

Concurrerend commentaar

Een methodevergelijkend onderzoek
tussen de commentaren op een
feedbackbutton en Infocus

P.A.V. Hubbers
3134490

Masterscriptie Communicatiestudies
Universiteit Utrecht 2010

Begeleider: Leo Lentz

P.A.V. Hubbers - 3134490
Communicatiestudies
Faculteit Geesteswetenschappen
Universiteit Utrecht
Augustus 2010

Voorwoord

Na mijn masterstage bij communicatieadviesbureau Pastoor BV, waarbij ik een meetsysteem voor het evalueren van gemeentelijke websites mocht analyseren en herontwikkelen, ben ik erg geïnteresseerd geraakt in dat onderzoeksonderwerp. Omdat mijn stagebegeleider Sanne Elling dit opmerkte en door kon spelen aan Leo Lentz, leverde laatstgenoemde mij het thema voor dit huidige masteronderzoek: een methodevergelijkend onderzoek tussen twee verschillende evaluatiemethoden voor gemeentelijke websites.

Wat mij erg bevallen is in het werken met gemeentelijke websites is de maatschappelijke relevantie. Naast het wetenschappelijk onderzoek kreeg ik het gevoel echt dienstverlenend werk te verrichten, wat inzicht geeft waar gemeentes daadwerkelijk verder mee kunnen.

Het was erg leerzaam me te verdiepen in de twee evaluatiemethoden: de feedbackbutton en Infocus. Door vergelijkend onderzoek te doen, komen erg helder de plus- en minpunten van beide methoden naar voren. Ik heb met veel plezier mogen werken aan dit masterverslag. Mijn begeleider Leo Lentz heeft mij de vrijheid gegeven waarin ik graag werk. Hij leverde zinvolle ideeën aan waar ik goed mee verder kon werken. Bij deze wil ik hem daar hartelijk voor bedanken. Ook het contact met medewerkers van gemeente Utrecht heeft veel inzicht gegeven in de feedbackfaciliteit van de website. Tot slot wil ik ook mijn vriend Mark bedanken voor het vertrouwen in mij van het begin tot het eind. Hij heeft me goed geholpen en erg gemotiveerd.

Er kan nog veel ontdekt worden wat betreft het evalueren van gemeentelijke websites. Ik ben erg blij en dankbaar dat ik deel heb mogen uitmaken van het zoekproces naar de optimale methode.

Samenvatting

In dit onderzoek is gekeken in hoeverre de webevaluatiemethoden de feedbackbutton en Infocus van elkaar verschillen. Door het commentaar van beide methoden met elkaar te analyseren en te vergelijken is een congruente validering gemaakt. De volgende variabelen zijn daarbij betrokken geweest: dimensie, relevantie, kosten en de kwaliteit van de verbaliseringen.

Wat betreft de dimensie blijkt het commentaar van Infocus zich op veel meer dimensies te richten dan dat van de feedbackbutton. Daar waar de reacties op de feedbackbutton vooral gaan over informatie die niet gevonden kan worden en links die niet werken, gaan de reacties op Infocus daarnaast ook over onder andere een onduidelijke vormgeving, onopvallende links en een onverzorgde tekst. Bij Infocus wordt een groot aantal kleine problemen door een klein aantal mensen benoemd, terwijl bij de feedbackbutton een klein aantal grote problemen door een grote groep websitebezoekers gemeld wordt. Dit brengt ons direct bij de relevantie van de reacties.

Omdat de feedbackbutton echte gebruikersproblemen toont en Infocus ook kleine oneffenheden meet, hebben de reacties op de feedbackbutton een grotere maat van ernst. Deze berichten zijn relevanter voor de kwaliteit van de website. De problemen die aangekaart worden in deze berichten moeten daardoor ook vaak eerder opgelost worden. Dit brengt ons dan weer bij de kosten.

Veel genoemde punten bij Infocus hebben een lage prioriteit, dus hoeven niet direct verholpen te worden. Daarnaast is het direct duidelijk waar de punten zich op focussen dus is het traceren (en verhelpen) van de problemen wellicht geen lastige en dure klus. Bij de feedbackbutton ligt dat anders. Hier moet eerst een verdeling worden gemaakt tussen de relevante en irrelevante berichten en worden alle berichten van bezoekers beantwoord. Qua kosten aan werknemers binnen de gemeente is de feedbackbutton dus een stuk duurder uit. Echter, bij Infocus worden de proefpersonen betaald. De websitegebruikers die uit zichzelf een reactie plaatsen op de feedbackbutton doen dit vrijwillig. Dit brengt ons tot slot bij de kwaliteit van de verbaliseringen.

De reacties op de feedbackbutton zijn voornamelijk in de vorm van klachten, omdat bezoekers iets niet kunnen vinden. De reactie komt dus puur voort uit eigenbelang. Het commentaar van de proefpersonen bij Infocus is vooral in de vorm van opmerkingen en tips, omdat zij de rol van evaluator hebben toegewezen gekregen.

Inhoudsopgave

<u>Inleiding</u>	8
<u>Theoretisch kader</u>	12
- <u>Feedbackbutton</u>	12
- <u>Infocus</u>	14
<u>Onderzoeksopzet</u>	18
- <u>Feedbackbutton</u>	18
- <u>Infocus</u>	19
- <u>Stappenplan</u>	19
- <u>Congruente validering</u>	20
<u>Resultaten</u>	22
- <u>Feedbackbutton</u>	22
- <u>Infocus</u>	23
- <u>Congruente validering</u>	25
<u>Parkeren</u>	27
<u>Zwemmen</u>	31
<u>Afval</u>	36
<u>Afname</u>	41
<u>Verwerking</u>	41
<u>Afhandeling</u>	41
<u>Conclusie</u>	48
<u>Discussie</u>	52
<u>Literatuur</u>	56
<u>Bijlagen</u>	

Inleiding

Urenlang zoeken op internet naar dat ene antwoord. Die bron van informatie waar alles toch snel te vinden zou moeten zijn. Helaas is dit niet altijd het geval en zorgen websites vaak voor frustratie. Een website waar informatie bovenal snel en gemakkelijk gevonden moet worden is de gemeentelijke website. Het aanvragen van vergunningen, controleren welke evenementen er plaatsvinden, nakijken hoe duur de dichtstbijzijnde parkeergarage kost: een gemeentelijke website moet op diverse terreinen op simpele wijze duidelijkheid bieden aan burgers. Aan inwonenden en toeristen. Aan ondernemers en werkzoekenden.

Er zijn diverse manieren om te controleren in hoeverre een gemeentelijke website voldoet. Het evalueren kan overgelaten gelaten worden aan experts, waarbij we spreken over een *expertgerichte evaluatie*. Zo zijn er verschillende richtlijnen opgesteld, ook wel heuristische genoemd, bijvoorbeeld vanuit de overheid. Hierbij scoort een professional de website aan de hand van de opgestelde punten, die vaak ook gecreëerd zijn door experts op ict-gebied. In alle richtlijnen komt globaal naar voren dat de bezoeker snel en eenvoudig terecht moet komen op de pagina waar de voor hem relevante informatie beschikbaar is. Door een goede structuur aan te bieden in de vorm van een duidelijke en uniforme navigatie, moet dat volgens de Stijlgids van EGEM (kennisplatform voor gemeentelijke vraagstukken rond informatievoorziening, e-dienstverlening en organisatieontwikkeling) mogelijk zijn (<http://www.egem-iteams.nl/stijlgids-gemeenten-03>). Ook uit het onderzoek van Van den Haak, de Jong en Schellens (2007) blijkt dat probleem nummer 1 van websitebezoekers een lastige structuur is.

Maar ook het oordeel van 'echte' websitegebruikers is relevant. Hier spreken we van *doelgroepgerichte evaluatie*. In doelgroepgericht onderzoek naar de kwaliteit van websites staan vijf methoden centraal. Deze kunnen worden onderscheiden op de dimensie 'in-use' versus 'non-use' en de dimensie laboratoriumonderzoek versus on-line onderzoek (www.gemeentewebsites.nl). Het verschil tussen in-use en non-use methoden ligt in de taak van de evaluator: in het eerste geval gebruikt hij de tekst zoals hij dat normaal ook zou doen, in het tweede geval is de evaluator een beoordelaar van de tekst. Bij laboratoriumonderzoek bevindt de evaluator zich in een testsituatie, bij online-onderzoek bevindt de evaluator zich in een natuurlijke gebruikerssituatie. Dit leidt tot vier typen doelgroepgerichte methoden, waarbinnen vijf methoden ingedeeld kunnen worden. Onderstaande tabel toont deze typen en de bijbehorende methoden.

	In-use methoden	Non-use methoden
Laboratoriumonderzoek	Hardopdenkmethode Oogbewegingsregistratie	Infocus
Online onderzoek	Feedbackbutton	Vragenlijst

Tabel 1: Overzicht doelgroepgerichte evaluatiemethoden

In dit onderzoek zal gefocust worden op twee van bovenstaande gebruikersgerichte evaluatiemethoden: *de feedbackbutton* en *Infocus*. Beide methoden richten zich op de ervaring van de gebruikers van de website. Omdat gemeentelijke websites gebruiksvriendelijk en toegankelijk willen en dienen te zijn, is onderzoek naar de ervaring van de gebruiker een zeer geschikte methode om dit te onderzoeken. Naar de twee bovengenoemde methoden is nog maar weinig onderzoek gedaan en er is relatief nog maar weinig literatuur over bekend. Door de twee evaluatiemethoden met elkaar te vergelijken zal inzicht gegeven worden in de methoden an sich en kunnen de voor- en nadelen ten opzichte van elkaar benoemd worden.

Voor dit onderzoek zal commentaar op de gemeentelijke website van Utrecht gebruikt worden. Het commentaar van de feedbackbutton van www.utrecht.nl, aangeduid op de website als 'Reageer op deze pagina', zal geanalyseerd worden en vervolgens vergeleken worden met het commentaar van Infocus op www.utrecht.nl, verzameld door Henssen (2010). Hierbij staat de volgende onderzoeksvraag centraal:

Welk commentaar brengen beide evaluatiemethoden teweeg bij de gebruikers?

Hierbij zullen voor beide corpora de volgende deelvragen een uitgebreid antwoord op de hoofdvraag vormen:

Op welke dimensies richt het commentaar zich?

In hoeverre is het commentaar relevant?

Wat zijn de kosten en opbrengsten van beide meetmethoden?

Is er een verschil in kwaliteit van de verbalisering van de gebruikers en proefpersonen?

Bovenstaande deelvragen worden in de Onderzoeksopzet uitgebreid toegelicht. De antwoorden op de deelvragen bieden meer inzicht in de twee evaluatiemethoden, waardoor voor vervolgonderzoek naar webevaluatie en de kwaliteit van websites een betere keuze kan worden gemaakt.

Dit onderzoek geeft inzicht in de voor- en nadelen van beide meetmethoden; welke methode roept welk commentaar wel en welk commentaar niet op? Daarnaast biedt dit onderzoek praktische informatie voor gemeentes. Er kan aan de hand van deze gegevens advies gegeven worden welk commentaar zij

kunnen verwachten via beide meetmethoden en welk commentaar zij dus vrijwel niet kunnen/zullen ontvangen via een methode ten opzichte van de andere evaluatiemethode.

