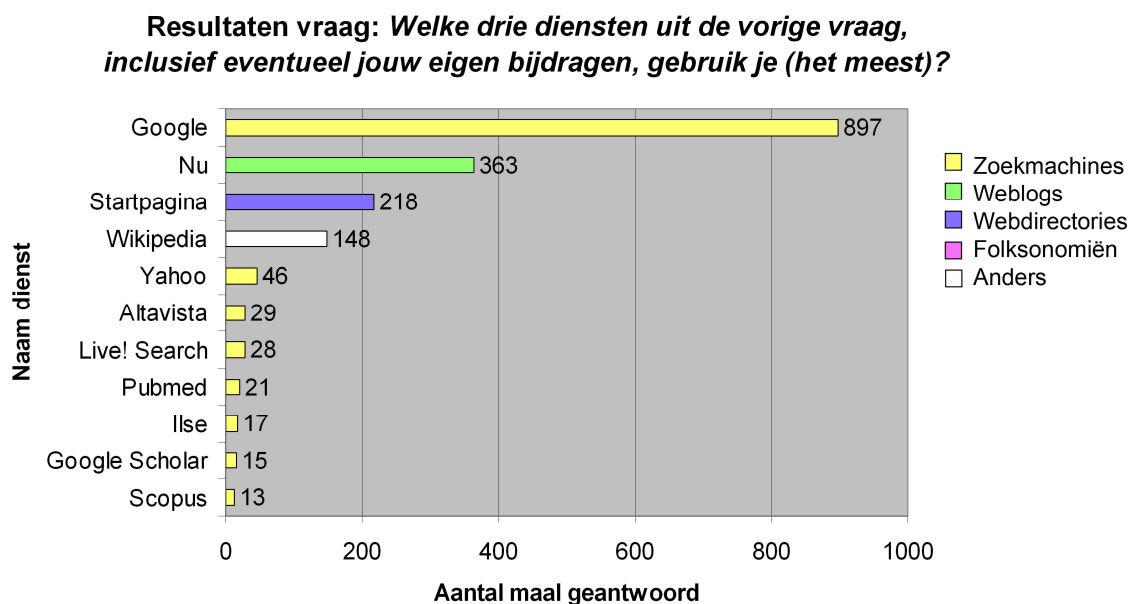
Figuur 38: Frequentie van gebruik zoekmethode gerangschikt aan de hand van het aantal uren dat per week op internet besteed wordt

5.4.6 Aanvulling door respondenten

De respondenten is aan het begin van de survey gevraagd welke drie diensten ze het meest gebruiken.

Als antwoord op deze vraag zijn in totaal 199 diensten genoemd. Van de respondenten gaven er 897 Google als antwoord, 363 gaven Nu.nl als antwoord en Startpagina werd 218 keer geantwoord. Dit zijn respectievelijk een zoekmachine, een weblog en een webdirectory.

Een opvallende uitkomst van deze vraag is dat, van de diensten die eruit naar voren kwamen, 81% (163/199) slechts éénmaal voorkwam in de lijst van antwoorden.



Figuur 39: De meest gebruikte zoekdiensten door respondenten

Op de vraag welke diensten volgens de gebruikers ontbraken in de lijst werd van de 681 antwoorden door een groot aantal respondenten Wikipedia geantwoord, namelijk drieëntwintig procent.

Hier wordt meteen een gevoelige plek in het opgestelde raamwerk van zoekmethoden zichtbaar. Wikipedia is niet te plaatsen binnen dit kader, vanwege het feit dat het niet voldoet aan een aantal eisen: de dienst Wikipedia beheert haar eigen content en de dienst is er niet op toegespitst om door te verwijzen naar elke soort content. Wel blijkt deze dienst dus een belangrijke rol in te nemen in het zoekgedrag van de respondenten.

Wikipedia word in deze lijst van ontbrekende diensten pas op grote afstand gevolgd door de nummer twee en drie, respectievelijk Tweakers (een informatica weblog, ±3 procent) en Google Scholar (zoekmachine voor wetenschappelijke content, ±3 procent).

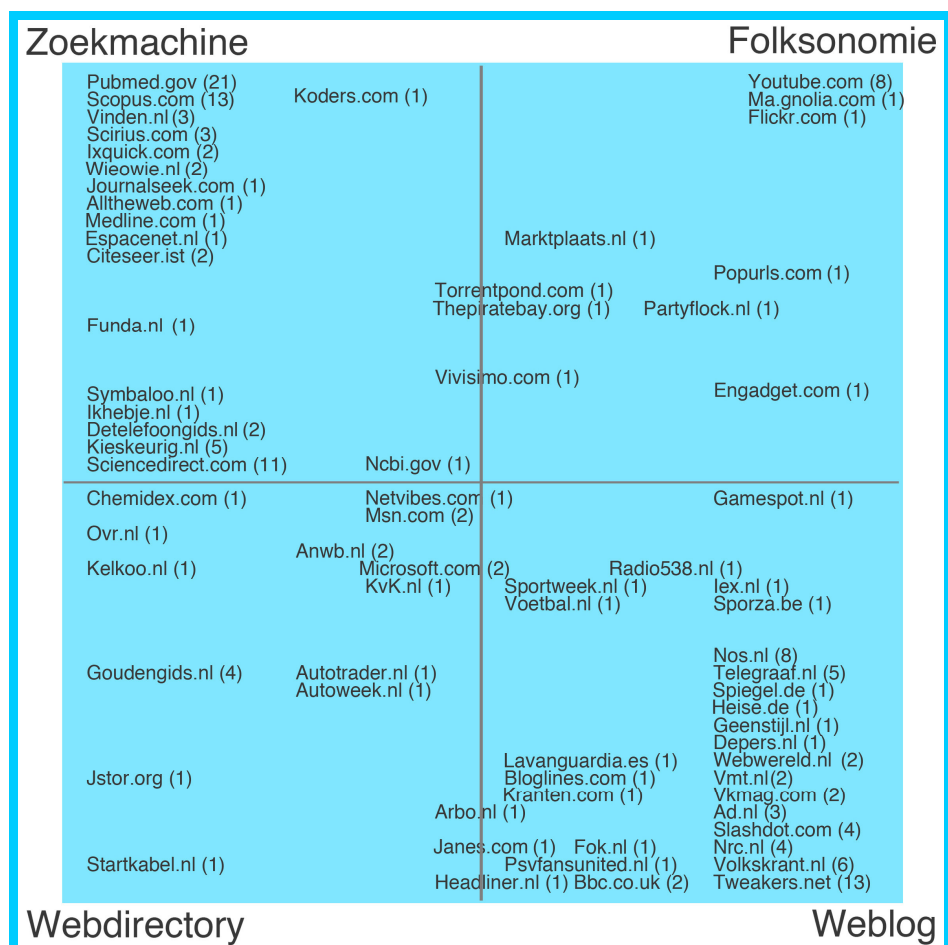
Na deze drie diensten liep de lijst al snel terug naar een lange lijst van diensten die slechts 1 of 2 keer als ontbrekend werden gezien. Er is hier duidelijk zichtbaar dat er naast een Long Tail van zoekresultaten op internet, er ook een Long Tail van zoekdiensten bestaat.

Aan de hand van deze vraag over gebruikte diensten blijkt de pragmatische selectie van diensten voor de dienstenanalyse H4 erg succesvol. De gekozen diensten, met Wikipedia als uitzondering, blijken voor de respondenten inderdaad de meest toonaangevende en/of gebruikte diensten van informatieontsluiting op internet te zijn.

De diensten die in deze Long Tail van zoekdiensten voorkomen variëren tussen digitale kranten, websites over voetbal, vergelijkingssites voor consumentenartikelen, sociale netwerksites

etcetera. Hiernaast (Figuur 40) is overzichtelijk gemaakt welke diensten er geantwoord zijn en op welke wijze deze informatie ontsluiten. Er zaten tussen de antwoorden vanzelfsprekend websites welke niet binnen de voor het onderzoek gestelde raamwerk passen; niet alle diensten op

internet zijn ontworpen om gebruikers te ondersteunen in het behalen van specifieke



Figuur 40: Overzicht van geantwoordde diensten in gebruikersonderzoek gerangschikt aan de hand hun methode van informatieontsluiting

zoekdoelen. Van diensten die dit wel als doel hebben kan een overgrote meerderheid goed geplaatst kan worden binnen het opgestelde kader van zoekmachines, folksonomieën, weblogs en webdirectories. In sommige gevallen bestaan de diensten uit een combinatie van twee of meer van deze zoekmethoden.

Toch is er een duidelijke afwijking zichtbaar: de diensten die in de Longtail van zoekdiensten naar voren komen zijn veelal gespecialiseerd op één onderwerp, één soort content of een combinatie van de twee. Een helder voorbeeld hiervan is bijvoorbeeld de dienst *Ikhebje*, een zoekmachine gespecialiseerd op het zoeken naar de persoonlijke gegevens van mensen op internet, zoals (geheime) telefoonnummers, adressen en dergelijke. Nog een voorbeeld is *Autoweek*; een hybride website die zowel een weblog over autonieuws aanbiedt en daarnaast ook de gebruiker voorziet van een zoekmachine welke gespecialiseerd is in doorzoeken van het aanbod van tweedehands auto's op internet.

Er is in het methodologisch kader reeds naar voren gebracht dat er lichte vertekening werd verwacht in het onderzoek richting hoger-opgeleide internetgebruikers, omdat deze een groot deel vertegenwoordigen van groep die primair benaderd is voor het onderzoek. In de resultaten van ontbrekende diensten lijkt deze veronderstelling te ondersteunen; van de twintig diensten die het meest geantwoord zijn als ontbrekend, zijn er vier met een overwegend wetenschappelijk specialisme.

5.5 Navigatieel, informatieel en transactioneel gebruik

In het volgende deel van de resultaten van de survey wordt er dieper ingegaan op het gebruik van zoekportals.

In de paragraaf hiervoor kunt u lezen hoe in de survey de respondenten gevraagd is een selectie te maken van drie diensten welke ze het meest gebruiken. Aan de hand hiervan is er door middel van driemaal acht vragen onderzocht welke dienst(en) nu precies voor welke zoekdoel(en) worden gebruikt.

De eerste van deze acht stellingen heeft betrekking op behalen van het navigatiele zoekdoel. Stelling 2 t/m 5 hebben betrekking op het informatiele zoekdoel, met

respectievelijk de subdoelen *directed*, *undirected*, *advice* en *locate*. Stelling 6 t/m 8 hebben betrekking op het transactionele zoekdoel, met respectievelijk de subdoelen *interact*, *entertainment* en *download*.

Door de geantwoordde diensten te categoriseren in de vier zoekmethoden (zoekmachines, folksonomieën, webdirectories en weblogs), is het inzichtelijk te maken welke van deze zoekmethoden door gebruikers ter hand wordt genomen wanneer zij een webadres (*navigatieel zoekdoel*), informatie (*informatieel zoekdoel*) of een interactie (*transactioneel zoekdoel*) proberen te behalen.

In de grafieken en tabellen is de relatieve gebruiksfrequentie zichtbaar. Op deze wijze is de frequentie van gebruik relatief ten opzichte van de andere methoden, om zo eventuele vervoming door het absolute gebruik te voorkomen.

5.5.1 Navigatieel

Als eerste is er gekeken naar de mate van het gebruik van de verschillende methoden voor het behalen van het navigatiele zoekdoel.

Uit de resultaten komt naar voren dat zoekmachines voor dit zoekdoel veruit het populairst zijn. 34% van de respondenten geeft aan deze methode vaak (+) voor dit doel te gebruiken en zelfs 43% geeft aan altijd (++) deze methode te gebruiken. Naast het feit dat een zeer grote groep respondenten de dienst met grote regelmaat gebruiken, is er ook maar een zeer kleine groep respondenten die deze methode nooit (-) gebruikt, namelijk slechts 3%.

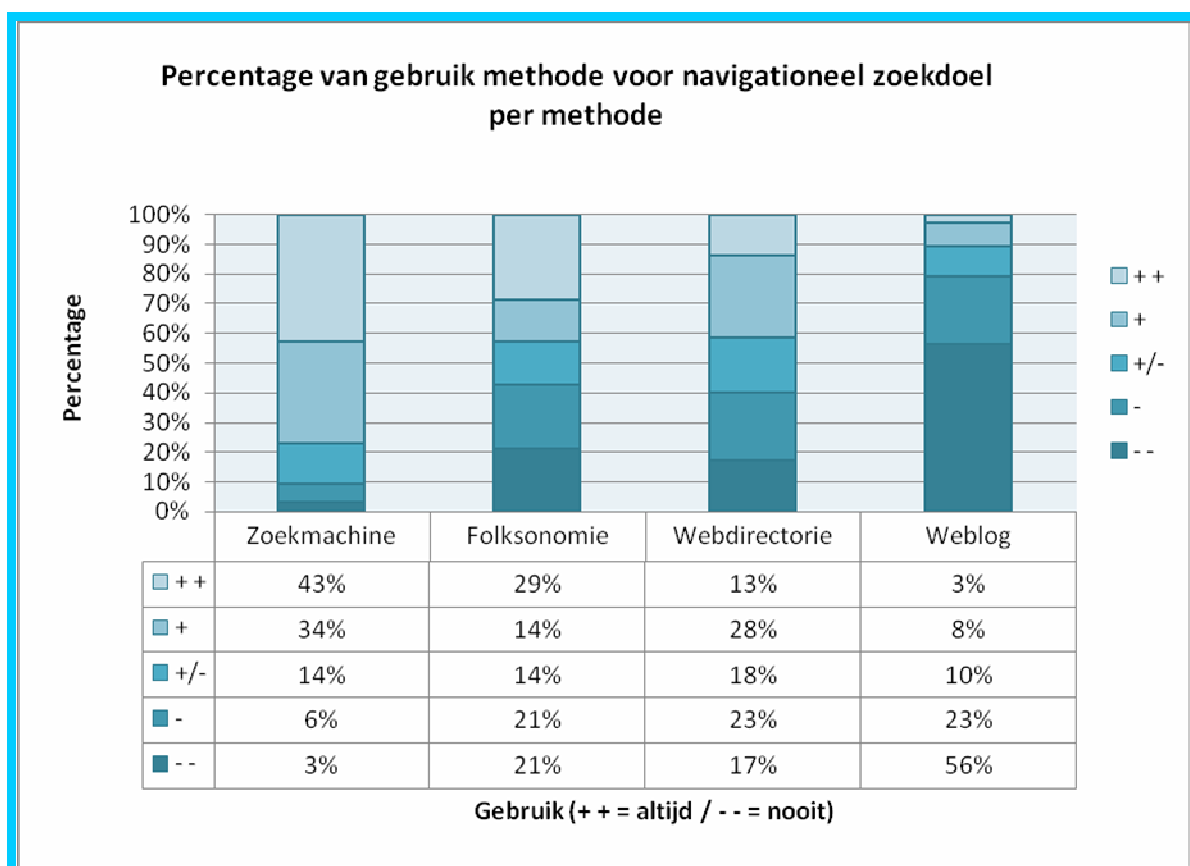
De resultaten wat betreft het gebruik van folksonomieën voor het behalen van dit zoekdoel geeft een verdeeld beeld. 42% van de respondenten geeft aan deze methode nooit (-) of slechts zelden (-) te gebruiken. Daartegenover staat echter een groep respondenten (29%), die antwoorden deze methode altijd (++) te gebruiken. Tussen deze twee uitersten is een duidelijk kloof zichtbaar.

De antwoorden over het gebruik webdirectories voor dit zoekdoel duidt op een gelijkmatiger verdeeld gebruik. Rond de 15% van de respondenten geeft aan deze methoden altijd (++) of

nooit(++) te gebruiken (respectievelijk 13% en 17%). 28% gebruikt deze methode vaak (+), tegenover 23% die de methode zelden (-) gebruikt.

De methode die voor dit zoekdoel het minst gebruikt wordt door de respondenten zijn weblogs. 23% van de respondenten geeft aan deze methode zelden (-) te gebruiken voor het behalen van dit zoekdoel en zelfs meer dan de helft (56%) geeft aan deze methode nooit (- -) te gebruiken voor dit zoekdoel. Slechts 3% zegt deze methode altijd (+ +) te gebruiken.

Hieronder een (samenvattend) overzicht



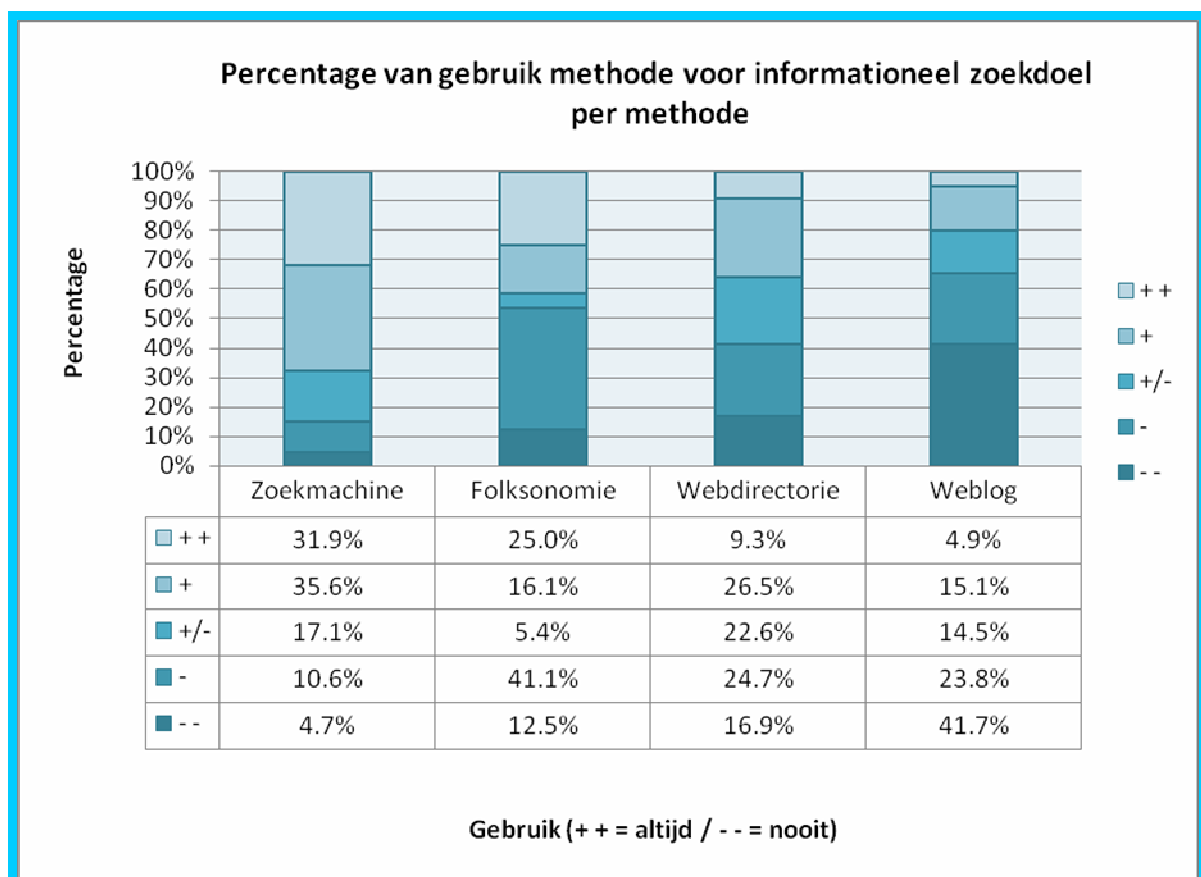
Figuur 41: Relatief gebruikspercentage van zoekmethoden voor het behalen van het navigatieveel zoekdoel

Er is voor dit zoekdoel een segmentatie gemaakt binnen de groep respondenten aan de hand van bepaalde kenmerken, zoals geslacht, leeftijd (<35 vs. >35jr) of aantal uren internetgebruik per week om er achter te komen of er verschillen zijn in gebruik tussen deze groepen. Uit deze resultaten van deze segmentatie kwamen slechts één noemenswaardig verschil naar voren welke afweek van de algemene verschillen in gebruik: de groep welke

met grote regelmaat (vaak(+)) en altijd(++)) folksonomieën gebruikt om dit doel te behalen bestaat voor bijna tachtig procent uit mannen.

5.5.2 Informatieel

Uit de vragen die betrekking hebben op het informatiele zoekdoel en de keuze van zoekmethoden komt de volgende tabel/grafiek naar voren.



Figuur 42: Relatief gebruikspercentage van zoekmethoden voor het behalen van het informatieel zoekdoel

Er is te zien dat het gebruik van zoekmachines voor het informatiele doel vrijwel gelijk is als het gebruik voor een navigationeel zoekdoel; bijna 70% van de respondenten geeft aan deze methode vaak (+) of altijd (++) te gebruiken. Er is wederom slechts een kleine groep respondenten die antwoord deze methode nooit (--) te gebruiken.

De weblogs scoren bij dit zoekdoel het zwakst; bijna 75% van de respondenten zegt deze methode zelden (-) of nooit (-) te gebruiken voor dit doeleinde.

De webdirectories hebben ook bij dit zoekdoel een redelijk gelijkmatige verdeling van het gebruik met een geleidelijke afname in het percentage richting de extremen (nooit/altijd). In het gebruik van de folksonomische methode voor het informationele zoekdoel is een duidelijke tweedeling zichtbaar. Iets meer dan de helft van de respondenten geeft aan dit nooit of zelden te gebruiken, maar daar tegenover staat een groep van een kwart die de methode altijd gebruiken.

Opvallend is bij dit zoekdoel is wederom het verschil wanneer er gekeken wordt naar het gebruik van mannen en vrouwen. Waar een, zij het kleine, groep mannen een overwegend positief antwoord geeft op het gebruik van folksonomieën, is er vrijwel geen enkele vrouw die een dienst met deze zoekmethode in haar top drie heeft genoemd. De vrouwen die dit wel hebben gedaan, geven aan deze zoekmethode zelden of nooit te gebruiken voor dit zoekdoel.

Wat betreft het gebruik van weblogs en webdirectories is het verloop andersom. Relatief gezien hebben meer vrouwen dan mannen diensten met deze zoekmethode in hun lijst met meest gebruikte diensten. Binnen de groep van respondenten die aangeven deze methoden wél gebruiken, geven de vrouwen ook aan deze gemiddeld vaker te gebruiken dan de mannen.

Wat betreft het gebruik van zoekmachines is er tussen mannen en vrouwen weinig verschil in de mate van gebruik te ontdekken.