Om bovenstaande onderzoeksvragen te beantwoorden zullen eerst beide evaluatiemethoden onder de loep genomen worden in het Theoretisch kader. Vervolgens zal de Onderzoeksmethode beschreven worden. Daarna volgt in de Resultaten een beschrijving en analyse van het commentaar van de feedbackbutton. Dit wordt gevolgd door een beschrijving en analyse van het verzamelde Infocuscommentaar van Henssen (2010). Vervolgens zal de vergelijking tussen beide evaluatiemethoden uitgeschreven worden in de congruente validering, waarna de Conclusie en Discussie dit onderzoek zullen sluiten.

Theoretisch kader

In dit Theoretisch kader zullen de feedbackbutton van www.utrecht.nl en het evaluatiesysteem Infocus aan bod komen. Hierbij worden de eerdere onderzoeken van Vu Anh (2007) en Henssen (2010) betrokken als aanleiding. Vu Anh heeft zich gericht op de feedbackbutton van www.utrecht.nl en enkele aanbevelingen voor een verbeterd ontwerp gegeven. Henssen heeft een elftal pagina's van www.utrecht.nl geanalyseerd met Infocus. Met behulp van de gegevens van beide onderzoeken wordt een uitgebreid beeld gegeven van de gemeentelijke website van Utrecht. Deze informatie zal uitmonden in de vraag welke informatie beide meetmethodes daadwerkelijk geven en waarin zij verschillen en overeenkomen. De verschillen en overeenkomsten in het commentaar van beide meetmethodes komen in latere hoofdstukken aan bod.

Feedbackbutton

Zoals in de Inleiding beschreven staat, is de feedbackbutton een in-use meetmethode die online wordt afgenomen. Dit houdt in dat de gebruiker de webpagina zelfstandig gebruikt in plaats van enkel beoordeelt. Daarnaast bevindt deze gebruiker zich in een natuurlijke situatie en beoordeelt hij de website vrijwillig.

De feedbackbutton op www.utrecht.nl staat op elke pagina, afgezien van de homepage. De rode pijl op onderstaand bewerkte afbeelding toont de feedbackbutton ('Reageer op deze pagina').



Afbeelding 1: Feedbackbutton op www.utrecht.nl

Wanneer een bezoeker op de feedbackbutton klikt komt hij in het volgende scherm terecht:

The screenshot shows the 'Reactie' (Feedback) page on the website of Gemeente Utrecht. The page layout includes a navigation menu on the left, a search bar at the top right, and a central form for submitting feedback. The form fields are labeled 'Naam:*', 'E-mail:*', and 'Reactie:*'. A red banner indicates that fields with an asterisk are required. The page also features contact information for Gemeente Utrecht on the right side, including the address, post address, telephone number, and email link. The footer contains links for 'Colofon' and 'Proclaimer'.

Afbeelding 2: Reactiepagina op www.utrecht.nl

Per dag komen er gemiddeld tien berichten binnen. Alle verstuurd berichten komen in tweevoud terecht bij gemeente Utrecht. Elk bericht wordt één keer verzonden naar de desbetreffende medewerker die zich bezighoudt met de pagina waarop gereageerd is. Dit noemt gemeente Utrecht de decentrale bank. Binnen gemeente Utrecht werken plusminus 200 werknemers die zich bezighouden met de ontwikkeling en content van www.utrecht.nl. De functies van deze werknemers verschillen echter enorm. De gemeentelijke website van Utrecht heeft ruim 8000 webpagina's, waarvan elke werknemer het zicht houdt over één tot 100 pagina's. Daarnaast wordt elk bericht één keer verstuurd naar de centrale bank, alwaar alle berichten die binnenkomen opgeslagen worden. Met de informatie uit deze bank kan gewerkt worden voor projecten zoals dit huidige onderzoek.

Vu Anh (2007) heeft eerder al onderzoek gedaan naar de feedbackfaciliteit van gemeente Utrecht. Zij zou het commentaar van de websitebezoekers in haar onderzoek aanvankelijk categoriseren aan de hand van drie hoofdcategorieën *Inhoud*, *Bereikbaarheid* en *Vormgeving* van de Web Evaluation Questionnaire (WEQ) (<http://www2.hum.uu.nl/onderzoek/weg/index.html>). De categorie *Inhoud* is in de WEQ onderverdeeld in Relevantie, Begrijpelijkheid en Volledigheid, de *Bereikbaarheid* in Gebruikersgemak, Hyperlinks, Structuur, Snelheid en Zoekmachine. De categorie *Vormgeving* bevat geen enkele subcategorieën. Echter, deze protocolanalyse bleek in het onderzoek van Vu Anh te beperkt, omdat de verkregen reacties veel meer

dimensies en verschillende onderwerpen bleken te betreffen. Er is in haar onderzoek besloten dat de protocolanalyse zo dicht mogelijk bij de reacties zelf moet staan. Daarom is er een andere categorieverdeling gemaakt tijdens het analyseren van het feedbackcommentaar. Onderstaande tabel toont de door Vu Anh gecreëerde rubrieken met bijbehorende classificaties:

Rubrieken	Classificaties
Informatie	1. Informatie is niet te vinden 2. Informatie is niet te vinden (via zoekmachine) 3. Informatie is niet correct 4. Informatie is niet duidelijk 5. Informatie is niet volledig 6. Informatie wekt verwachtingen die niet worden waargemaakt.
Navigatie	7. De link werkt niet 8. De link moet aangepast/gewijzigd
Verzoek	9. Verzoek om extra functionaliteit 10. Verzoek om ook op de site vermeld te worden
Technisch mankement	11. Probleem door technisch mankement
Vormgeving	12. Suggesties over de vormgeving van de site
Overig	13. Overige feedback

Tabel 2: Categorieverdeling voor het feedbackcommentaar op www.utrecht.nl, Vu Anh (2007)

Vu Anh (2007) zegt dat de feedbackbutton nog niet optimaal gebruikt wordt door burgers. Volgens haar is de feedbackbutton op enkele punten te verbeteren. Uit haar resultaten blijkt verder dat slechts ongeveer één derde deel van de reacties op de feedbackbutton daadwerkelijk betrekking heeft op de kwaliteit van de website. De meeste reacties gaan volgens haar over de bereikbaarheid van informatie op de website of de problemen door een technisch mankement.

Infocus

Infocus verschilt op diverse punten van de feedbackbutton. Zo is het geen online-onderzoek, maar een laboratoriumonderzoek. Dit betekent dat de evaluator zich in een geconstrueerde testsituatie bevindt. Daarnaast is het een non-use onderzoek in plaats van een in-use onderzoek. Dit geeft aan dat de evaluator de website puur evalueert in plaats van gebruikt zoals hij dat normaal zou doen.

Het onderzoeksprogramma Infocus is doorontwikkeld vanuit het tekstevaluatieinstrument Focus van De Jong en Lentz (2001). Met Infocus werd de mogelijkheid gecreëerd websites te evalueren. Het evaluatieinstrument is ontwikkeld in het kader van het onderzoeksprogramma 'Evaluatie van gemeentelijke websites'. Infocus werkt als volgt: Proefpersonen worden, al dan niet aan de hand van een taak, gevraagd kritiek te leveren op de pagina's van een website. Op ieder willekeurig moment kunnen zij tijdens het bekijken van de website op de zogenaamde commentaarknop klikken. Er verschijnt dan een venster waarin de proefpersoon zijn commentaar kan melden. De proefpersoon kan vervolgens zelf aangeven in welke categorie zijn commentaar valt. Ook kan deze visueel markeren op welk punt van de website zijn commentaar betrekking heeft. Deze functionaliteit maakt Infocus uniek ten opzichte van andere webevaluatiemethoden. Zodra de proefpersoon zijn commentaar opslaat wordt het verzonden naar de database en beschikbaar gesteld voor verdere analyse (Henssen, 2009).

Onderstaande afbeelding toont het commentaarscherm van Infocus (Henssen, 2009):



Afbeelding 3: Commentaarscherm van Infocus op website www.utrechtourway.nl, Henssen (2009)

De data die worden gebruikt voor dit huidige onderzoek zijn reeds opgenomen en geïnclassificeerd door Henssen (2010). In zijn onderzoek werd het commentaar van 74 proefpersonen op diverse pagina's binnen www.utrecht.nl geanalyseerd. Het commentaar is onderverdeeld in twaalf subcategorieën onder de hoofdcategorieën *Inhoud*, *Bereikbaarheid*, *Vormgeving* van de Web Evaluation Questionnaire (WEQ)

(<http://www2.hum.uu.nl/onderzoek/weg/index.html>) en de restcategorie *overig*. In een eerder onderzoek van Henssen (2009), waarin hij Infocus toepast op de gemeentes Nijmegen, Apeldoorn en Dordrecht, worden enkel de hoofdcategorieën *Inhoud*, *Bereikbaarheid* en *Vormgeving* uit de WEQ gebruikt. Omdat dit onduidelijkheid gaf bij de proefpersonen bood Henssen de aanbeveling de drie hoofdcategorieën te voorzien van subcategorieën. De categorie *Inhoud* kende voor het onderzoek naar www.utrecht.nl vijf subcategorieën, *Bereikbaarheid* kende er vier en *Vormgeving* kende er drie. In schema zijn de volgende categorieën gebruikt:

WEQ	Henssen
Inhoud	<ol style="list-style-type: none"> 1. Tekst is niet relevant 2. Tekst is niet begrijpelijk 3. Tekst is niet volledig 4. Tekst is niet gestructureerd 5. Tekst is niet verzorgd
Bereikbaarheid	<ol style="list-style-type: none"> 6. Onduidelijke link 7. Link/menu niet opvallend 8. Gebrekkige werking menu 9. Gebrekkige navigatiestructuur
Vormgeving	<ol style="list-style-type: none"> 10. Letters moeilijk te lezen 11. Gebrekkige pagina-indeling 12. Gebrekkige vormgeving
/	13. Overig

Tabel 3: Categorieverdeling WEQ en Henssen (2010)

De resultaten van het onderzoek van Henssen (2010) zijn in de conclusie echter weer bijeengebracht tot de drie hoofdcategorieën uit de WEQ om een globaler overzicht te kunnen schetsen. Uit het onderzoek bleek, wat betreft de *Inhoud*, dat het gros van de informatie doorgaans begrijpelijk was. Een enkele keer werden termen en regels onvoldoende toegelicht of waren de teksten onverzorgd. Wat betreft de *Bereikbaarheid* vonden proefpersonen de navigatiestructuur veelal gebrekkig. Ook vielen links niet altijd voldoende op. Maar over het algemeen bleken er vrij weinig ernstige navigatieproblemen te zijn. Ook over de *Vormgeving* bleken de proefpersonen tevreden. Alleen pagina's die afweken van de huisstijl werden niet gewaardeerd. Daarnaast bleek de rechterkolom niet bij iedereen voldoende op te vallen. Onder de categorie *Overig* viel vooral commentaar op technische problemen, zoals interne zoekmachines die niet reageerden.

In het volgende hoofdstuk zal uiteengezet worden hoe in dit huidige onderzoek het commentaar van beide evaluatiemethoden geanalyseerd wordt.

Onderzoeksopzet

Zoals eerder in het Theoretisch kader al gezegd, verschillen de twee evaluatiemethoden op diverse punten. Ook zijn zij beide op andere manieren toegepast wat betreft het classificeren van het commentaar. In deze Onderzoeksopzet zal eerst aan bod komen hoe het nieuwe commentaar van de feedbackbutton verzameld en gecategoriseerd is. Vervolgens zal uit het onderzoek van Henssen (2010) herhaald worden hoe dat commentaar verzameld en gecategoriseerd is. Tot slot worden de verschillen en overeenkomsten tussen beide analyseprocessen toegelicht. Hierbij wordt ook uitgelegd hoe de deelvragen in deze categorisering betrokken zijn.

Feedbackbutton

Het feedbackcommentaar dat in dit huidige onderzoek wordt gebruikt, is afkomstig uit de periode 29-05-2007 tot en met 28-10-2009. Er is geen selectie gemaakt wat betreft opleiding, geslacht en leeftijd van de websitegebruikers. Deze gegevens zijn onbekend. Het commentaar, bestaande uit een kleine 9000 reacties, komt van de centrale bank en betreft het dus commentaar op diverse pagina's. Bij het aanleveren van het commentaar is dus nog geen selectie gemaakt in pagina's.

Om de reacties te analyseren en te categoriseren, is het essentieel te weten hoe deze reacties opgebouwd zijn en uit welke onderdelen zij bestaan (Vu Anh, 2007). Onderstaande reactie toont hoe de reacties aangeleverd kwamen:

Onderwerp: Reactie op pagina: X

Verzonden: dd-mm-jjjj 00:00

Naam: AAA

E-mail: AAA@BBB.nl

Reactie: <test>

Ingevuld vanaf: <http://www.utrecht.nl/smartsite.dws?id=X>

Een reactie bestaat dus uit vijf onderdelen, te benoemen als:

1. Het onderwerp van de reactie
2. De datum en tijd van de verzonden reactie
3. Naam en e-mailadres van de gebruiker
4. De inhoud van de reactie
5. De link van de pagina waar vandaan de gebruiker de reactie invult

Infocus

Bij het onderzoek van Henssen (2010) is gekozen voor een 'vrije' opzet van het Infocusonderzoek. Dit houdt in dat de proefpersonen, in dit geval allemaal hoogopgeleide studenten, zelfstandig aan de slag gingen met drie of vier taken. De proefpersonen kregen vijf tot acht minuten per taak om deze te volbrengen en ondertussen feedback te geven op de bezochte webpagina's. Wanneer een proefpersoon niet binnen de gegeven tijd het antwoord kon vinden wees de onderzoeksleider deze op de juiste pagina. Op deze wijze kon deze proefpersoon alsnog een oordeel geven op de inhoud van de gezochte informatie (Henssen, 2010).

Stappenplan

Ten eerste is een tweedeling gemaakt in het aangeleverde commentaar van de feedbackbutton tussen commentaar dat relevant is voor de kwaliteit van de website en commentaar zonder relevantie voor de website. De relevantie is bepaald vanuit de vraag 'Kan een medewerker van gemeente Utrecht hier iets mee wat betreft de kwaliteit van de website?'. Bij reacties waarin een fout of mogelijke aanpassing gemeld wordt is dit het geval. Bij reacties waarin bijvoorbeeld geklaagd wordt op een referendum is dit niet het geval. Onder het irrelevante commentaar valt daarnaast non-commentaar (reclame en onleesbare teksten), reacties van burgers die geen betrekking hebben op de website (zoals leningaanvragen of klachten op individuele wethouders) en testreacties van medewerkers binnen gemeente Utrecht. Al het relevante commentaar is geclassificeerd op de drie WEQ-hoofdcategorieën: *Inhoud*, *Bereikbaarheid* en *Vormgeving*.

Ten tweede is eenzelfde afbakening gemaakt zoals bij het onderzoek van Henssen (2010) gedaan werd. Dit om een zo valide mogelijke vergelijking te maken. In het onderzoek van Henssen is namelijk een selectie gemaakt van de meest bezochte webpagina's van www.utrecht.nl. De volgende deelsites zijn in zijn onderzoek gebruikt:

- Zwemmen
- Parkeren
- Vacatures
- Afval
- U-pas
- Winkels
- Verhuizen
- Reizigersvaccinaties
- Leidsche Rijn
- WOZ / Gemeentebelastingen
- Vrijmarkt

In dit huidige onderzoek wordt dus ook enkel het commentaar op bovenstaande pagina's verder onderzocht. De reacties op de feedbackbutton op bovenstaande pagina's zijn vervolgens dieper geclassificeerd en onderverdeeld in de eerder genoemde dertien subcategorieën uit, eveneens, het onderzoek van Henssen (2010).

Ten derde is op het hierboven aangegeven afgebakende commentaar aangegeven of het commentaar afkomstig was van een medewerker binnen gemeente Utrecht of een burger. Aan de hand van de verplicht genoteerde e-mailadressen in de reacties kon vastgesteld worden of het commentaar al dan niet afkomstig was van een medewerker: alle e-mailadressen van werknemers binnen gemeente Utrecht eindigen op ...@utrecht.nl. Het commentaar afkomstig van werknemers die zich bewust bezig houden met de kwaliteit van de website werd uitgesloten van het verdere onderzoek. Dit onderzoek richt zich immers op het verschil tussen de in-use methode de feedbackbutton en de non-use methode Infocus. Het commentaar van de werknemers van gemeente Utrecht is niet afkomstig van 'normale' bezoekers, dus valt het niet onder het benodigde materiaal voor de vergelijking tussen de verschillende evaluatiemethoden. De gemeentelijke website van Utrecht bevat tevens een Stadsgids. Begin 2008 heeft de gemeente een brief gestuurd naar alle genoemden in de Stadsgids met de vraag hun gegevens te controleren. Het commentaar op deze gegevens (akkoordmeldingen of in te voeren mutaties) is ook uitgesloten voor het verdere onderzoek omdat dit wederom geen puur in-use commentaar is. Tot slot werd ook dubbel commentaar, wanneer er duidelijk sprake was van een technische fout, verwijderd.

Congruente validering

In de Inleiding is beschreven dat de feedbackbutton en Infocus met elkaar vergeleken worden. Dit kan het beste gedaan worden als ook het commentaar van beide evaluatiemethoden op dezelfde manier gecategoriseerd en geanalyseerd wordt. Per deelvraag zal hier behandeld worden hoe dit tot uitvoering is gebracht.

Op welke dimensies richt het commentaar zich?

Het commentaar van Infocus is op drie (+ overig) hoofdrubrieken geclassificeerd, wat uiteindelijk tot dertien subcategorieën leidt, maar het commentaar van de feedbackbutton is in het eerdere onderzoek door Vu Anh (2007) op vijf (+ overig) geclassificeerd, wat in haar onderzoek ook tot dertien subcategorieën leidt. Echter, omdat het lastig is een vergelijking te maken tussen vijf en drie hoofdcategorieën, zal in dit huidige onderzoek het feedbackcommentaar van www.utrecht.nl op dezelfde wijze als het Infocuscommentaar van Henssen (2010) ingedeeld worden aan de hand van de eerder genoemde drie hoofdcategorieën en de dertien subcategorieën. Op deze wijze kan het nieuwe commentaar van de feedbackbutton en het

bestaande Infocuscommentaar van Henssen (2010) op heldere wijze met elkaar vergeleken worden en daardoor inzicht bieden in de twee verschillende evaluatiemethodes.

In hoeverre is het commentaar relevant?

Om de relevantie van het commentaar specifiek in kaart te kunnen brengen is ook de ernst van het commentaar genoteerd. Hierbij heeft de 'severity rating' van Nielsen als handvat gediend (www.useit.com). Door de mate van ernst van problemen in kaart te brengen kan er gestructureerd een oplossing gevonden worden. Daarnaast geven deze gegevens inzicht in wat bezoekers tijdens het gebruik van websites prettig vinden en waar zij tegenaan lopen. De ernst van een gebruikersprobleem is te bepalen aan de hand van de volgende punten: de frequentie van het probleem, de impact van het probleem en de persistentie van het probleem. Wanneer bepaald is welke mate van ernst het probleem betreft, kan dit ingedeeld worden op een schaal van 0 tot 4, waarbij 0 staat voor 'dit is geen gebruikersprobleem' en 4 staat voor 'catastrofaal probleem, dient direct verholpen te worden' (www.useit.com).

Om een duidelijke vergelijking van de ernst van de problemen te kunnen maken, zal van de pagina's Parkeren, Zwemmen en Afval het commentaar uitgebreid vergeleken worden. De reden waarom het commentaar van juist deze pagina's voor deze deelvraag gebruikt wordt is omdat via beide evaluatiemethoden er relatief veel reacties op deze pagina's geplaatst zijn. Dit zal blijken uit de Resultaten.

Wat zijn de kosten en opbrengsten van beide meetmethodes?

Getracht zal worden na te gaan hoeveel Infocus en de feedbackbutton kosten en wat het eigenlijk opbrengt. Zodra deze gegevens bekend zijn kan afgewogen worden welke meetmethode economisch het meest efficiënt is.

Is er een verschil in kwaliteit van de verbaliseringen van de gebruikers?

Een groot verschil tussen de twee evaluatiemethoden is dat de feedbackbutton een in-use methode in een natuurlijke situatie is en Infocus een non-use methode in een laboratorium. Omdat de gebruikers in beide onderzoeken compleet anders evalueren geeft dat wellicht ook een andere vorm van commentaar. Hoe verbaliseren websitegebruikers hun commentaar als zij gevraagd worden commentaar te geven en als zij uit eigen initiatief een reactie plaatsen? En is daar een patroon in te zien?

Voor deze deelvraag zal wederom enkel het commentaar van de pagina's Parkeren, Zwemmen en Afval geanalyseerd en vergeleken worden, omdat deze data bij beide evaluatiemethoden in grote aantallen aanwezig zijn. Dit zal wederom blijken uit de Resultaten.

Resultaten

Feedbackbutton

Uit de eerste verdeling tussen relevante en irrelevante reacties, bleken 3104 reacties relevant te zijn voor de kwaliteit van de website en 5090 irrelevant. Onderstaand figuur toont waaruit deze algemene verdeling bestaat:



Figuur 1: Verdeling van het commentaar op de feedbackbutton in aantallen en procenten van totaal 8194 reacties

Bovenstaand figuur toont dat de meeste berichten op de feedbackbutton irrelevant zijn (62 %). Een groot deel van de relevante berichten is afkomstig van medewerkers binnen gemeente Utrecht (39 %). Daarnaast richt een deel van de relevante berichten van burgers zich op de Stadsgids (33 % van de 1888 externe relevante reacties).

Wanneer we ons focussen op de eerder genoemde afbakening, dus enkel het commentaar op de deelpagina's grondig analyseren, de categorieën uit het onderzoek van Henssen (2010) gebruiken en de berichten van medewerkers binnen gemeente Utrecht, berichten omtrent de Stadsgids en dubbele berichten uitsluiten, blijven er 390 bruikbare reacties over. Deze brengen de volgende resultaten, wanneer de reacties geïnclassificeerd worden in de dertien subcategorieën:

WEQ	Subcategorieën	Aantal reacties
Inhoud (28 %)	Tekst is niet relevant	13 (3,3 %)
	Tekst is niet begrijpelijk	35 (9 %)
	Tekst is niet volledig	51 (13,1 %)
	Tekst is niet gestructureerd	7 (1,8 %)
	Tekst is niet verzorgd	4 (1 %)
Bereikbaarheid (14 %)	Onduidelijke link	0 (0 %)
	Link/menu niet opvallend	0 (0 %)
	Gebrekkige werking menu	53 (13,6 %)
	Gebrekkige navigatiestructuur	3 (0,8 %)
Vormgeving (1 %)	Letters moeilijk te lezen	0 (0 %)
	Gebrekkige pagina-indeling	1 (0,2 %)
	Gebrekkige vormgeving	3 (0,8 %)
Overig (56,4 %)	/	220 (56,4 %)
Totaal (100 %)		390 (100 %)

Tabel 4: Resultaten van het afgebakend commentaar op de feedbackbutton per subcategorie in aantallen en procenten

Uit bovenstaande tabel valt direct op dat bij lange na niet alle reacties van de bezoekers in te delen zijn in de door Henssen opgestelde subcategorieën. De problemen die het meest genoemd worden uit bovenstaande lijst die wél in te delen zijn, zijn de gebrekkige werking van een menu en een onvolledige tekst. Het overgrote deel van de reacties zijn niet in te schikken in de subcategorieën en vallen daardoor onder *Overig*. Bezoekers blijken in een in-use onderzoek vooral te klagen over informatie die niet gevonden kan worden ('waar vind ik...') of niet die klopt ('het parkeertarief is reeds verhoogd'). Het niet kloppen of de afwezigheid van informatie (los van een onvolledige tekst) van informatie staan echter niet als subcategorieën in de lijst van Henssen (2010). Wanneer informatie verouderd is, zou een subcategorie *Tekst is niet actueel* wel passen, maar de subcategorie *Tekst is niet relevant* past niet.