Wanneer er gekeken wordt naar het aantal uren dat een respondent op internet is en de zoekmethode is er één afwijking waar te nemen, namelijk wat betreft het gebruik van folksonomieën voor het informationele zoekdoel. Er zijn slechts enkele respondenten die minder dan twintig uur per week op internet zijn (waarvan geen enkele onder de tien uur) die deze methode gebruiken. Het merendeel van de gebruikers van deze methode om informationele zoekdoelen te behalen besteedt veertig of zestig uur per week op internet. Dit ten opzichte van een gemiddeld internetgebruik van rond de 16 uur per week.

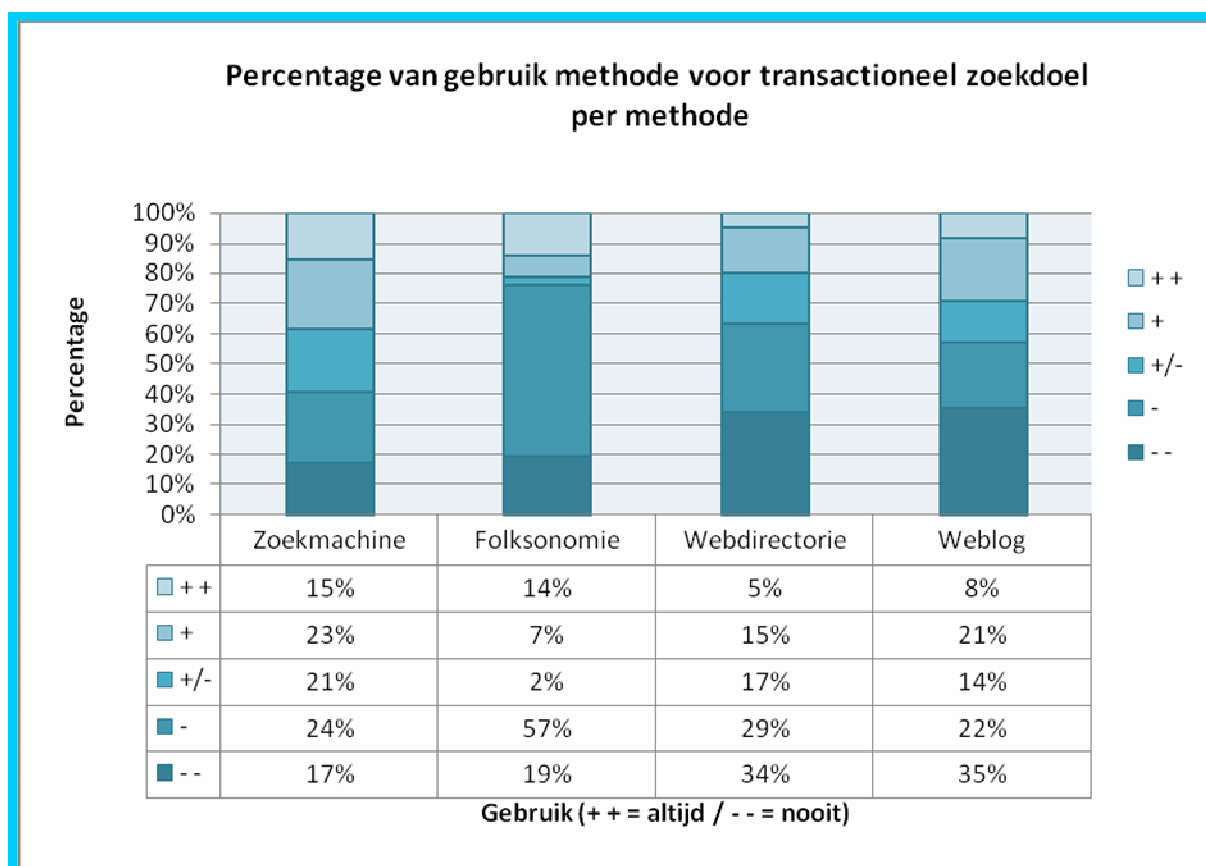
Er is geen relevant verschil op te merken tussen de keuze van zoekmethoden voor dit en de leeftijd van de respondent.

Zoals eerder in het onderzoek besproken, is er binnen het informationele zoekdoel ook nog onderscheid te maken in directed, undirected en advice en locate (pag. 34).

Binnen deze subdoelen zijn de verhoudingen van gekozen methoden vrijwel gelijk aan die van de het overkoepelende zoekdoel. Het enige subdoel waar een verschil opgemerkt kan worden is wanneer er onderscheid wordt gemaakt tussen respondenten met een leeftijd van boven en onder de 35 jaar. In de laatste van deze twee groepen wordt met de zoekmachine methode zo'n 10% meer gezocht met het *advice* zoekdoel.

5.5.3 Transactioneel

Voor het transactionele zoekdoel ziet het gebruik van zoekmethode er als volgt uit.



Figuur 43: Relatief gebruikspercentage van zoekmethoden voor het behalen van het transactioneel zoekdoel

Er wordt in de bovenstaande figuur over het gebruik van de verschillende methoden voor het behalen van een transactioneel zoekdoel een behoorlijk andere verdeling zichtbaar in vergelijking met het navigationale en transactionele zoekdoel.

De meest opvallende hiervan is de afname in eerdere dominantie van de zoekmachines bij andere zoekdoelen. Een deel hiervan wordt overgenomen door de weblogs, die bij de categorie vaak (+) een tweede plaats behalen met een gebruik van 21%. Ook de

webdirectories scoren qua gebruik lager dan bij het navigatieve en informatieve zoekdoel. De hoeveelheid respondenten die deze dienst nooit gebruikt voor dit zoekdoel verdubbelt zelfs bijna ten opzichte van de andere zoekdoelen.

Ook bij dit zoekdoel wordt het gebruik van folksonomieën gekarakteriseerd door een tweestrijd. De twee grote pieken zijn voor de categorie soms (-) en de categorie altijd (+ +). Wanneer er gekeken wordt naar verschillen binnen leeftijdsgroepen, geslacht of het aantal uren internetgebruik komen de volgende verschillen naar voren:

Het meest opvallende verschil is te zien tussen het gebruik door mannen en vrouwen.

Vrouwen gebruiken voor dit zoekdoel webdirectories zo'n 8 procent meer dan mannen.

Het tweede verschil is dat zichtbaar wordt dan de groep respondenten onder de 35 jaar weblogs een stuk vaker gebruikt dan de groep respondenten van boven de 35 jaar oud. Dit verschil ligt respectievelijk rond de 28% versus 13%. Dit verschil in gebruik kwam reeds eerder in het gebruikersonderzoek ook duidelijk naar voren.

Ook dit zoekdoel bestaat uit enkele subdoelen, zijnde: download, entertainment en interact.

Wanneer er gekeken wordt naar deze subdoelen komen er nog enkele (kleine) verschillen naar voren die afwijken van de eerder aangegeven verschillen en/of overeenkomsten. Het eerste verschil dat naar voren komt uit de data is dat de mannelijke groep respondenten de folksonomische methode relatief (nog) vaker gebruikt voor het behalen van het subdoel *download*, oftewel voor het downloaden van bijvoorbeeld software of videobestanden. Een tweede verschil is zichtbaar in het gebruik voor het *entertainment* subdoel; vrouwen gebruiken zoekmachines meer voor het behalen van dit zoekdoel dan mannen.

5.6 Conclusie gebruikersonderzoek

Uit het gebruikersonderzoek zijn een divers aantal zaken naar voren gekomen die helpen antwoord te krijgen op de opgestelde hoofd- en deelvragen van dit onderzoek.

Er is gebleken dat respectievelijk de zoekdiensten Google, Startpagina en Yahoo! de grootste naamsbekendheid genieten binnen de responsgroep.

De diensten die door de geënquêteerde internetgebruikers het meest worden gebruikt zijn Google (zoekmachine), gevolgd door Startpagina (webdirectory) en Nu.nl (weblog). Andere diensten volgen in gebruik deze “big three” op grote afstand.

In het onderzoek wordt Wikipedia door veel respondenten aangevuld als ontbrekende dienst en wordt onderschreven als een belangrijke schakel in het zoekproces.

Verder bevestigt het gebruikersonderzoek de juistheid van de selectie van diensten voor de dienstenanalyse (Pag. 50)

In de frequentie van gebruik van de verschillende zoekmethoden worden er enkele verschillen duidelijk wanneer verschillende groepen worden gesegmenteerd.

Het verschil in frequentie van gebruik van de zoekmethoden tussen mannen en vrouwen is evenredig aan het aantal uren dat deze groepen op internet besteden; mannen besteden enkele uren per week meer op internet en geven aan de diensten dan ook met iets grotere regelmaat te gebruiken. Uitzondering hierop zijn de webdirectories, deze worden door vrouwen meer gebruikt.

Wanneer er gekeken wordt naar het verschil in leeftijd en de frequentie van gebruik komt een opvallend beeld naar voren. De jongere respondenten (<35 jaar) maken meer gebruik van zoekmethoden waarvan de werkwijze relatief nieuw is; met name de weblogs doen het goed binnen deze leeftijdsgroep. Er is een duidelijk correlatie zichtbaar tussen het gebruik

van webdirectories en de leeftijd van de respondent; hoe ouder de respondent, hoe meer deze de webdirectories gebruikt.

Het verband tussen het aantal uren dat per week op internet wordt besteed en de frequentie van gebruik van zoekmethoden werd eerder al aangestipt; des te meer uren een respondent gemiddeld gezien op internet doorbrengt, des te vaker hij of zij gebruikt maakt van de verschillende zoekmethode. Toch is er één uitzondering; de groep internetgebruikers die meer dan zesenvestig uur per week op internet besteedt maakt duidelijk minder gebruik van weblogs. Uit de beschikbare data is het helaas niet mogelijk om hier een eenduidige verklaring voor te geven.

De internetgebruikers die aan het onderzoek hebben meegewerkt geven aan voor het behalen van een *navigatieve* zoekdoel overwegend zoekmachines te gebruiken. De *webdirectories* blijken voor dit zoekdoel de tweede keuze te zijn. De weblogs worden minimaal gebruikt voor dit zoekdoel.

Voor het behalen van *informatieve* zoekdoelen wordt door de geïnterviewden wederom de zoekmachine het meest gebruikt. Op de tweede plaats komen de webdirectories. In de categorie altijd (+ +) worden de zoekmachines echter opvallend dicht op de voet gevolgd door de folksonomieën. De afwijking gaat gepaard met nog een afwijking op een iets ander vlak. Er komt uit deze vraag namelijk naar voren dat deze folksonomieën vrijwel uitsluitend door de mannelijke respondenten, die bovengemiddeld veel tijd op internet zijn, ter hande worden genomen. Waar gemiddeld gezien het vrouwelijk gebruik redelijk gelijk op gaat met die van de mannen, blijft deze hier ver achter.

Voor het behalen van *transactieve* zoekdoelen worden zoekmachines wederom het meest gebruikt. Wel moet deze zoekmethode een groot deel inleveren qua gebruik aan de weblogs. Waar bij de andere zoekdoelen deze methode grotendeels achterbleef, is deze hier een stuk meer aanwezig. Ook zijn de folksonomieën, in de categorie altijd (+ +) ook hier weer gewaagd aan de zoekmachine, met zelfs maar één procent verschil.