Een andere veel voorkomende reactie is het geven van tips ('een plattegrond zou handig zijn geweest'). Reacties als deze vallen echter niet onder *Tekst is onvolledig* omdat er niet duidelijk staat dat er relevante tekst ontbreekt. Ook technische problemen worden veelvuldig gemeld. Hier valt bijvoorbeeld onder dat de interne zoekfunctie tot weinig of geen resultaten leidt of dat aanvraagdocumenten of taxatieverslagen niet geopend kunnen worden. Verder wijzen een hoop bezoekers de gemeente op verlopen beveiligingscertificaten. Een ander punt dat niet ingedeeld kan worden in bovenstaande subcategorieën is wanneer bezoekers een foutmelding krijgen bij het versturen van een e-mail. Tot slot klagen een hoop bezoekers wanneer zij een pagina willen printen, maar er een lege A4 uit de printer rolt.

De reacties die hierboven beschreven worden, hebben allemaal betrekking op het eigenbelang van de bezoeker: de bezoeker wil iets weten, maar kan de informatie niet vinden. De bezoeker wil iets aanvragen, maar kan niet bij het document terecht. De bezoeker wil iets uitprinten, maar krijgt geen tekst op papier. De bezoeker heeft een doel om de pagina te bezoeken en zal zich dus blijkbaar ook enkel druk maken over elementen die dit doel dwarsbomen. Hij maakt zich amper druk om bijvoorbeeld de verzorgdheid van de tekst of de vormgeving van de pagina. Daarnaast zijn bovenstaande categorieën gecreëerd voor het non-use laboratorium evaluatiemiddel Infocus, terwijl de feedbackbutton in-use is in een natuurlijke situatie. Dit blijkt een groot verschil teweeg te brengen in het commentaar. Bij Infocus worden de bezoekers indien nodig geholpen door de proefleider. Wanneer zij zelf een link in eerste instantie niet hebben gezien, kunnen zij een reactie geven als 'de link is niet opvallend'. Maar een echte bezoeker van de website zal de link of niet vinden en een klacht geven als 'ik kan de informatie niet vinden' of hij ziet de link en vindt zijn informatie. Het voordeel van Infocus is in dit geval dat het duidelijk is waarom bezoekers informatie niet zouden kunnen vinden, namelijk door de onopvallende link.

Infocus

Via Infocus zijn op de elf eerder genoemde pagina's 1119 reacties geplaatst. Onderstaande tabel toont de verdeling van het afgebakende commentaar op de elf deelpagina's wat ingedeeld is in de dertien subcategorieën.

WEQ	Subcategorieën	Aantal reacties
Inhoud (54,8 %)	Tekst is niet relevant	54 (4,8 %)
	Tekst is niet begrijpelijk	151 (13,5 %)
	Tekst is niet volledig	215 (19,2 %)
	Tekst is niet gestructureerd	99 (8,9 %)
	Tekst is niet verzorgd	94 (8,4 %)
Bereikbaarheid (21,5 %)	Onduidelijke link	61 (5,5 %)
	Link/menu niet opvallend	54 (4,8 %)
	Gebrekkige werking menu	44 (3,9 %)
	Gebrekkige navigatiestructuur	82 (7,3 %)
Vormgeving (13,1 %)	Letters moeilijk te lezen	13 (1,2 %)
	Gebrekkige pagina-indeling	40 (3,6 %)
	Gebrekkige vormgeving	93 (8,3 %)
Overig (10,6 %)	/	119 (10,6 %)
Totaal (100 %)		1119 (100 %)

Tabel 5: Resultaten afgebakend commentaar op Infocus per subcategorie in aantallen en procenten

Vergeleken met de procenten van de algemene categorieverdeling van al het relevante feedbackbuttoncommentaar valt vooral op dat er bij Infocus dertien maal zoveel commentaar wordt gegeven op de vormgeving (13,1 % tegenover 1 %). Als men de opdracht krijgt een website te evalueren lijkt men zich dus meer te richten op de vormgeving dan wanneer zij voor hun eigen doel een website bezoeken.

De subcategorieën met de hoogste scores zijn die van de onbegrijpelijke en onvolledige tekst. Maar ook een onverzorgde, niet-gestructureerde tekst en een gebrekkige vormgeving van de pagina blijken de proefpersonen op te vallen, in tegenstelling tot de gebruikers bij de feedbackbutton. Wat dus direct opvalt, is dat bij Infocus in elke categorie reacties zijn ingedeeld, met een laagste score van 1,3 %. Wederom komt hier naar voren dat als mensen gevraagd worden commentaar te leveren, zij een stuk uitgebreider commentaar geven dan wanneer zij vanuit zichzelf reacties plaatsen. De enige punten die duidelijk bekritiseerd worden bij de feedbackbutton zijn onvolledige teksten en gebrekkige links. Echter, het meeste commentaar is het niet kunnen vinden van informatie, maar deze optie is niet te vinden in de subcategorieën van Infocus. Dat is te begrijpen omdat bij Infocus de participanten (uiteindelijk) naar de betreffende pagina worden geleid door de proefleider. Een webpagina kan gewoonweg bij Infocus dus niet 'niet gevonden worden'. De opdrachten van Infocus zijn daarnaast zo gevormd dat de antwoorden in teksten te vinden zijn. Maar als mensen reageren op de feedbackbutton betreft dat niet altijd informatie via teksten. De subcategorieën onder *Inhoud* gaan dus niet altijd op voor de feedbackbutton omdat deze enkel geformuleerd zijn over teksten.

Dit is ook terug te zien in de grote verschillen in de categorie *Overig* (56,4 % bij de feedbackbutton tegenover 10,6 % bij Infocus). Bij de feedbackbutton zijn een hoop reacties die niet in één van de dertien subcategorieën van Henssen geplaatst kunnen worden. Dit betreffen voornamelijk reacties over technische mankementen, verouderde informatie en informatie die gewoonweg niet aanwezig is op de website. Omdat bij Infocus de opdrachten ontwikkeld zijn, is de kans erg klein dat de gevraagde informatie helemaal niet op de website staat. Proefpersonen van Infocus zullen hier dus ook niet over klagen. Echte bezoekers bezoeken de website voor diverse doelen. Bij deze gebruikers valt het dus direct op als zij de voor hen relevante informatie niet kunnen vinden, deze verouderd is of als zij dankzij een technisch probleem niet bij de informatie kunnen komen.

Congruente validering

Voor de congruente validering is het commentaar afgebakend op een elftal pagina's. De geselecteerde pagina's in dit onderzoek zijn gekozen omdat deze pagina's ook in het onderzoek van Henssen bezocht worden. Volgens statistieken bleken deze pagina's het meest bezocht te worden. In het Infocusonderzoek

zijn alleen die pagina's bezocht en is dus ook alleen op die pagina's commentaar geplaatst. In dit onderzoek hebben de gebruikers geen pagina's aangewezen gekregen.

Onderstaande tabel toont het verschil in aantal reacties per pagina op beide meetmethoden.

Pagina	Aantal reacties feedbackbutton	Aantal reacties Infocus
Zwemmen	68 (17,3 %)	191 (17,1 %)
Parkeren	93 (23,8 %)	140 (12,5 %)
Vacatures	34 (8,7 %)	103 (9,2 %)
Afval	67 (17,2 %)	173 (15,5 %)
U-Pas	15 (3,9 %)	64 (5,7 %)
Winkels	6 (1,6 %)	156 (13,9 %)
Verhuizen	46 (11,8 %)	46 (4,1 %)
Reizigersvaccinaties	1 (0,3 %)	41 (3,7 %)
Leidsche Rijn	7 (1,8 %)	67 (6,9 %)
WOZ / Gemeentebelastingen	47 (12,1 %)	102 (9,1 %)
Vrijmarkt	6 (1,5 %)	36 (3,2 %)
Totaal	390 (100 %)	1119 (100 %)

Tabel 6: Reacties op de feedbackbutton en Infocus op de genoemde pagina's in aantallen en procenten

Direct valt op dat Infocus een stuk meer commentaar oproept in vergelijking met de feedbackbutton (1119 tegenover 390). Dit is opvallend aangezien het verzamelde commentaar op de feedbackbutton afkomstig is uit een periode van zeventien maanden, terwijl de reacties op de feedbackbutton in ongeveer een week tijd verzameld zijn. Dit zou een groot verschil in het aantal commentaar kunnen betekenen, met de feedbackbutton in het voordeel. Echter, dit verschil wordt omgedraaid omdat de reacties op de feedbackbutton zelfstandig gepost zijn. Bij Infocus hebben de evaluatoren de opdracht gekregen commentaar te leveren. Wanneer men de opdracht krijgt opmerkingen en/of kritiek te leveren, blijkt men dus veel meer geneigd commentaar te plaatsen, dan wanneer dit een vrijwillige keus is. Later in dit hoofdstuk zal verder behandeld worden of dit verschil ook invloed heeft op de relevantie van het commentaar.

Opvallend is ook dat niet alle gekozen pagina's uit het onderzoek van Henssen ook meteen veel commentaar van de echte websitebezoekers krijgen. Veel bezoekers betekent dus niet direct veel reacties. De meeste reacties uit de data van dit huidige onderzoek waren geplaatst op de volgende pagina's: Stadsgids, Evenementen, Routeplanner, Hondenweides, Stadsplannen, Openbare melding en de Bibliotheek. Alleen de pagina's Zwemmen, Parkeren en Afval uit de lijst meest bezochte pagina's uit het

onderzoek van Henssen (2010) kregen ook in dit onderzoek veel reacties. Voor enkele deelvragen zal dus ook het commentaar op deze drie pagina's gebruikt worden als vergelijkbaar materiaal.

Om een heldere vergelijking te krijgen tussen de feedbackbutton en Infocus zullen hieronder de deelvragen verder beantwoord worden.

Op welke dimensies richt het commentaar zich?

Uit bovenstaande tekst en tabellen, over het commentaar op de afgebakende pagina's, komt al naar voren dat aangewezen evaluatoren op een veel breder vlak een website bekritisieren. Alles wat hen opvalt benoemen zij. Echte gebruikers reageren meestal alleen als hen iets in de weg staat. Zij gaan voornamelijk naar de website om iets op te zoeken en als die informatie niet gevonden kan worden of niet duidelijk is zullen zij dit melden. De overige informatie op de pagina gaat hen weinig aan. Uit bovenstaande resultaten blijken de enige dimensies uit de categorieënlijst van Henssen (2010) waarop echte gebruikers zich richten dan ook te maken te hebben met de volledigheid van informatie en de gebrekkige werking van links. De andere dimensies worden vrijwel niet benoemd.

De vraag hierbij is wel of alle subcategorieën uit het onderzoek van Henssen er wel toe doen. Het zijn punten die volgens de proefpersonen uit zijn onderzoek soms wat verbeterd mogen worden, maar échte gebruikers lijken er vrijwel geen waarde aan te hechten. En als de echte gebruikers zich ergens toch niet aan storen, dan is de vraag of daar dan wel tijd en geld ingestoken zou moeten worden. En zijn mensen die aangesteld worden als evaluator niet té streng, zodat zij geforceerd puntjes gaan zoeken en benoemen die wel wat verbeterd zouden mogen worden, zonder dat zij zich er echt aan storen?

In hoeverre is het commentaar relevant?

Voor deze deelvraag is enkel het commentaar op de pagina's Parkeren, Zwemmen en Afval gebruikt, zodat een duidelijke vergelijking gemaakt kan worden. Deze pagina's zullen respectievelijk behandeld worden.

Parkeren

Onderstaande tabel toont welke unieke problemen genoemd zijn bij beide meetmethoden op de pagina's over Parkeren. Problemen genoemd door één persoon en complimenten zijn hierbij niet betrokken, vandaar dat de reacties bij elkaar opgeteld lager uitvallen dan de totaalscore. Bij Kans staat vermeld hoeveel reacties over het betreffende probleem geplaatst zijn. Bij overeenkomende problemen wordt ook het aantal in procenten van het totaal aantal reacties vermeld. Bij Ernst staat vermeld welke mate van ernst

het betreffende probleem heeft. Vervolgens worden eerst de reacties van Infocus toegelicht en daarna zullen de reacties van de feedbackbutton aan bod komen.

Parkeren					
Infocus	Kans	Ernst	Feedbackbutton	Kans	Ernst
Parkeergebieden onduidelijk	7 (5 %)	4	Parkeergebieden onduidelijk	44 (47 %)	4
Plattegrond opent niet	3	2	Informatie klopt niet	9	4
Startpagina geen startpagina	2	1	Kaartje parkeergarages ontbreekt	9	3
Links niet toegelicht	2	2	Vergunningaanvragen onduidelijk	6	2
Link valt niet op	2	2	Technische problemen	5	3
'Rayon' niet bekend	2	2			
'Spelregels' onpassend	2	1			
Redactionele foutjes	12	1			
Totaal Gemiddeld	140	1,6	Totaal Gemiddeld	93	3,2

Tabel 7: Aantal vermeldingen en ernst per unieke problemen bij Infocus en de feedbackbutton op pagina Parkeren

Opvallend is dat maar één probleem bij beide meetmethoden genoemd wordt, namelijk de onduidelijkheid over de parkeergebieden. De redactionele foutjes op de pagina worden bij Infocus meer opgemerkt. Daarnaast is direct te zien dat de feedbackbutton op de pagina Parkeren minder diverse reacties oproept, maar wel veel meer reacties krijgt over dezelfde onderwerpen. Hieronder zal dit verschil toegelicht worden.

Bij Infocus moesten de proefpersonen via de pagina Parkeren drie taken op zien te lossen. Zij moesten zich oriënteren op parkeermogelijkheden in Utrecht, het aanvragen van een parkeervergunning en bezoekersvergunningen. Onderstaand worden de meest voorkomende reacties uit de 140 reacties van de proefpersonen op de pagina's Parkeren genoemd en beoordeeld middels de severity rating van Nielsen (www.useit.com).

Bij het zoeken naar parkeerplaatsen maakten enkele proefpersonen de melding dat de link naar de plattegrond niet in een nieuw venster opent. De plattegrond opent dus wel, maar bezoekers lijken het prettiger te vinden als dit in een nieuw venster gebeurt. Dit probleem heeft om die reden een lichte mate van ernst, dus krijgt een 2 als schaalbeoordeling (klein gebruikersprobleem, verhelpen heeft lage prioriteit).

Een ander opvallend punt op de website Parkeren is volgens enkele proefpersonen dat de startpagina van Parkeren geen startpagina is. Dit betreft niet zozeer het gebruik, maar duidelijkheid binnen de structuur en vormgeving. Daarom krijgt dit commentaar de schaalbeoordeling 1 (cosmetisch probleem,

alleen verhelpen als daar tijd voor is). Een probleem met dezelfde mate van ernst is dat het kaartje met parkeergelegenheden voor de auto ook fietsenstallingen toont.

Op de pagina met links naar parkeergarages vonden enkele proefpersonen het storend dat er nergens vermeld staat dat het externe links naar websites van derden zijn. Dit betreft het gebruik en kan als storend ervaren worden, maar belemmert het gebruik verder niet. Daarom heeft dit probleem de schaalwaardering 2.

Bij de zoektocht naar de tarieven van een bezoekersvergunning stuitten twee proefpersonen op een probleem: De link naar de tarieven staat op een erg onopvallende plek. Dit betreft een gebruikersprobleem, maar aangezien maar weinig proefpersonen de link niet opvallend genoeg vonden, krijgt ook dit probleem de schaalwaardering 2.

Een punt dat relatief veel proefpersonen (zeven) meldden betrof de onduidelijkheid over de parkeergebieden bij de parkeervergunningaanvragen. Er wordt wel gesproken van gebied 1, 2 en 3, maar nergens wordt toegelicht welke stadsdelen/straten dit betreft. Dit is een groot onduidelijk punt in het gebruik en aangezien relatief veel proefpersonen dit opvalt krijgt dit probleem de schaalwaardering 4 (catastrofaal voor gebruikers, dient direct te worden verholpen). Een ander punt dat onduidelijkheid gaf binnen de parkeervergunningaanvragen was dat het woord 'rayon' voor enkele proefpersonen niet bekend was. Het is geen enorm probleem, maar kan wel grote onduidelijkheid geven. Daarom krijgt dit probleem schaalwaardering 2.

De regels omtrent de parkeervergunning werden aangeduid met het woord 'spelregels'. Enkele proefpersonen konden dit niet waarderen en zullen een formelere term prefereren. Het woord 'spelregels' belemmert echter niet het gebruik, dus krijgt dit probleem schaalwaardering 1.

Enkele proefpersonen vielen wat kleine redactionele foutjes op. Dit belemmert het gebruik niet en heeft daardoor een erg kleine mate van ernst. Ook dit probleem kan dus geclassificeerd worden met de schaalwaardering 1.

Uit bovenstaande opsomming blijkt dat het enige probleem dat een hoge prioriteit heeft de onduidelijkheid over de parkeergebieden is. En dat punt wordt bovendien slechts in zeven van de 140 reacties benoemd. De gemiddelde ernst van de problemen bij de proefpersonen van Infocus is 1,6 op de schaal van Nielsen. Henssen (2010) heeft in zijn onderzoek alle commentaarpunten beschreven, ook al werden de meeste slechts door één proefpersoon gemeld. Hierbij is het de vraag of een punt dat slechts één persoon opvalt, eigenlijk wel relevant is. Een website moet voor 'de gemiddelde bezoeker' helder en overzichtelijk zijn. Wanneer één persoon een pagina niet overzichtelijk vindt, zegt dat dan meer over de pagina of meer over de proefpersoon?

Hieronder zullen de meest voorkomende reacties op de pagina Parkeren van de feedbackbutton beschreven worden. Ook hier zullen ze ingedeeld worden op mate van ernst middels de severity rating van Nielsen.

Binnen de 390 relevante berichten betroffen 93 reacties de pagina Parkeren. Bijna de helft van de reacties (44) betroffen commentaar op de onduidelijkheid van de parkeergebieden. Net als bij Infocus is dit dus het meest grote probleem bij de websitegebruikers. Desalniettemin komt dit bij de feedbackbutton een stuk duidelijker naar voren. De bezoekers van de website is in dit geval namelijk niet gevraagd commentaar te geven. Zij reageren enkel op een pagina als zij vinden dat er iets gemeld moet worden. Dat nu bijna de helft van de commentaren op de pagina Parkeren de onduidelijkheid van de parkeergebieden betreft, geeft aan dat dit toch echt wel een ernstig probleem is en spoedig aangepast zou moeten worden. Daarom zou dit probleem hier de schaalwaardering 4 krijgen.

Andere problemen die via de feedbackbutton regelmatig genoemd worden (beide 9 keer) voor de pagina Parkeren zijn niet kloppende informatie (parkeertarieven die te laag staan) en het gebrek aan een kaartje waar alle parkeergarages op staan aangegeven. Beide problemen betreffen belangrijke gebruikersproblemen die grote consequenties met zich mee dragen. Wanneer men bijvoorbeeld rekent op een uurtarief van 2,- euro, maar dit blijkt achteraf 4,- euro te zijn, dan levert dit een hoop agressie en problemen. Daarom wordt het eerstgenoemde probleem beoordeeld met schaalwaardering 4. Het gebrek aan een kaartje met parkeergarages is geen catastrofaal probleem, maar dient zeker verholpen te worden. Een gemeentelijke website hoort voor bewoners en bezoekers uitnodigend en duidelijk te zijn. Daarom krijgt deze klacht de schaalwaardering 3 (groot gebruikersprobleem, verhelpen heeft hoge prioriteit).

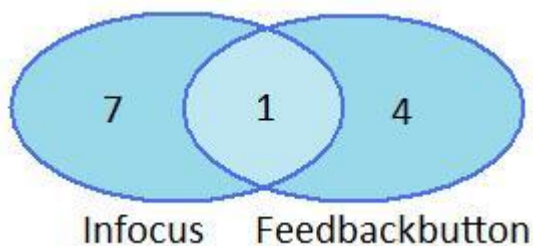
Enkele bezoekers klaagden over de onduidelijkheid omtrent het aanvragen/omzetten van de parkeervergunning. Ook al is de groep maar klein die de informatie onduidelijk vindt, desalniettemin moet dit in het achterhoofd meegenomen worden. Iedere websitebezoeker zou immers makkelijk informatie moeten vinden en begrijpen op een gemeentelijke website. Daarom krijgt dit probleem toch de schaalwaardering 2.

Tot slot werden enkele malen technische problemen aan het licht gebracht. Zo werkten diverse links niet en kon de e-mail niet verstuurd worden. Ook al worden deze problemen sporadisch genoemd, het zijn wel punten die direct opgelost zou moeten worden, zodat de werking van de website vlekkeloos verloopt. Om die reden worden deze punten afgeboekt met schaalwaardering 3.

Uit bovenstaande gegevens blijkt wederom dat echte bezoekers een stuk minder diverse problemen benoemen en zich voornamelijk focussen op de echt storende punten. De gemiddelde ernst van deze punten is 3,2 op de schaal van Nielsen. De ernst van deze problemen is dus twee keer zo groot als bij Infocus. De proefpersonen van Infocus hebben een kritischer oog en benoemen alles wat hen opvalt, terwijl de echte bezoekers enkel de voor hen relevante informatie willen vinden. Lukt dit niet, dan hebben zij iets

om over te klagen. Het nadeel van de feedbackbutton is dat niet alle te verbeteren punten op de website genoemd worden, maar het voordeel is juist dat wel alle écht belangrijke punten aan de orde komen. Daar waar de feedbackbutton gehoor geeft aan de echte gebruikers voor echte problemen, geeft Infocus juist weer een breder inzicht over kleinere oneffenheden.