Het onderzoek laat ook blijken dat de geënquêteerde internetgebruiker overwegend positief is wat betreft de mogelijkheden en de werking van zoeken en vinden op internet. Ze geven aan er relatief weinig moeite mee te hebben, en vrij snel te kunnen vinden waarnaar ze opzoek zijn. Ook komt uit de stellingen naar voren dat Google het populairst is: meer dan vijftig procent geeft aan het eens of redelijk eens te zijn dat ze met deze dienst alles kunnen vinden wat ze zoeken.

6. Verdieping onderzoeksresultaten

In deze paragraaf zijn er verbanden gelegd tussen de analyse van de zoekmethoden uit het onderzoek (Hoofdstuk 4) en het gebruik van deze methode door de respondenten (Hoofdstuk 5).

6.1 Zoekmachines

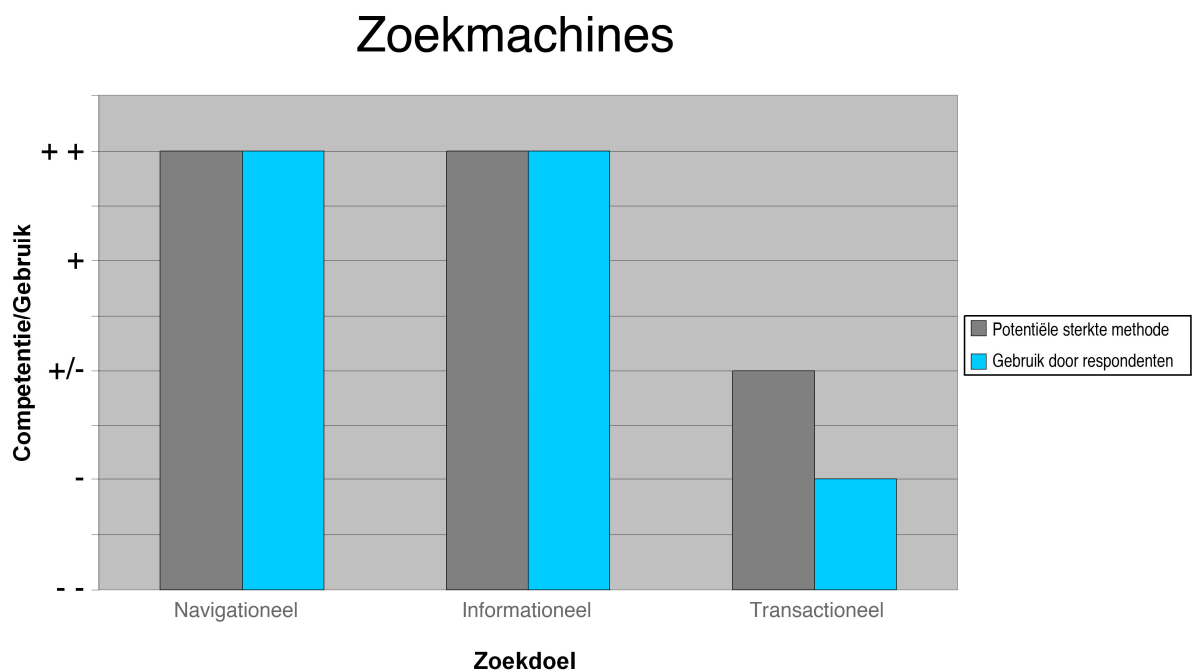
Eerder in dit onderzoek is gekeken naar de werkwijze waarmee zoekmachines informatie die op internet staat ontsluiten en op welke wijze ze de gebruikers ondersteunen in de zoektocht naar deze informatie. Hieruit kwam naar voren dat zoekmachines, zoals Google, Yahoo! en LiveSearch, voornamelijk erg sterk zijn op het gebied van het indexeren en vindbaar maken van tekstuele content. De zoekmachines kunnen ook zeer snel metadata over content analyseren, maar het ontbreekt aan de controle over het waarheidsgehalte van deze metadata. Ook wordt voor dit doel vaak een waardeoordeel over de content gezocht, een oordeel dat door computers niet of nauwelijks te maken valt.

De techniek die zoekmachines gebruiken voor informatieontsluiting maakt deze methode voornamelijk geschikt voor het behalen van navigatieve (het vinden van een webadres) en informatieve (het vinden van informatie) zoekdoelen. Op het gebied van transactionele zoekdoelen (het zoeken naar een interactie op internet, zoals afbeeldingen, spellen etc.) zijn de zoekmachines minder sterk.

Wanneer er gekeken wordt naar het gebruik van zoekmachines door de ondervraagde internetgebruikers wordt dit beeld bevestigd. De meerderheid van de respondenten gebruikt zoekmachines voor het bevredigen van het navigatieve zoekdoel. Drieënveertig procent van de gebruikers geeft zelfs aan de zoekmachine *altijd* voor dit zoekdoel te gebruiken.

Ook voor het behalen van informationele zoekdoelen wordt de zoekmachine het meest gebruikt, ongeveer een derde gebruikt de diensten hier altijd voor, nog een derde geeft aan deze methode er *vaak* voor te gebruiken.

Wanneer het gaat om het behalen van een transactioneel zoekdoel moeten de zoekmachines een deel van hun populariteit inleveren; maar een op de zes respondenten geeft aan zoekmachines hier altijd voor te gebruiken.



Figuur 44: Sterkten zoekmachines versus gebruik door respondenten

6.2 Folksonomieën

De werkwijze van folksonomieën, zoals Delicious, Digg en Nujij waarin gebruikers content kunnen aanraden aan andere gebruikers of deze kunnen taggen heeft ook een aantal duidelijke sterkten en zwakten.

Op het gebied van het navigationale doeleinde kwam naar voren dat de folksonomieën niet erg geschikt zijn voor het behalen van dit zoekdoel. De folksonomieën richten zich voornamelijk op de inhoud van de content. Wel zijn de folksonomieën sterk in het *terugvinden* van webadressen (*retrieval*), veelal omdat deze hier door de gebruikers zelf geplaatst kunnen worden onder de terminologie die zij zelf hanteren.

Wat betreft het behalen van het informationele zoekdoelen hebben de folksonomieën een redelijke potentie. De content die via deze methode gevonden kan worden heeft het vermogen goed aan te sluiten bij de wensen van de gebruikers, aangezien deze zelf verantwoordelijk zijn voor de selectie van content. Een nadeel aan deze methode is dat deze niet op kan tegen de snelheid van indexering van de computerprogramma's van zoekmachines. Ook qua hoeveelheid geïndexeerde content zijn de folksonomieën niet altijd volledig genoeg.

Op het gebied van het behalen van transactionele doeleinde komt de kracht van de folksonomieën duidelijk naar voren. Bij transactionele doeleinden gaat het vaak om een menselijk waardeoordeel over content, iets wat gebruikers vanzelfsprekend beter doen dan computerprogramma's.

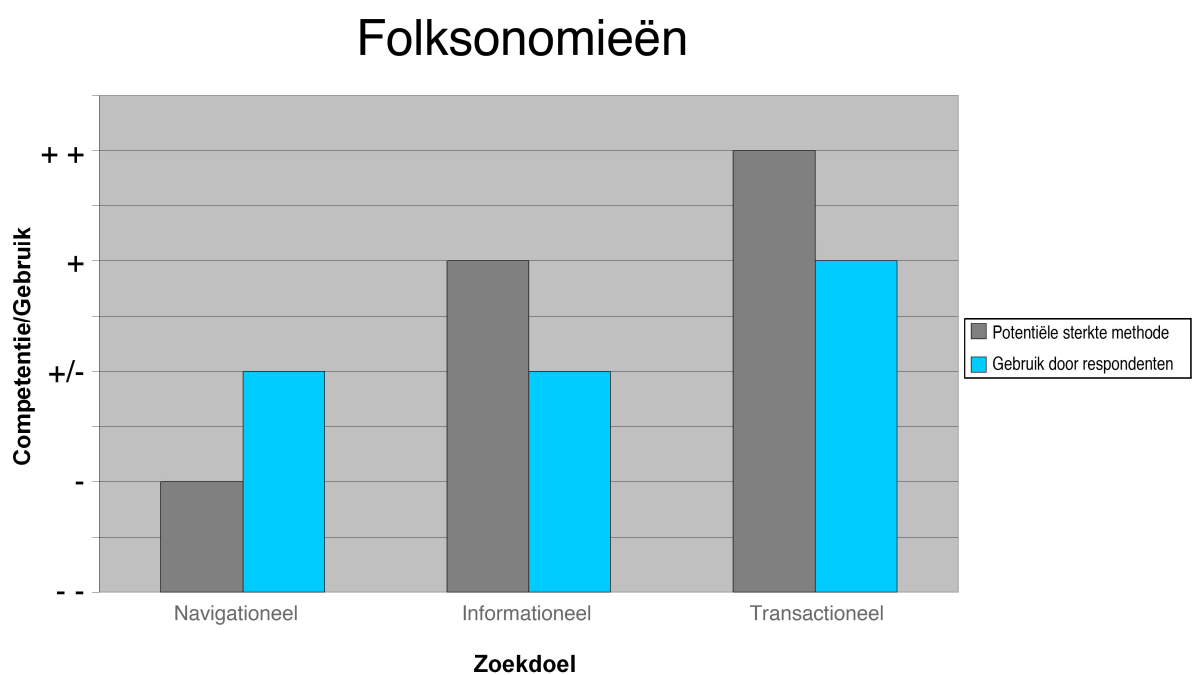
In het gebruik van folksonomische diensten door gebruikers komt een vergelijkbaar beeld naar voren.

Het gebruik van folksonomieën door internetgebruikers om navigationale doeleinden te behalen is relatief laag. Alleen bij de groep gebruikers die zeggen deze methode altijd te gebruiken scoort de dienst onevenredig hoog, wat hoogstwaarschijnlijk te danken is aan de eerdergenoemde mogelijkheid om eerder gevonden content terugvindbaar te maken middels deze diensten.

Op het gebied van informationele zoekdoelen scoren de folksonomieën erg laag, slechts een klein deel van de gebruikers geeft aan deze diensten te gebruiken. Ook bij dit zoekdoel scoort de groep die het altijd zegt te gebruiken onevenredig hoog.

Op het gebied van transacties scoort de folksonomie naar verwachting; deze scoort hier relatief erg hoog, en komt als methode op de tweede plaats wanneer het het behalen van dit zoekdoel betreft.

Er lijkt bij folksonomieën een tweestrijd zichtbaar, een grote groep maakt relatief weinig of zelfs geen gebruik van deze diensten, tegenover een kleine groep die er juist veel gebruik van maakt. In het gebruikersonderzoek kwam reeds naar voren dat de mensen die relatief veel tijd op internet doorbrengen deze methode vaker gebruiken dan mensen die minder tijd op internet besteden. Deze tweestrijd zou dus veroorzaakt kunnen worden door de vaardigheid die de omgang ermee vereist. Deze vaardigheid is wellicht bij minder frequente internetgebruikers niet optimaal ontwikkeld.



Figuur 45: Sterkten folksonomieën versus gebruik door respondenten

6.3 Webdirectories

In de analyse van de methodiek van webdirectories is naar voren gekomen dat deze manier om content te ontsluiten een wat oudere manier is om informatie te indexeren. De werkwijze van diensten als Startpagina en Open Directory Project is vrijwel identiek aan die van methoden die ook offline worden ingezet om informatie vindbaar maken. Deze diensten alfabetiseren of rangschikken informatie op onderwerp, vergelijkbaar met dat van een bibliotheek. Voordelen hiervan zijn de vaardigheid die veel internetgebruikers hebben met deze werkwijze en de specialisatie op onderwerpen. Een nadeel ervan is dat deze manier vaak erg rigide is, en het erg complex is om een passend raamwerk op te stellen waarin alle geïndexeerde content logisch geordend is.

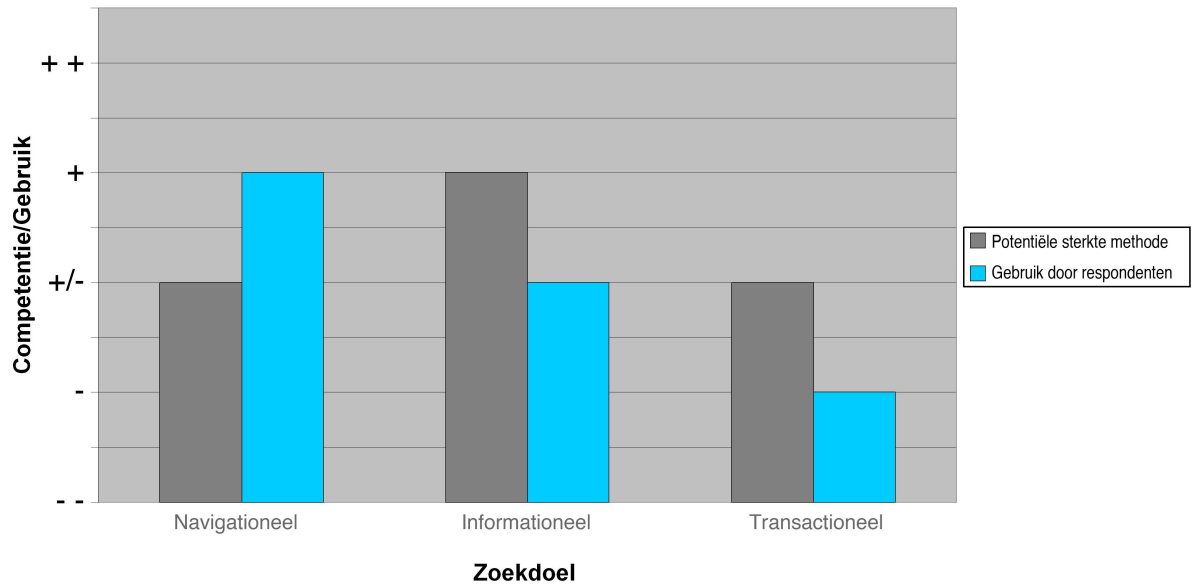
In de analyse wordt naar voren gebracht dat deze methode, vanwege de specialisatie en selectie op onderwerp, voornamelijk geschikt is voor het informationele doeleinde en vanwege de, vaak, menselijke tussenkomst op het gebied van transactionele zoekdoelen. Het navigationale zoekdoel scoort minder hoog, omdat de directory-pagina over de gezochte website eerst gevonden moet worden, voordat de stap naar het gezochte webadres gemaakt kan worden.

In de gebruikersanalyse komt echter een ander beeld naar voren. Juist op het gebied van het navigationale doeleinde scoort de webdirectory relatief hoog. Ongeveer een op de drie respondenten geeft aan de dienst hier vaak voor te gebruiken.

Op het gebied van het informationele doeleinde scoort de methode gemiddeld; een middelgrote groep gebruikers geeft aan de dienst met enige regelmaat of vaak te gebruiken voor dit zoekdoel. Slechts een zeer klein deel geeft aan deze dienst hier altijd voor te gebruiken.

Ook wat betreft het transactionele zoekdoel staat de methode analyse gedeeltelijk haaks op de uitkomsten van het gebruikersonderzoek. Een op de drie gebruikers zegt deze methode hier maar soms voor te gebruiken, en nog een derde geeft aan deze methode er nooit voor te gebruiken. Dit is op zijn minst opvallend te noemen, aangezien er door de eerdergenoemde menselijke tussenkomst in de plaatsing van content binnen een directory er impliciet een waardeoordeel wordt gegeven over de inhoud ervan.

Webdirectories



Figuur 46: Sterkten webdirectories versus gebruik door respondenten

6.4 Weblogs

Als laatste komen we aan bij de werking en het gebruik van de weblogs als methode van informatieontsluiting.

Bij de analyse van de methode komt naar voren dat diensten die deze werkwijze gebruiken, waaronder Spitsnieuws en Nu.nl, veelal zwak zijn in het bevredigen van het navigationale doeleinde. Weblogs zijn voornamelijk gericht op het aanbieden van nieuwe, gespecialiseerde content op chronologische volgorde. Dit is een methode die zich niet goed leent om te gebruiken in de zoektocht naar een webadres; er is geen garantie dat het gezochte webadres ook daadwerkelijk voorkomt in de database van dit soort diensten.

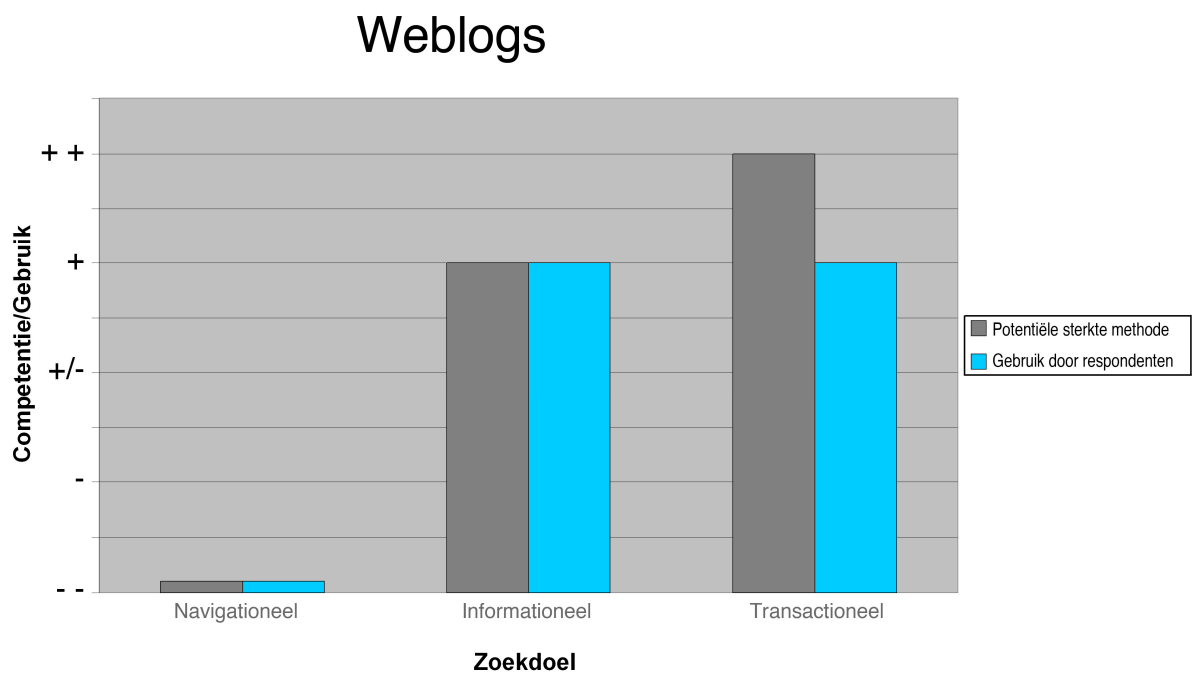
Op het informationele vlak heeft een weblog meer te bieden. De constante stroom van nieuwe informatie die er geplaatst wordt, maakt het voor de gebruiker eenvoudig op de hoogste te blijven van nieuwe informatie en/of content. En gebrek hierbij is dat de content niet gericht gezocht wordt, een pull situatie, maar juist naar de gebruiker gebracht wordt, een push situatie. Hierdoor is de gebruiker voor selectie van informatie afhankelijk van de persoon of personen die de weblog onderhouden. Wanneer dit goed overeenkomt kan de gebruiker op eenvoudige wijze veel informatie zoeken en vinden (of zelfs, krijgen) via deze methode.

Voor het behalen van transactionele doeleinden hebben weblogs ook een groot potentieel. Dit zoekdoel betekent vaak dat de gebruiker op zoek is naar een bepaalde vorm van vermaak, of dit nu spellen, muziek of afbeeldingen zijn. Deze manier van het aanbieden van content maakt het mogelijk voor de gebruiker om een relatief passieve houding aan te nemen in plaats van een actieve om zijn transactionele zoekdoelen te bevredigen.

Wanneer er gekeken wordt naar het gebruik van weblogs voor het navigationale zoekdoel worden de verwachtingen waargemaakt; meer dan de helft van de gebruikers geeft aan deze methode *nooit* te gebruiken voor het zoeken naar webadressen. Wat betreft het informationele zoekdoel scoort de weblog enigszins lager dan verwacht. De kracht om informatie te filteren en aan te bieden wordt door gebruikers niet gezien als een goede

methode om op zoek te gaan naar informatie; meer dan zestig procent van de respondenten geeft dan ook aan deze werkwijze nooit of zelden te gebruiken.

De prognose over het gebruik van deze methode voor transactionele zoekdoelen komt passend overeen met het daadwerkelijke gebruik van de respondenten. De weblogs nemen voor het behalen van dit zoekdoel een significant groter deel van het gebruik in beslag dan bij de voorgaande zoekdoelen. Een op de drie gebruikers geeft aan deze methode vaak of altijd voor dit zoekdoel te gebruiken.



Figuur 47: Sterkten weblogs versus gebruik door respondenten

6.5 Conclusie

Het gebruik van de verschillende zoekmethoden voldeed veelal goed aan de verwachtingen die waren afgeleid uit de sterkten en zwakten van de werkwijze van de methode.

De *zoekmachines* nemen zowel bij het navigationele als bij het informationele zoekdoel, respectievelijk het zoeken naar een webadres of informatie, een groot deel van het gebruik voor hun rekening.

Wanneer er een waardeoordeel vereist is in de afweging van de relevantie van content, zoals dit vaak het geval is bij transactionele zoekopdrachten, schiet deze werkwijze echter te kort. Dit is ook zichtbaar in het gebruik.

De *folksonomieën* wijken enigszins af van de prognoses in het gebruik. Het gebruik voor navigationele zoekdoelen scoort hoger dan verwacht. Dit is waarschijnlijk te wijten aan de mogelijkheid dat gebruikers hier vaak reeds gevonden content terugvindbaar kunnen maken aan de hand van hun eigen metadata.

Voor het informationele zoekdoel scoort deze methode laag, slechts een kleine groep gebruikers maakt met regelmaat gebruik van deze methode, de overgrote meerderheid kiest voor een andere methodiek.

Op het gebied van transactionele zoekdoelen maakt de folksonomie de belofte waar, de kracht van de menselijke opinie en indexering, komt duidelijk naar voren in de frequentie van gebruik.

Er is binnen het gebruik van folksonomieën een kloof zichtbaar; gebruikers met een hoog internetgebruik geven veelal aan deze methode als een van hun favorieten te zien, waar minder ervaren gebruikers laten zien deze methode juist vrijwel niet te gebruiken. Dit zou kunnen duiden op een relatief hoge toetredingsdrempel voor dit type dienst.