Het is duidelijk dat Infocus meer commentaar teweeg brengt, maar dat deze over veel meer diverse en kleinere problemen gaan. Er zijn een stuk minder feedbackreacties op de pagina Parkeren dan bij Infocus, maar deze gaan vrijwel allemaal over dezelfde onderwerpen, waardoor de ernst ervan ook direct een stuk groter is. Onderstaande afbeelding geeft dit helder weer:



Figuur 2: Weergave van de unieke en overeenkomende problemen bij Infocus en de feedbackbutton op de pagina's Parkeren

De afbeelding geeft weer dat Infocus acht unieke problemen (met uitzondering van de problemen die door één enkele proefpersoon zijn en complimenten) teweeg heeft gebracht op de pagina's over Parkeren en de feedbackbutton vijf. Een van de problemen wordt bij beide meetmethoden genoemd. Dit is tevens het probleem met de hoogste ernst, namelijk de schaalwaardering 4. Het betreft in dit geval de onduidelijkheid over de parkeergebieden.

Opvallend is dat bij Infocus slechts 5 % van de reacties van de proefpersonen dit probleem benoemen en van de reacties van de echte gebruikers 47 % dit probleem aanstipt. Dit heeft echter niets met de mate van ernst te maken, maar met de mate van kans. De kans is dus groter dat echte gebruikers de onduidelijkheid over de parkeergebieden benoemen, dan dat aangewezen proefpersonen dat doen.

Zwemmen

Onderstaande tabel toont welke unieke problemen genoemd zijn bij beide meetmethoden op de pagina's over Zwemmen. Problemen genoemd door één persoon en complimenten zijn hierbij niet betrokken, vandaar dat de reacties bij elkaar opgeteld lager uitvallen dan de totaalscore. Bij Kans staat vermeld hoeveel reacties over het betreffende probleem geplaatst zijn. Bij overeenkomende problemen wordt ook het aantal in procenten van het totaal aantal reacties vermeld. Bij Ernst staat vermeld welke mate van ernst

het betreffende probleem heeft. Vervolgens worden eerst de reacties van Infocus toegelicht en daarna zullen de reacties van de feedbackbutton aan bod komen.

Zwemmen					
Infocus	Kans	Ernst	Feedbackbutton	Kans	Ernst
Tarieven en openingstijden	3 (1,6 %)	4	Tarieven en openingstijden	17 (25 %)	4
Activiteitenschema mist	4 (2,1 %)	3	Activiteitenschema mist	15 (22,1 %)	3
Extra link openingstijden	2 (0,6 %)	2	Extra link openingstijden	11 (16,2 %)	2
Voorzieningen te beperkt	3	3	Doorstromen onduidelijk	4	3
Te veel links	3	3	Informatie verouderd	4	2
Te weinig foto's	3	2	Activiteitenrooster opent niet	4	2
Link om terug te gaan mist	2	3	Spelfouten	4	1
Link banenzwemmen mist	2	3			
'Juiste baan' onduidelijk	2	2			
Rode ballonnen onduidelijk	2	2			
Onopvallende links	2	2			
Roosters onoverzichtelijk	2	3			
Redactionele foutjes	23	1			
Totaal Gemiddeld	191	2,5	Totaal Gemiddeld	68	2,6

Tabel 8: Aantal vermeldingen en ernst per unieke problemen bij Infocus en de feedbackbutton op pagina Zwemmen

Ook hier is direct zichtbaar dat bij de drie bovenste overeenkomende problemen, de feedbackbutton een stuk meer reacties heeft verzameld. Dit is onder andere te zien bij de extra link voor de openingstijden. Voordat de openingstijden ingezien kunnen worden moet er een pdf-bestand geopend worden. De echte bezoekers blijken dit een stuk vervelender te vinden dan de proefpersonen bij Infocus. De redactionele foutjes worden wederom wel een stuk meer opgemerkt bij Infocus.

Bij Infocus moesten de proefpersonen via de pagina Zwemmen ook drie taken op zien te lossen. Zij moesten zich oriënteren op zwemverenigingen, de openingstijden en –dagen van de zwembaden in Utrecht en de bereikbaarheid van de zwembaden. Onderstaand worden de meest voorkomende reacties uit de 191 reacties van de proefpersonen op de pagina's Zwemmen genoemd en beoordeeld middels de severity rating van Nielsen (www.useit.com).

Enkele proefpersonen stelden dat het niet direct duidelijk was welke rode ballonnen op Google Maps bij welk zwembad hoorden. Dat zij dit niet direct konden vinden, geeft aan dat dit op een gegeven moment wel duidelijk is geworden. Om die reden heeft dit probleem de schaalwaardering 2.

Een aantal proefpersonen maakte opmerkingen van redactionele aard. Dit zijn geen grote onduidelijkheden. Daarom krijgen deze problemen de schaalwaardering 1.

Twee proefpersonen merkten op dat er geen link is waarmee terug kan worden gekeerd naar de pagina waar alle zwembaden te zien zijn. Ook al kan de backbutton van de browser gebruikt worden; een link om terug te keren is een stuk overzichtelijker. Omdat een website zo gebruiksvriendelijk mogelijk moet zijn, heeft dit probleem de schaalwaardering 3.

Een andere link die door twee proefpersonen gemist wordt, is een link over banenzwemmen. Een link naar een pagina specifiek over banenzwemmen is dus niet aanwezig. Omdat ook dit een stuk directer, sneller en overzichtelijker zou kunnen, krijgt ook dit probleem schaalwaardering 3.

De zin 'kies de juiste baan' riep bij twee proefpersonen vragen op. Zij begrepen niet wat 'de juiste baan' dan is en hadden graag wat toelichting gezien. Dit is een klein inhoudelijk probleem omdat het geen essentiële basisinfo betreft. De websitebezoekers zullen geen grote problemen ondervinden bij het zoeken naar informatie dankzij deze zin. Daarom krijgt dit probleem schaalwaardering 2.

Informatie die wel tot de basisinformatie behoort en onduidelijk was in de ogen van vier proefpersonen was de ontbrekende tekst over creatief zwemmen. Niemand leek te vinden wanneer iedereen creatief mag zwemmen. Omdat dit belangrijke informatie betreft krijgt dit probleem schaalwaardering 3.

Drie proefpersonen waren van mening dat er te veel menu's en links op de pagina stonden. Dit ten nadele van het globale overzicht. Omdat een overzichtelijke toegankelijke website een groot punt is, krijgt dit probleem schaalwaardering 3.

Aansluitend zagen twee proefpersonen enkele links aanvankelijk over het hoofd. Maar omdat zij deze links alsnog eigenhandig wel zagen zitten is dit geen zeer ernstig probleem en krijgt het schaalwaardering 2.

Twee proefpersonen vonden de roosters onoverzichtelijk, omdat er voor elke vakantie een nieuw rooster was aangemaakt. Omdat de informatie wel aanwezig is, maar wellicht niet heel mooi gesorteerd staat, krijgt dit probleem schaalwaardering 2. De informatie is er immers wel, maar misschien niet voor elke bezoeker direct duidelijk te overzien.

Twee proefpersonen vonden het overbodig dat er een extra link is voor de openingstijden. Deze zouden volgens hen al direct getoond kunnen worden. Omdat ook dit geen groot probleem, maar wel een verbetering is, krijgt ook dit probleem schaalwaardering 2.

Drie proefpersonen vonden de informatie over de voorzieningen te beperkt. Nergens staat duidelijk of er eetmogelijkheden, glijbanen en/of sauna's aanwezig zijn. Omdat dit basisinformatie van een zwembad is, krijgt dit probleem schaalwaardering 3.

Ook de foto's gaven geen volledig beeld van het zwembad. Deze hadden wat uitgebreider mogen zijn. Maar omdat foto's in dit geval minder essentieel zijn dan de tekstuele informatie krijgt dit probleem schaalwaardering 2.

Drie proefpersonen konden de openingstijden en tarieven helemaal niet vinden. Omdat dit de belangrijkste informatie op de wegpagina is, krijgt dit probleem schaalwaardering 4. Wanneer bezoekers op de website niet kunnen vinden voor welk bedrag en wanneer ze kunnen zwemmen, is er immers sprake van een groot probleem.

Uit bovenstaande opsomming blijkt het probleem met de hoogste prioriteit de onvindbaarheid van de tarieven en openingstijden. Dat punt wordt slechts in drie van de 191 reacties benoemd. De gemiddelde ernst van de problemen bij de proefpersonen van Infocus is 2,5 op de schaal van Nielsen.

Hieronder zullen de meest voorkomende reacties op de pagina Zwemmen van de feedbackbutton beschreven worden. Ook hier zullen ze ingedeeld worden op mate van ernst middels de severity rating van Nielsen.

Binnen de 390 relevante berichten zijn er 68 geplaatst op de pagina's over Zwemmen. Zeventien van deze reacties gaan over de onduidelijkheid van de openingstijden. Omdat dit wellicht de meest belangrijke informatie is op een pagina over een zwembad, krijgt dit probleem schaalwaardering 4. Als bezoekers niet kunnen vinden wanneer zij er terecht kunnen, is er sprake van een groot probleem. Dit zelfde probleem werd ook bij Infocus genoemd.

Aansluitend hieraan wordt er elf keer geklaagd over het niet kunnen openen van de openingstijden. De openingstijden staan niet op de pagina zelf, maar in een pdf-bestand dat eerst gedownload moet worden. Veel bezoekers blijken dit vervelend te vinden en melden dat via de feedbackbutton. Echter, dit kan niet als een catastrofaal probleem bestempeld worden omdat de openingstijden immers wel aanwezig zijn, hetzij achter een link. Daarom krijgt dit probleem schaalwaardering 2. Ook dit probleem viel de proefpersonen van Infocus op.

Een ander groot probleem bleek de onduidelijkheid van de activiteitenkalender. Vijftien bezoekers bleken niet helder te kunnen zien wanneer er bijvoorbeeld baantjes getrokken konden worden of wanneer het zwemmen voor vrouwen plaatsvond. Dit probleem werd ook gemeld bij Infocus, maar in mindere mate. De proefpersonen bij Infocus misten een directe link naar de informatie over banenzwemmen en recreatief zwemmen, terwijl de echte websitegebruikers gewoonweg niet kunnen vinden wanneer welke activiteiten plaatsvinden. Enkele bezoekers geven dit aan met beschrijvingen als: "Zojuist van een teleurstelling bekomen: baantjeszwemmen in De Kwakel was vandaag maar tot 10:00 uur i.p.v. 11:00 uur mogelijk. Dit was bijzonder onduidelijk tot niet weergegeven op de website". Omdat het zeer belangrijk is voor

bezoekers te weten wanneer zij welk soort zwemmen kunnen beoefenen, krijgt dit probleem schaalwaardering 3.

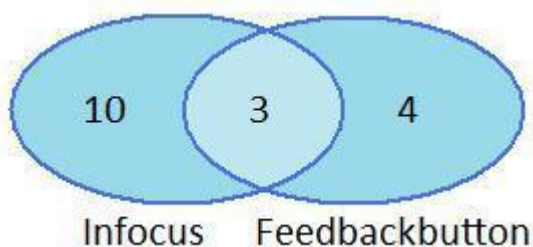
De overige punten werden door een stuk minder personen aangestipt. Vier proefpersonen konden de activiteitenkalender niet openen. Ook dit krijgt schaalwaardering 2, omdat de informatie wel op de pagina aanwezig is, maar door een technisch mankement niet gelezen kan worden. Vier bezoekers viel het op dat enige informatie verouderd was. Dit is geen groot probleem, aangezien het geen zeer belangrijke informatie betrof, zoals openingstijden en tarieven. Maar omdat het wel onzorgvuldig is, krijgt dit probleem schaalwaardering 2.

De informatie die verder bij vier proefpersonen mist was informatie over het doorstromen bij zwemdiploma's. Het is op de website niet duidelijk wanneer de zwemlessen plaatsvinden en op welk niveau gezwommen kan worden. Dit is belangrijke informatie voor als ouders hun kind willen inschrijven voor zwemlessen. Daarom krijgt dit probleem schaalwaardering 3.