De *webdirectories* verassen binnen dit onderzoek. De kracht van webdirectories ligt technisch gezien op het zichtbaar organiseren van de meest relevante content, veelal gebaseerd op menselijk inzicht. Dit zou moeten betekenen dat deze methode hoog zou moeten scoren op informationeel en transactioneel gebied. De cijfers spreken dit echter tegen; veel gebruikers nemen deze methode ter handen om het navigationele doeleinde te

bevredigen en in mindere mate het informationele. Het transactionele gebruik ligt beduidend lager.

Wat betreft het gebruik van *weblogs* wordt er door de gebruikers recht gedaan aan de sterkten en zwakten van de methode om informatie te ontsluiten. Op het gebied van navigationale zoekdoelen scoort de methode zeer slecht. Het gebruik voor informationele zoekdoelen is gematigd, maar bij de kracht van deze methode, het transactionele zoekdoel, scoort de werkwijze een stuk hoger.

Samenvattend kan er gesteld worden dat de zoekmachines voor alle zoekdoelen het meest gebruikt worden, maar het te behalen zoekdoel heeft onmiskenbaar een sterke invloed op de selectie van diensten door de gebruiker.

Daarnaast lijkt de selectie van methoden die gebruikers maken voor het behalen van een specifiek zoekdoel sterke overeenkomsten te vertonen met de competenties van de geselecteerde methode.

7. Conclusie

7.1 Onderzoeksverloop

In deze thesis is er geprobeerd antwoord te geven op de centrale vraag van het onderzoek, namelijk:

Welke methoden worden door de toonaangevende diensten online aangeboden aan gebruikers om content op internet te ontsluiten, wat zijn de voor- en nadelen van deze methoden en diensten en met welk zoekdoel worden de verschillende methoden door gebruikers gehanteerd en ervaren?

Om deze vraag te kunnen beantwoorden zijn de volgende drie deelvragen geformuleerd:

- Welke type methoden bieden toonaangevende diensten aan op internet voor het ontsluiten van content?
- Wat zijn de voor- en nadelen van deze methoden en met welk zoekdoel worden verschillende methoden door gebruikers gehanteerd?
- Hoe wordt het zoeken naar informatie op internet door gebruikers ondervonden?

Om een antwoord te krijgen op de centrale probleemstelling van deze thesis is er een tweeledig onderzoek verricht.

Ten eerste is er voor dit onderzoek een analyse gemaakt welke diensten het meest gebruikt en/of toonaangevend zijn in het ondersteunen van internetgebruikers wanneer deze op zoek zijn naar informatie op internet.

Deze toonaangevende diensten zijn vervolgens geanalyseerd en aan de hand van hun werkwijze gecategoriseerd. Hierdoor ontstond een raamwerk van vier verschillende *zoekmethoden* die de diensten boden; *zoekmachines*, *folksonomieën*, *webdirectories* en *weblogs*. De voor- en nadelen van de diensten zijn vervolgens in kaart gebracht.

Vervolgens zijn door middel van literatuuronderzoek de potentiële zoekdoelen van gebruikers gedefinieerd. Deze zijn het navigeren naar een webadres (*navigatieve zoekdoel*), zoeken naar informatie (*informatieve zoekdoel*) of een zoeken naar interacties

met het internet (*transactioneel zoekdoel*). Er is een eerste inschatting gemaakt hoe effectief de verschillende methoden zijn voor de gedefinieerde zoekdoelen.

Als tweede is er, om de benodigde data betreffende het zoekgedrag en ervaringen van internetgebruikers te verkrijgen, een gebruikersonderzoek opgesteld. Met het gebruikersonderzoek zijn de potentiële zoekdoelen, de methoden en de effectiviteit getoetst.

7.2 Resultaten

Wat betreft de keuze in zoekdiensten kan gezegd worden dat de respondenten hoofdzakelijk de zoekmachine Google gebruiken. Verdere diensten die relatief veel gebruikt worden zijn respectievelijk de weblog Nu.nl en de webdirectory Startpagina.

Wanneer er gekeken wordt naar het gebruik van zoekmethoden door de respondenten komen de volgende zaken naar voren.

Zoekmachines en webdirectories worden voor het zoeken naar een webadres, het *navigatieve* zoekdoel, het meest gebruikt.

Zoekmachines worden voor het zoeken naar inhoudelijke informatie, het *informatieve* zoekdoel, ook het meest gebruikt, met op de tweede plaats webdirectories.

Voor het behalen van een *transactioneel* zoekdoel, zoals entertainment of downloads, wordt de zoekmachine wederom als meeste gebruikt. Voor dit zoekdoel worden qua gebruik de zoekmachines op de voet gevolgd door weblogs.

Uit het onderzoek komt naar voren dat jongere respondenten (<35jaar) meer nieuwere zoekmethoden zoals weblogs en folksonomieën gebruiken. Oudere respondenten (>35 jaar) zijn conservatiever in hun keuze, vooral de webdirectory scoort binnen deze groep relatief hoog.

De mannelijke respondenten besteden meer tijd op internet, en maken van alle zoekmethoden vaker gebruik, behalve van webdirectories. Deze worden door vrouwen meer gebruikt.

Er is, op één uitzondering na, geen verband tussen hoeveel tijd een persoon op internet besteedt en de keuze van zoekmethoden. De uitzondering hierop is het gebruik van folksonomieën. Deze worden relatief gezien bovenmatig veel gebruikt door mannen die veel tijd op internet besteden.

Op de vraag hoe internetgebruikers het zoeken naar informatie en content op internet ondervinden, antwoordt de grote meerderheid dat ze hier relatief weinig of zelfs geen problemen mee ondervinden. Het zoeken naar informatie wordt ook overwegend als gemakkelijk ervaren door de respondenten. Het merendeel geeft hiernaast ook aan hier niet veel tijd voor nodig te hebben.

Er is door de respondenten duidelijk aangegeven dat het vrijwel altijd lukt om datgene te vinden wat ze zoeken.

Wanneer er gekeken wordt naar de uitkomsten van het onderzoek kan er gesteld worden dat, bij de ondervraagde internetgebruikers, de versnellende toename in de hoeveelheid en diversificatie van content weinig negatieve effecten heeft veroorzaakt. De meeste gebruikers geven aan geen last te ondervinden van een information-overload.

Hiermee lijken de scenario's die O'Reilly (1980) en McLuhan (1967) schetsen over de information-overload voor de internetgebruikers van dit onderzoek vooralsnog minder werkelijkheid te zijn geworden dan het scenario dat Benkler (2006) schetst.

De gebruikers geven ook aan vaker wel dan niet hetgeen te vinden wat ze op internet zoeken en slechts de minderheid vindt dat er te veel informatie op internet staat, waardoor men niet vindt wat men zoekt. Hiermee wordt het idee dat de vindbaarheid van kwalitatief goede informatie achteruit zal gaan met de komst van een te grote kwantiteit ten minste gedeeltelijk wordt ontkracht (Keen 2008).

Het bestaan van Andersons Long Tail (2004) wordt met dit onderzoek nogmaals bevestigd. De diensten die de respondenten aangeven te gebruiken voor het zoeken geven duidelijk weer dat deze Long Tail net zo goed opgaat voor zoekdiensten als voor de content die deze

diensten vindbaar maken. Er bestaat een lange staart aan content- en/of onderwerpspecialistische versies van de zoekmethoden die voor dit onderzoek zijn afgebakend.

Ook is er in deze thesis de suggestie gedaan om aan het door Rose (2004) en Eppler (2002) opgestelde raamwerk aan potentiële zoekdoelen van gebruikers aan het *navigationale* zoekdoel het subdoel *retrieval* toe te voegen.

7.3 Kritische reflectie en vervolgonderzoek

Een aantal aspecten van dit onderzoek zou bij een vervolgonderzoek anders of grondiger aangepakt kunnen worden. In de opzet van dit onderzoek is de gebruikersgroep bepaald door verspreiden van enquêtes via de zogenaamde ‘sneeuwbalmethode’ via een aantal specifieke netwerken (zie hiervoor paragraaf 3.2.2), hetgeen zou kunnen leiden tot een niet representatieve groep respondenten aangezien er geen controle is op de groep deelnemers. Door een andere manier van selectie van de respondenten is het wellicht mogelijk om een representatievere doelgroep samen te stellen. Daarnaast is de lijst van kenmerken waarop de methoden zijn beoordeeld (H3) een eerste aanzet. Deze zou in vervolg onderzoek uitgebreider kunnen worden getoetst en eventueel worden vervolledigd. Ten slotte zou in plaats of naast een enquête een onderzoeksmethode ingezet kunnen worden om inzicht te krijgen in mogelijke hiaten of nuances in het onderzoek. De enquête heeft de vooraf opgesteld vragen beantwoord, maar niet gevalideerd of dat de juiste vragen zijn geweest.

Zoals aan het begin van deze thesis reeds naar voren kwam is dit een exploratief onderzoek. De wetenschappelijke basis die voor dit onderzoek gebruikt is, is in veel gevallen verre van definitief. Hierdoor blijft er vooralsnog een ruime marge over voor discussie. Een van de punten van discussie is de tijdseigenheid van het raamwerk van Rose (2004) en Eppler (2002). Deze onderzoeken zijn reeds enkele jaren geleden gedaan. De ontwikkeling van het internet heeft echter sinds deze onderzoeken niet stilgestaan. Het internet neemt een steeds socialere rol in en er is een grote variatie aan nieuwe mogelijkheden ontwikkeld voor

en door gebruikers. Deze veranderingen zouden mogelijk dan ook een gevolg gehad kunnen hebben voor de mogelijke zoekdoelen die gebruikers proberen te behalen. Een vervolgonderzoek waarin de hedendaagsheid van deze doelen onderzocht wordt, zou dan ook een goede toevoeging kunnen zijn aan het discours.

In dit onderzoek zijn zowel de zoekmethoden als de zoekdoelen onderzocht. Een vervolgonderzoek naar de zoekmotieven van gebruikers zou, in combinatie met het voorgaande, mogelijk tot een nieuwe inzichten en een extra verdieping in de kennis over internetgebruik(ers) kunnen leiden en de Quaero Quaestio, oftewel de zoektocht naar het zoeken, worden voortgezet.

Literatuurlijst

Ackoff, R. L. 1989. From Data to Wisdom. *Journal of Applied Systems Analysis*, volume 16, p-3-9.

Anderson, C. 2004. The Long Tail. *Wired Magazine*, 12 (10).
<http://www.wired.com/wired/archive/12.10/tail.html>. Geraadpleegd 27 april 2008.

Anderson, C. 2006. *The Long Tail: Why the Future of Business Is Selling Less of More*. Hyperion.

Baarda, D. & Goede, M. de. 1997. *Methoden en technieken*. Noordhoff Uitgevers B.V..

Bateman, S., Brooks, C. & McCalla, G. 2006. *Collaborative Tagging Approaches for Ontological Metadata in Adaptive E-Learning Systems*. University of Saskatchewan, Canada. [electronische versie] http://www.win.tue.nl/SW-EL/2006/camera-ready/02-bateman_brooks_mccalla_SWEL2006_final.pdf

Bellinger, G., Castro, D. & Mills, A. 2004. Data, Information, Knowledge, and Wisdom. <http://www.outsights.com/systems/dikw/dikw.htm>. Geraadpleegd 3 augustus 2008.

Bellis, M. 1997. The History of HTML. *About.com*. Geraadpleegd op 2 maart 2009.

Benkler, Y. 2006. *The wealth of networks: how social production transforms markets and freedom*. Yale University Press, 2006.