Tot slot vonden drie proefpersonen enkele spelfouten. Dit is een cosmetisch probleem, zoals Nielsen dat noemt, en krijgt daarom schaalwaardering 1.

Uit bovenstaande gegevens blijkt dat echte bezoekers de meeste problemen hebben met onduidelijke openingstijden en activiteitentijden. De gemiddelde ernst van alle genoemde punten is 2,6 op de schaal van Nielsen. Dit verschilt, in tegenstelling tot de pagina Parkeren, miniem met de gemiddelde ernst van de punten genoemd bij Infocus.

Wat opvalt is dat de proefpersonen bij Infocus veel meer problemen benoemen en dat maar een klein aantal de grote problemen aanstipt, terwijl de bezoekers die reageren op de feedbackbutton meer overeenkomende grote problemen aankaarten. Echter, omdat hier de problemen die aangestipt zijn door één enkele proefpersoon of bezoeker niet meegenomen zijn, verschillen de gemiddelde waardes van ernst niet veel van elkaar. Als bij Infocus alle kleine problemen met een lage schaalwaardering genoemd worden en meegenomen worden in de berekening naar de gemiddelde ernst, dan zou de gemiddelde ernst een stuk lager liggen. Onderstaande afbeelding geeft een helder inzicht in de diversiteit van de problemen:



Figuur 3: Weergave van de unieke en overeenkomende problemen bij Infocus en de feedbackbutton op de pagina's Zwemmen

De afbeelding geeft weer dat Infocus wederom meer diverse reacties heeft opgeroepen dan de feedbackbutton. Bij het onderdeel Zwemmen komen drie problemen van Infocus en de feedbackbutton

overeen. Het probleem dat de openingstijden en tarieven niet gevonden konden worden heeft een mate van ernst van 4. De andere twee problemen (activiteitschema ontbreekt en de extra link voor de openingstijden is vervelend) krijgen schaalwaardering 3. Hier komt dus wederom naar voren dat de overeenkomende problemen de meest ernstige problemen zijn. Ook komt naar voren dat de meest ernstige problemen op de pagina's Zwemmen bij de feedbackbutton veel vaker aangestipt worden dan bij Infocus (22,1 % tegenover 2,1 %, 25 % tegenover 1,6 % en 16,2 % tegen 0,6 %). Het begint erop te lijken dat de meest ernstige problemen op een website het beste benoemd worden door echte gebruikers. De mate van ernst is hoog bij veel van de genoemde problemen, maar ook de kans is erg groot, aangezien het percentage bijbehorende reacties beduidend hoger ligt dan bij Infocus. De volgende analyse van de pagina's over Afval zal hier een definitief inzicht in kunnen geven.

Afval

Onderstaande tabel toont welke unieke problemen genoemd zijn bij beide meetmethoden op de pagina's over Afval. Problemen genoemd door één persoon en complimenten zijn hierbij niet betrokken, vandaar dat de reacties bij elkaar opgeteld lager uitvallen dan de totaalscore. Bij Kans staat vermeld hoeveel reacties over het betreffende probleem geplaatst zijn. Bij overeenkomende problemen wordt ook het aantal in procenten van het totaal aantal reacties vermeld. Bij Ernst staat vermeld welke mate van ernst het betreffende probleem heeft. Vervolgens worden eerst de reacties van Infocus toegelicht en daarna zullen de reacties van de feedbackbutton aan bod komen.

Afval					
Infocus	Kans	Ernst	Feedbackbutton	Kans	Ernst
Ophaaltijden niet compleet	10 (5,8 %)	4	Ophaaldagen onbekend	13 (19,4 %)	4
Zoekapplicatie valt niet op	6	2	Ophaaltijden feestdagen	15	4
Link afvalwijzer valt niet op	5	2	Plattegrond afvalpunten	10	2
Kaartje afvalscheidingstations	5	2	Verouderde informatie	8	2
Afvallijst onduidelijk	4	2	Aanvraagformulier werkt niet	5	2
Link grofvuil onopvallend	4	3	Uitleg over afval	4	3
Ophaalicoon onduidelijk	4	2	Gebrek aan klachtenformulier	3	2
Zoekapplicatie is omslachtig	3	2	Onduidelijkheid frituurvet	3	2
Plattegrond onduidelijk	3	2			
Dikgedrukt kopje	3	1			
Telefonische aanvraag	2	3			
Plattegrond ver uitgezocht	2	2			
Op afspraak bezoeken	2	3			
Totaal Gemiddeld	173	2,1	Totaal Gemiddeld	67	2,6

Tabel 9: Aantal vermeldingen en ernst per unieke problemen bij Infocus en de feedbackbutton op pagina Afval

Ook op de pagina's over Afval komt naar voren dat de echte bezoekers meer over dezelfde problemen klagen. Het enige punt waar beide over klagen zijn onduidelijke reguliere ophaaltijden en de onduidelijke plattegrondjes. Echter, in beide gevallen wordt er op verschillende wijze over de plattegrondjes geklaagd: de proefpersonen van Infocus vinden het kaartje te ver uitgezocht en te onduidelijk, terwijl de echte bezoekers op het kaartje centrale afvalpunten aangegeven willen hebben. Dit geeft weer duidelijk aan waarin beide meetmethoden verschillen. De feedbackbutton geeft vooral gehoor aan echte gebruikerswensen en –problemen, terwijl de proefpersonen van Infocus zich ook erg richten op vormgevingsproblemen. Deze verschillen worden hieronder toegelicht.

Bij Infocus moesten de proefpersonen via de pagina Afval drie taken op zien te lossen. Zij moesten zich oriënteren op het ophalen van huisvuil, het ophalen van grofvuil en afvalscheidingstations. Onderstaand worden de meest voorkomende reacties uit de 173 reacties van de proefpersonen op de pagina's Afval genoemd en beoordeeld middels de severity rating van Nielsen (www.useit.com).

Zes proefpersonen snaptten de werking van de applicatie waarmee de ophaaldagen van het huisvuil getoond werd niet. De informatie verschijnt links en dit hadden zij niet door. Dit is een probleem omtrent

onopvallendheid. De informatie is immers wel gewoon aanwezig. Daarom krijgt dit probleem de schaalwaardering 2.

Aansluitend meldden drie proefpersonen dat de zoekfunctie een omslachtige manier van zoeken is. Deze applicatie vraagt de bezoeker een adres in te vullen en geeft vervolgens de ophaaldagen en –tijden. De klacht over deze zoekfunctie is een idee voor een optimalisatie, maar betreft geen groot probleem. Ook daarom krijgt dit probleem schaalwaardering 2.

Maar liefst tien proefpersonen maakten melding van het probleem dat de ophaaltijden van het huisvuil niet voor alle adressen bekend zijn. Dit is een groot probleem. Bezoekers komen juist naar deze webpagina om te controleren wanneer hun huisvuil opgehaald wordt. Zij kunnen daar nu dus niet allemaal de benodigde informatie over vinden. Daarom krijgt dit probleem schaalwaardering 4.

Twee proefpersonen meldden dat de plattegrondjes waarop glasbakken, papierbakken en textielbakken staan aangegeven, te ver uitgezoomd zijn, waardoor een onduidelijk beeld ontstaat. Daarnaast constateerden drie proefpersonen dat de icoontjes te ver op elkaar staan, waardoor deze elkaar overlappen. Dit zijn kleine problemen wat betreft onoverzichtelijkheid. Omdat de informatie wel aanwezig is en ingezoomd kan worden op de plattegrondjes, krijgen deze punten beide schaalwaardering 2.

Vier proefpersonen herkenden het ophaal-icoon, waarmee zij de afvalpunten op de kaart konden tonen, niet. Het icoon bleek te klein en onbekend. Omdat dit een uiterlijk probleem is met kleine gebruikersgevolgen krijgt ook dit probleem schaalwaardering 2.

Vijf proefpersonen vonden de link ‘afvalwijzer’ in de lopende tekst te onopvallend. De link is belangrijk en mag volgens hen meer opvallen. Omdat de proefpersonen de link wel eigenhandig gevonden hebben en het punt geen grote gebruikersproblemen teweeg brengt, krijgt dit probleem schaalwaardering 2.

Vier proefpersonen constateerden dat de link voor de afvalwijzer 2010 tweemaal genoemd wordt, maar beide keren anders geformuleerd is. Dit scheidt verwarring. Wederom is dit een klein probleem en heeft een mate van ernst van 2 op de schaal van Nielsen.

De lijst waarop staat wat wel en wat niet als huisafval telt, vonden vier proefpersonen inconsistent en onduidelijk. De kleuren en lettertypes variëerden te veel waardoor onduidelijkheid ontstaat. Ook dit is een klein probleem wat zich richt op onoverzichtelijkheid. Dit probleem krijgt schaalwaardering 2.

Vier proefpersonen vonden het vreemd dat de link ‘grofvuil laten ophalen’ onderaan de pagina staat. Dit is volgens hen dé informatie waarvoor bezoekers naar de webpagina komen en mag dus wat prominenter geplaatst worden. Ook dit heeft met overzichtelijkheid te maken en heeft daarom schaalwaardering 2.

Aansluitend is het op de webpagina niet bekend waar en wanneer het grofvuil opgehaald kan worden. Hiervoor moet gebeld worden. Twee proefpersonen vinden dit onbegrijpelijk. Omdat deze informatie duidelijk aanwezig had moeten zijn en dit probleem gebruikersirritaties teweeg brengt, krijgt dit probleem schaalwaardering 3.

Drie proefpersonen snappen niet waarom 'legitimatie bewijs verplicht' (bij het bezoeken van een afvalscheidingstation) dikgedrukt is, terwijl dit geen kopje is. Omdat dit enkel een onduidelijkheid van lettertype is, krijgt dit probleem schaalwaardering 1. Verder werden er nog enkele redactionele suggesties gedaan. Ook deze problemen vallen onder schaalwaardering 1.

Bij de informatie over de afvalscheidingstations werd niet expliciet gemeld of deze alleen op afspraak bezocht konden worden. Hier stoorden twee proefpersonen zich aan. Het was een stuk duidelijker geweest voor de geïnteresseerde bezoeker als het er wel had gestaan. Nu is de belangrijke informatie incompleet. Daarom krijgt dit probleem schaalwaardering 3.

Ook het gebrek aan een kaartje met daarop alle afvalscheidingstations viel bij vijf proefpersonen op. Dit zou een nuttige toevoeging zijn geweest. Het verhoogt de gebruiksvriendelijkheid en krijgt daarom schaalwaardering 2.

Uit bovenstaande opsomming valt direct op dat, in tegenstelling tot de reacties op de pagina's Parkeren en Zwemmen van Infocus, meerdere proefpersonen dezelfde problemen benoemen. Één probleem wordt 'zelfs' door tien proefpersonen van de 173 reacties aangestipt. Dat punt betreft tevens het probleem met de grootste mate van ernst: de ophaaltijden van het huisvuil zijn niet voor alle adressen bekend. De gemiddelde ernst van de problemen bij de proefpersonen van Infocus is 2,1 op de schaal van Nielsen.

Hieronder zullen de meest voorkomende reacties op de pagina Afval van de feedbackbutton beschreven worden. Ook hier zullen ze ingedeeld worden op mate van ernst middels de severity rating van Nielsen.