Broder, A. 2002. A taxonomy of web search. *SIGIR Forum*, 2002, Vol. 36, No. 2. 2002, p-5. [Electronische versie].

Buresh, S. 2008. Search engine, a brief history. <http://www.searchengineguide.com/scott-buresh/askcom-search-engine-a-brief-history.php> Geraadpleegd 18 november 2008.

Carlson, C. 2003. *Information Overload, Retrieval Strategies and Internet User Empowerment*. The Good, the Bad and the Irrelevant - the user and the future of information and communication technologies. Conference Proceedings [Electronische versie]. Helsinki 2003, p-169-173. <http://eprints.rclis.org/archive/00002248/>

Centraal Bureau voor de Statistiek. 2008. *Mediaproducten steeds meer via internet*. Persbericht PB08-071 [Electronische versie].
<http://www.cbs.nl/NR/rdonlyres/B6BD7026-9407-4667-9E8E-F38072D8FE96/0/pb08n071.pdf>

Checkit.nl. 2008. *Checkit Nationale Search Engine Monitor December 2008*.
<http://www.checkit.nl/nationalesearchenginemonitor.html> Geraadpleegd 11 december 2008.

Eppler, M.J., Mengis, J. 2004. The Concept of Information Overload - A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines. *The Information Society: An International Journal*, 20(5), 2004, p-1-20.

Fowler, F.J. 2002. *Survey Research Methods*. Thousand Oaks/Sage Publications Inc.

Frissen, V. et al. 2008. *Naar een 'User Generated State'? De impact van nieuwe media voor overheid en openbaar bestuur*. TNO in opdracht van Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, p-26.

Fructuoso van der Veen, V.A. et al. 2007. *De digitale economie*. [Electronische versie]. Centraal Bureau voor de Statistiek, Voorburg.

Gantz, J. F. et al. 2008. The Diverse and Exploding Digital Universe. *EMC2*, p-3. [Electronische versie]. <http://www.emc.com/collateral/analyst-reports/diverse-exploding-digital-universe.pdf>

Giles, J. 2005. Internet encyclopaedias go head to head. *Nature*, 438, p-900-901.

Gray, M. 1995. Measuring the growth of the web. *Massachusetts Institute of Technology (MIT)*. <http://www.mit.edu/people/mkgray/growth/>

Haan, J. de. 2004. ICT en Samenleving. *In het zicht van de toekomst, Sociaal en Cultureel Rapport 2004*. Sociaal Cultureel Planbureau (SCP), p-179-222.

Hair J.F., Bush R.P. & Ortinau, D.J. 2003. *Marketing research: within a changing information environment*. McGraw-Hill/Irwin, New York.

Hansen, A. 1998. *Mass Communication Research Methods*. Palgrave, Basingstroke.

Hodge, G. 2001. Metadata Made Simpler. *National Information Standards Organization (NISO)* [Electronische versie], p-3. http://scigate.ncsi.iisc.ernet.in/raja/is206/sources/Metadata_Bklt.pdf

Internet Assigned Numbers Authority (IANA). 2007. *MIME Media Types* <http://www.iana.org/assignments/media-types/> Geraadpleegd 12 oktober 2008.

Kock, N.F., McQueen, R. & Corner, J. 1997. The Nature of Data, Information and Knowledge Exchanges in Business Processes: Implications for Process Improvement and Organisational Learning. *PACIS 1997 Proceedings*. Artikel 74. <http://www.cis.temple.edu/~kock/public/tlo97/kiflww6.htm> Geraadpleegd 6 juli 2008.

Laertius, D. (469 BC - 399 BC). *Lives of Eminent Philosophers*, volume 1, Loeb Classical Library, Harvard University Press, 1972.

Martijn, M. & Nihom, B. (red.). 2008. Andrew Keen: 'Experts, daar gaat het om in onze cultuur'. *Vrij Nederland*.
<http://www.vn.nl/KunstCultuur/FilmMedia/ArtikelFilmMedia/AndrewKeenExperts,DaarGaatHetOmInOnzeCultuur.htm>

Mathes, A. 2004. Folksonomies - Cooperative Classification and Communication Through Shared Metadata. USA, p-1.
<http://adammathes.com/academic/computermediated-communication/folksonomies.pdf>
Geraadpleegd 25 oktober 2008.

McLuhan, M. 1967. *The Best of Ideas on CBC Radio. Canada*.

Microsoft Technet. 2002. *Glossary of Networking Terms for Visio IT Professionals*.
[http://technet.microsoft.com/nl-nl/library/cc751329\(en-us\).aspx](http://technet.microsoft.com/nl-nl/library/cc751329(en-us).aspx) Geraadpleegd 10 oktober 2008.

O'Reilly, C.A. 1980. Individuals and information overload in organizations: Is more necessarily better? *Academy of Management Journal* 23, p-684–696.

O'Reilly, T. 2005. What Is Web 2.0.
<http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>.
Geraadpleegd 3 juni 2008.

Poortman, S. & Bierens, G. 2005. *InformatieProfessional*. Jaargang 9, nummer 7/8 juli/augustus 2005.

Put, E. 2008. Soorten websites: gids voor alle website types. *Edwinsdesignlab*.
<http://www.edwinsdesignlab.be/blog/2008/09/17/soorten-websites-gids-voor-alle-website-types/> Geraadpleegd 10 mei 2008.

Rose, D.E. & Levinson, D. 2004. Understanding user goals in Web search. *Proceedings of the Thirteenth International World Wide Web Conference 2004*.

Rosen, L. & Weil, M. 1997. *Technostress: Coping with Technology @ Work, @ Home, @ Play*. John Wiley & Sons, 1997.

Schick, A.G., Gordon, LA. & and Haka, S. 1990. Information overload: a temporal approach. *Accounting, Organizations and Society*, 15, 3 (1990), p-199-220.

Slot, M. & Frissen, V. 2007. Users In The 'Golden' Age Of The Information Society. *Observatorio (OBS*) Journal*, 3 (2007), 201-224, p-201.

Smith, A. 1776. *The Wealth of Nations; An Inquiry into the Nature and Causes of the Wealth of Nations*.

Tancer, B. 2008. Sizing up the Long Tail of Search.
http://weblogs.hitwise.com/bill-tancer/2008/11/sizing_up_the_long_tail_of_sea.html.
Geraadpleegd 22 november 2008.

Technische Universiteit Delft (Tudelft). *Glossary*.
<http://www.library.tudelft.nl/tulib/glossary/index.htm> Geraadpleegd 20 juli 2008.

Toffler, A. 1981. *The Third Wave*. New York, Bantam Books.

Vanderwal, T. 2005. Off the Top: Folksonomy Entries.
<http://www.vanderwal.net/random/category.php?cat=153> Geraadpleegd 12 september 2008.

Vickery, G. & Wunsch-Vincent, S. 2007. Participative Web And User-Created Content: Web 2.0 Wikis and Social Networking. *Source OECD Science & Information Technology*, 15, 2007, p-18.

Wall, A. 2006. *How Search Engines Work: Search Engine Relevancy Reviewed*.
www.seobook.com Geraadpleegd 23 juni 2008.

Weinberger, D. 2005. *Tagging and why it matters*. Harvard.
<http://cyber.law.harvard.edu/home/2005-07>. Geraadpleegd 10 november 2008.

WGBH. 2007. Open Content and Public Broadcasting. *Glossary*.
<http://opencontent.wgbh.org/report/glossary.html> Geraadpleegd 18 september 2008.

Wierenga, U. 2008. *Marktaandeel Zoekmachines 2007/2008*. <http://www.gintou.nl>.
Geraadpleegd 8 oktober 2008.

Overige websites

<http://www.Startpagina.nl>
<http://www.Craigslist.com>
<http://www.Google.com/dirhp>
<http://www.Google.com>
<http://www.Ilse.nl>
<http://dir.Yahoo.com>
<http://www.Altavista.com>
<http://www.Ask.com>
<http://www.Live.com>
<http://search.Yahoo.com>
<http://www.Nu.nl>
<http://www.Spitsnieuws.nl>
<http://www.Nujij.nl>
<http://www.delicious.com>
<http://www.Digg.com>
<http://www.Stumbleupon.com>
<http://www.Technorati.com>
<http://www.Alexa.com>
<http://www.Tastekid.com>
<http://www.Blinx.com>
<http://www.Podscope.com>
<http://www.Multiscope.nl>
<http://www.Wordle.net>

Bijlagen

B1. Lijst van figuren en tabellen

Figuren

<i>Figuur 1: Schematische weergave werkwijze zoekmethoden</i>	14
<i>Figuur 2: Hoeveelheid geproduceerde digitale informatie in exabytes</i>	19
<i>Figuur 3: The Long Tail</i>	21
<i>Figuur 4: Overzicht van de potentiële zoekdoelen en subdoelen van een zoekopdracht</i>	35
<i>Figuur 5: Verhouding tussen het geslacht van de respondenten</i>	45
<i>Figuur 6: Screenshot Google.com. (januari 2009)</i>	50
<i>Figuur 7: Screenshot search.yahoo.com. (januari 2009)</i>	51
<i>Figuur 8: Screenshot Live.com. (januari 2009)</i>	52
<i>Figuur 9: Screenshot Ask.com. (januari 2009)</i>	53
<i>Figuur 10: Screenshot Lycos.com. (januari 2009)</i>	54
<i>Figuur 11: Screenshot Altavista.com. (januari 2009)</i>	55
<i>Figuur 12: Screenshot Ilse.nl. (januari 2009)</i>	55
<i>Figuur 13: Resultaten Google.com, zoekterm: Google, resp. 12 december 2008 en 12 maart 2009</i>	56
<i>Figuur 14: Screenshot Delicious.com (januari 2009)</i>	61
<i>Figuur 15: Screenshot Digg.com (januari 2009)</i>	62
<i>Figuur 16: Screenshot Stumbleupon.com (januari 2009)</i>	63
<i>Figuur 17: Screenshot Nujij.nl (januari 2009)</i>	64
<i>Figuur 18: Bookmarkteller website Delicious</i>	65
<i>Figuur 19: Voorbeeld van een tagcloud</i>	65
<i>Figuur 20: De mogelijkheid tot het zoeken in andermans favorieten</i>	66
<i>Figuur 21: Spelfout in titel Digg.com</i>	66
<i>Figuur 22: Een voorbeeld van generieke metadata</i>	67
<i>Figuur 23: Voorbeeld discrepantie tussen verwijzing en beschrijving</i>	67
<i>Figuur 24: Screenshot Startpagina.nl (Januari 2009)</i>	71
<i>Figuur 25: Screenshot Graigslist.com (Januari 2009)</i>	72
<i>Figuur 26: Screenshot DMOZ (Januari 2009)</i>	73
<i>Figuur 27: Screenshot Yahoo.com (Januari 2009)</i>	74
<i>Figuur 28: Screenshot Nu.nl (januari 2009)</i>	79
<i>Figuur 29: Screenshot Spitsnieuws.nl (januari 2009)</i>	80
<i>Figuur 30: Overzicht van potentiële krachten van ontsluitingsmethoden</i>	84
<i>Figuur 31: Frequentie van leeftijd respondenten</i>	85
<i>Figuur 32: Frequentie van gemiddeld aantal uren internegebruik per week</i>	86
<i>Figuur 33*: Wordcloud met een overzicht van de antwoorden op de open vraag uit de survey.</i>	88
<i>Figuur 34: Frequentie van gebruik zoekdiensten</i>	90
<i>Figuur 35: Gebruiksfrequentie zoekmethoden.</i>	91
<i>Figuur 36: Verschillen in frequentie gebruik zoekmethoden mannen en vrouwen</i>	92
<i>Figuur 37: Frequentie in gebruik zoekmethode gerangschikt aan de hand van leeftijdsklassen</i>	93
<i>Figuur 38: Frequentie van gebruik zoekmethode gerangschikt aan de hand van het aantal uren dat per week op internet besteed wordt</i>	94
<i>Figuur 39: De meest gebruikte zoekdiensten door respondenten</i>	95
<i>Figuur 40: Overzicht van geantwoorde diensten in gebruikersonderzoek gerangschikt aan de hand hun methode van informatieontsluiting</i>	96
<i>Figuur 41: Relatief gebruikpercentage van zoekmethoden voor het behalen van het navigationeel zoekdoel</i>	99
<i>Figuur 42: Relatief gebruikpercentage van zoekmethoden voor het behalen van het informationeel zoekdoel</i>	100
<i>Figuur 43: Relatief gebruikpercentage van zoekmethoden voor het behalen van het transactioneel zoekdoel</i>	102

<i>Figuur 44: Sterkten zoekmachines versus gebruik door respondenten</i>	108
<i>Figuur 45: Sterkten folksonomieën versus gebruik door respondenten</i>	110
<i>Figuur 46: Sterkten webdirectories versus gebruik door respondenten</i>	112
<i>Figuur 47: Sterkten weblogs versus gebruik door respondenten</i>	114

Tabellen

<i>Tabel 1: Lijst van geanalyseerde diensten.</i>	40
<i>Tabel 2: Voor- en nadelen van werkwijze zoekmachines</i>	59
<i>Tabel 3: Waardering van de potentiële competentie van zoekmachines voor het behalen van zoekdoelen</i>	59
<i>Tabel 4: Voor- en nadelen van werkwijze folksonomieën</i>	69
<i>Tabel 5: Waardering van de potentiële competentie van folksonomieën voor het behalen van zoekdoelen</i>	70
<i>Tabel 6: Voor- en nadelen van werkwijze webdirectories</i>	77
<i>Tabel 7: Waardering van de potentiële competentie van webdirectories voor het behalen van zoekdoelen</i>	78
<i>Tabel 8: Voor- en nadelen werkwijze weblogs</i>	82
<i>Tabel 9: Waardering van de potentiële competentie van weblogs voor het behalen van zoekdoelen</i>	83
<i>Tabel 10: Resultaten reactie respondenten op voorgelegde stellingen</i>	87
<i>Tabel 11: Naamsbekendheid zoekdiensten onder respondenten</i>	89

B2. Topiclist dienstenanalyse

Topiclijst Dienstenanalyse

Diensttype

Zoekmachine, Folksonomie, Webdirectory, Weblog

Naam dienst

Eigenaar dienst

Aansturing dienst

Aantal bezoekers

Historie

Taal

Oorsprong

Omvang

Content

Marktaandeel

Visueel:

Screenshot

Visuele opvallendheden

Zoekopties

Zichtbare zoekopties

Tekst

Taal

Afbeelding

Video

Audio

Personen

Bedrijven

Overige

Verdere zoekopties

Overige

Resultaten

Afkomst

Bronnen

Samenwerkingsverbanden/ Diensten

Methoden

Overige

Technieken

Overige opmerkingen

Relevantie

Bepaling

Populariteit, chronologisch, voorkomendheid

Beïnvloeding

Kleuring

Commercieel/Gesponsord

Doelgroep/Volggroep

Overige opmerkingen

-

B3. Survey

IOThesis2008

pagina1

*** 1. Welke van deze diensten ken je om op internet informatie te zoeken en/of vinden en hoe vaak gebruik je ze?**

(Ontbreken jouw favoriete diensten in deze lijst? Zet deze er dan in het lijstje onder deze vraag bij!)

	Nooit van gehoord	Gebruik ik zelden tot nooit	Gebruik ik soms	Gebruik ik regelmatig	Gebruik ik vaak	Gebruik ik altijd
Nu.nl	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Spitsnieuws.nl	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
lIse.nl	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Vindex.nl	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Startpagina.nl	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Craigslist.com	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
DMoz/ODP/Google Directory (google.com/dirhp)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Yahoo Directory (dir.yahoo.com)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Nuijij.nl	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Del.icio.us	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Digg.com	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Yahoo.com	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Stumbleupon	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Google.com	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Lycos.nl	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Altavista.com	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ask.com	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Microsoft Live Search (live.com)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

2. (vervolg vraag 1)

Hier kan je eventueel ontbrekende diensten invullen:

-

-

-

3. .

	Gebruik ik zelden tot nooit	Gebruik ik soms	Gebruik ik regelmatig	Gebruik ik vaak	Gebruik ik altijd
-	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
-	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
-	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

4. Welke 3 diensten uit de vorige vraag, inclusief eventueel jouw eigen bijdragen, gebruik je (het meest)?

Onthoud even welke je invult, je hebt ze nodig bij de volgende vraag!

(Gebruik je minder dan 3 van deze diensten, vul dan niets in bij het desbetreffende vakje.)

1.
2.
3.

5. Er volgt nu drie keer (dezelfde) lijst met 9 vragen. Wil je zo vriendelijk zijn deze beantwoorden voor de diensten in jouw top 3?

Bij minder dan drie diensten bij de vorige vraag vul je alleen voor die diensten de 9 vragen in.

De dienst die ik op internet het MEEST gebruik om informatie te vinden, gebruik ik omdat..:
(-- nooit/++ altijd)

	--	-	+/+	+	++
...ik weet waar ik heen wil, alleen ik weet het precieze web-adres(URL) niet of het is gemakkelijker op deze manier. (Bv. De website van Coca Cola, Erasmus Universiteit)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...ik iets specifiek zoek. (Waarom glimt metaal?, Wat is het speelschema voor het EK?)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...ik algemene kennis wil opdoen over een onderwerp. (Bv. Ik wil meer weten over het menselijk lichaam.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...ik op zoek ben naar advies, ideeën, suggesties of instructies. (Bv. Hoe maak ik een gerecht?, Hoe kan ik stoppen met roken?)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...ik wil weten of/ waar waar een product of dienst te verkrijgen is in de 'echte' wereld. (Waar kan ik Nike Air Max maat 42 kopen?, Zit er een kapper in mijn woonwijk?)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...ik iets wil downloaden dat op mijn computer (of ander apparaat) moet staan om nuttig te zijn. (Limewire/Kazaa, Firefox)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...ik vermaakt wil worden, simpelweg door de inhoud van de site te bekijken. (Een webcam, foto's, grappige filmpjes, erotiek)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...ik een site zoek, die mij in staat stelt om actueel of interactief een ander programma/dienst op internet te gebruiken. (radarbeelden van knmi op buienradar.nl, valuta omrekenen met google calculator, een tele-atlas routeplanner op goudengids.nl)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

6. Dit is het tweede deel van de lijst met negen vragen:

De dienst die op de **TWEEDE** plaats komt als het gaat om hoevaak ik deze gebruik om informatie te vinden op internet, gebruik ik omdat..:

(-- nooit/++ altijd)

	--	-	-/+	+	++
...ik weet waar ik heen wil, alleen ik weet het precieze web-adres(URL) niet of het is gemakkelijker op deze manier. (Bv. De website van Coca Cola, Erasmus Universiteit)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...ik iets specifieks zoek. (Waarom glimt metaal?, Wat is het speelschema voor het EK?)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...ik algemene kennis wil opdoen over een onderwerp. (Bv. Ik wil meer weten over het menselijk lichaam.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...ik op zoek ben naar advies, ideeën, suggesties of instructies. (Bv. Hoe maak ik een gerecht?, Hoe kan ik stoppen met roken?)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...ik wil weten of/ waar waar een product of dienst te verkrijgen is in de 'echte' wereld. (Waar kan ik Nike Air Max maat 42 kopen?, Zit er een kapper in mijn woonwijk?)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...ik iets wil downloaden dat op mijn computer (of ander apparaat) moet staan om nuttig te zijn. (Limewire/Kazaa, Firefox)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...ik vermaakt wil worden, simpelweg door de inhoud van de site te bekijken. (Een webcam, foto's, grappige filmpjes, erotiek)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...ik een site zoek, die mij in staat stelt om actueel of interactief een ander programma/dienst op internet te gebruiken. (radarbeelden van knmi op buienradar.nl, valuta omrekenen met google calculator, een tele-atlas routeplanner op goudengids.nl)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

7. Dit is het derde deel van de lijst met negen vragen.

Hou nog héél even vol, alleen deze nog en een paar stellingen, dan ben je klaar.

De dienst die op de DERDE plaats komt als het gaat om hoe vaak ik deze gebruik om informatie te vinden op internet, gebruik ik omdat..:

(-- nooit/++ altijd)

	--	-	-/+	+	++
...ik weet waar ik heen wil, alleen ik weet het precieze web-adres(URL) niet of het is gemakkelijker op deze manier. (Bv. De website van Coca Cola, Erasmus Universiteit)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...ik iets specifiek zoek. (Waarom glimt metaal?, Wat is het speelschema voor het EK?)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...ik algemene kennis wil opdoen over een onderwerp. (Bv. Ik wil meer weten over het menselijk lichaam.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...ik op zoek ben naar advies, ideeën, suggesties of instructies. (Bv. Hoe maak ik een gerecht?, Hoe kan ik stoppen met roken?)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...ik wil weten of/ waar waar een product of dienst te verkrijgen is in de 'echte' wereld. (Waar kan ik Nike Air Max maat 42 kopen?, Zit er een kapper in mijn woonwijk?)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...ik iets wil downloaden dat op mijn computer (of ander apparaat) moet staan om nuttig te zijn. (Limewire/Kazaa, Firefox)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...ik vermaakt wil worden, simpelweg door de inhoud van de site te bekijken. (Een webcam, foto's, grappige filmpjes, erotiek)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...ik een site zoek, die mij in staat stelt om actueel of interactief een ander programma/dienst op internet te gebruiken. (radarbeelden van knmi op buienradar.nl, valuta omrekenen met google calculator, een tele-atlas routeplanner op goudengids.nl)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

*** 8. In hoeverre ben jij het eens met de volgende stellingen?**
 (-- oneens/ ++ eens)

	--	-	-/+	+	++
Als ik iets zoek op internet, kost het me veel tijd om het te vinden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ik kan met Google alles vinden wat ik nodig heb.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Als ik iets zoek op internet dan ga ik altijd hetzelfde te werk.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mijn omgeving (vrienden/collega's etc.) kunnen mij beter vertellen waar ik iets op internet kan vinden dan google.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ik laat de manier waarop ik zoek afhangen van wat ik zoek.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Er staat te veel informatie op internet, hierdoor kan ik vaak niet meer vinden wat ik zoek.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Er staat erg veel informatie op internet, maar er in zoeken is gemakkelijk.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Als ik iets zoek op internet, lukt het me altijd om dit te vinden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

9. Hoe vind jij je weg door de grote hoeveelheid informatie?

10. Je bent nu echt bijna klaar!

Alleen nog even wat informatie over jou:

Wat is je leeftijd?

Wat is je geslacht
(m/v)

Hoeveel uur per week
zit je op internet?
(werk + privé)

Wil je op de hoogte
gehouden worden van
het onderzoek? Vul
dan hier je
emailadres in. (niet
verplicht)