Binnen de 390 relevante berichten zijn er 67 geplaatst op de pagina's over Afval. Vijftien reacties daarvan betreffen de onduidelijkheid over de ophaaldagen tijdens de weken van Kerst en de jaarwisseling. Deze informatie lijkt nergens te vinden te zijn, terwijl de bezoekers juist dan erg de behoefte hebben te weten wanneer hun huisvuil opgehaald wordt. Omdat dit belangrijke informatie is en het gebrek aan deze informatie gebruikersproblemen teweeg brengt, krijgt dit probleem schaalwaardering 4.

Aansluitend werden er dertien reacties geplaatst over de onduidelijkheid van de reguliere ophaaldagen van afval. Er wordt gemeld dat men de gemeente moet bellen om te vragen naar de ophaaldagen, in plaats van dat er gewoon een schema staat: "Om erachter te komen wanneer ik het huisafval aan de straat moet zetten, moet ik u bellen begrijp ik (uiteraard alleen tijdens kantooruren). Ik zou graag zien dat deze informatie gewoon op de website aanwezig is". Ook dit probleem krijgt schaalwaardering 4 omdat dit de meeste essentiële informatie op deze webpagina is. Deze zou niet mogen ontbreken.

Eenzelfde probleem als wat gemeld wordt bij Infocus is het gebrek aan een duidelijke schemakaart waarop alle afvalpunten staan aangegeven. Tien mensen klagen hierover. Omdat dit in de meeste gevallen een tip is ("Het zou handig zijn als er een plattegrond aan de site gekoppeld is waarop de

afvalscheidingstations zijn aangegeven.”) en geen direct gebruikersprobleem is, krijgt dit punt schaalwaardering 2.

Acht proefpersonen klagen over verouderde informatie. Dit staat volgens hen onzorgvuldig op een gemeentelijke website. Omdat er geen belangrijke informatie mist in dit geval, krijgt dit probleem schaalwaardering 2.

Verder wordt er door vijf bezoekers geklaagd over een onduidelijk en niet-werkend aanvraagformulier voor het ophalen van grofvuil. Dit moet uiteraard werken, maar omdat de bezoekers ook nog de mogelijkheid hebben om te bellen, krijgt dit probleem schaalwaardering 2. Dat de gemeente de bezoeker meerdere opties geeft om het grofvuil te laten ophalen is immers een goed punt.

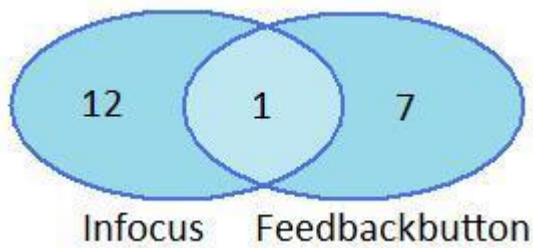
Vier mensen weten niet wat onder elke vorm van afval valt en zouden daar graag wat informatie over willen, maar deze ontbreekt. Omdat dit basisinformatie betreft, krijgt dit probleem schaalwaardering 3. Op een pagina over afval mag verwacht worden dat er staat wat elke soort afval inhoudt.

Daarnaast klagen drie mensen over het gebrek aan een klachtenformulier wanneer het huisvuil niet is opgehaald. Het klagen werd om die reden gedaan via de feedbackbutton. Omdat dit geen groot probleem betreft, maar de bezoekers wel de mogelijkheid zouden moeten hebben om een klacht achter te laten, krijgt dit probleem schaalwaardering 2.

Tot slot blijken drie bezoekers niet te weten wat er gedaan moet worden met frituurvet. Zij dachten dat op de pagina Afval van de gemeentelijke website te vinden, maar helaas was dit niet het geval. Echter, dit is algemene informatie en valt op internet elders te vinden. De informatie is niet puur gemeente gerelateerd. Deze informatie had er echter wel mogen staan, dus daarom krijgt dit probleem schaalwaardering 2.

Uit bovenstaande gegevens blijkt dat echte bezoekers de meeste problemen hebben met de onduidelijkheid over de reguliere en afwijkende ophaaldagen. Dit zal ook de belangrijkste reden zijn waarom inwoners van een gemeente deze pagina's zullen bezoeken. Wanneer de informatie dan niet gevonden kan worden, is er sprake van een groot probleem en dus een hoge mate van ernst. De gemiddelde ernst van alle genoemde punten is 2,6 op de schaal van Nielsen. Dit is een iets hogere score dan bij Infocus.

Het is wederom duidelijk dat Infocus meer commentaar teweeg brengt, maar dat deze over veel meer diverse en kleinere problemen gaan. Er zijn een stuk minder problemen geïdentificeerd op de pagina Afval dan bij Infocus. Onderstaande afbeelding geeft dit helder weer:



Figuur 4: Weergave van de unieke en overeenkomende problemen bij Infocus en de feedbackbutton op de pagina's Afval

De afbeelding geeft weer dat Infocus wederom meer diverse reacties heeft opgeroepen dan de feedbackbutton. Slechts één probleem wordt bij beide meetmethoden genoemd. Dit is overigens weer het probleem met de grootste mate van ernst, namelijk de onduidelijkheid over de reguliere ophaaldagen van het afval. Dit probleem heeft schaalwaardering 4 gekregen. Net als bij de pagina's Parkeren en Zwemmen, gaan ook hier een stuk meer reacties van de feedbackbutton over dit grootste probleem, in vergelijking met die van Infocus (19,4 % tegenover 5,8 %). Wederom geeft de feedbackbutton dus de duidelijkste resultaten wat betreft ernst én kans.

Kortom; de gemiddelde mate van *ernst* ligt hoger bij de reacties van de feedbackbutton op alle bovengenoemde pagina's. Daarmee geeft de feedbackbutton meer belangrijke problemen aan het licht dan Infocus. Daarnaast is de *kans* van de aanwezigheid van de problemen die genoemd worden bij de feedbackbutton een stuk groter, omdat deze een stuk vaker gemeld worden dan bij Infocus.

Overigens betekent dat niet dat de overige unieke problemen die aangekaart worden bij Infocus niet het overwegen waard zijn, maar de meest relevante, meest ernstige en meest frequente problemen kunnen het best bepaald worden aan de hand van de reacties op de feedbackbutton.

Wat zijn de kosten en opbrengsten van beide meetmethodes?

Helaas is het niet mogelijk geweest te achterhalen hoeveel de feedbackbutton en Infocus precies kosten. Echter, er kan wel een globaal beeld gegeven worden.

Afname

Wat betreft Infocus: voor deze evaluatiemethode moeten proefpersonen verzameld worden. Proefpersonen worden doorgaans betaald. Elke proefpersoon kreeg 10,- euro per deelname aan het onderzoek van Henssen (2010). Dan kost een onderzoek met 100 proefpersonen bij Infocus al 1000,- euro. Dan zijn de overige kosten nog niet eens meegenomen, zoals de labkosten en personeelskosten van onderzoeksmedewerkers.

Bij de feedbackbutton reageren websitebezoekers uit zichzelf, dus deze hoeven niet betaald te worden. De enige kosten die aanvankelijk gemaakt zijn, zijn back-office kosten voor het ontwerpen en plaatsen van de feedbackbutton. Als er 100 bezoekers een bericht hebben achtergelaten en daar wordt verder nog niks mee gedaan, heeft de gemeente daar dus nog geen cent hoeven aan te betalen. Op dit vlak is de feedbackbutton goedkoper.

Verwerking

Nadat de proefpersonen van Infocus hun bijdrage hebben geleverd is hun taak direct klaar. Bij de feedbackbutton ligt dat anders. Deze websitegebruikers laten vaak een reactie achter in de hoop dat er naar geluisterd en gehandeld wordt en verwachten vaak een antwoord. En gemeente Utrecht speelt hier netjes op in. De reacties die via de feedbackbutton binnen komen gaan automatisch naar de juiste vakafdeling die ze kan beantwoorden/afhandelen. Enkele reacties komen niet automatisch binnen bij de juiste vakafdeling omdat sommige redacteuren per ongeluk een foutje in het adres opgeven of omdat het adres nog niet gekoppeld is of vanwege een ander onbekend tijdelijk probleem. Deze reacties komen binnen in de inbox van de webmasters. Hieronder wordt uitgelegd wat er vervolgens mee gebeurt.

Afhandeling

Het commentaar van collega's/werknemers binnen gemeente Utrecht wordt, afhankelijk van het type commentaar, verwerkt/beantwoord of doorgestuurd naar de juiste vakafdeling. Voor het relevante commentaar van burgers (bijvoorbeeld een ontdekte typfout) wordt de burger bedankt voor zijn/haar reactie en vervolgens wordt de fout direct aangepast of doorgestuurd naar de juiste vakafdeling. Indien het niet bekend is waar het naar toe moet wordt het bericht doorgestuurd naar het Klant Contact Centrum (KCC). Zij nemen de vraag dan over en nemen contact op met de burger. Ook voor irrelevant commentaar van burgers (bijvoorbeeld aanvragen van leningen) wordt de burger bedankt voor zijn/haar reactie. Vervolgens wordt verteld waar ze wél terecht moeten met hun vraag. Ze worden dus naar de webpagina gestuurd waar het juiste formulier staat of zij krijgen de contactinformatie van de vakafdeling waar ze terecht kunnen. Indien het niet bekend is waar het naar toe moet, stuurt de werknemer het door naar het Klant Contact Centrum (KCC). Zij nemen de vraag dan over en nemen contact op met de burger. Het non-commentaar (bijvoorbeeld onleesbare teksten of reclames) wordt voor een deel afgevangen door een spamfilter. De overige berichten worden weer voor een deel gefilterd en na een paar dagen verwijderd (webmaster www.utrecht.nl, 02-06-2010).

Uit bovenstaande opsomming blijkt dat er erg veel tijd en werk zit in het afhandelen van de reacties op de feedbackbutton. Elke reactie van een bezoeker wordt beantwoord, of deze nou relevant of irrelevant is. Wat de kosten betreft kunnen hier de werkuren van de werknemers binnen gemeente Utrecht geteld worden.

Deze werknemers hebben wellicht meer werk in het afhandelen van het commentaar dan de onderzoekers bij Infocus. Daarnaast is Infocus een stuk helderder: proefpersonen kunnen exact markeren waar het probleem zich bevindt op de pagina. Deze punten kunnen dus ook snel en makkelijk aangepast worden. Bij de feedbackbutton moeten de werknemers eerst de betreffende pagina openen en opzoeken waar de websitebezoeker zijn commentaar op heeft gericht. Het verhelpen van problemen lijkt daardoor een lastigere klus dan bij Infocus.

Kortom, wat de kosten in proefpersonen betreft is de feedbackbutton voordeliger, maar wat de kosten in werkuren betreft is wellicht Infocus goedkoper. Wat beide meetmethoden opleveren is lastig te vergelijken aangezien zij, wat blijkt uit eerder genoemde resultaten, verschillende punten aanstippen. Daar waar Infocus handig is voor het aanpassen van kleine oneffenheden, geeft de feedbackbutton veel inzicht in grote gebruikersproblemen. Om die reden leveren beide methoden veel inzicht, hetzij op andere vlakken.

Is er een verschil in kwaliteit van de verbalisering van de gebruikers?

Zoals hierboven al beschreven, geeft Infocus meer inzicht in kleinere oneffenheden op een website en toont de feedbackbutton juist de grote ergernissen van websitegebruikers. Dit zijn twee verschillende soorten gegevens. Om te controleren hoe deze gegevens overgebracht worden, zijn de verbalisering van het commentaar van beide meetmethodes op de pagina Parkeren met elkaar vergeleken.

Omdat de echte gebruikers voornamelijk een reactie melden als zij de voor hen relevante informatie niet kunnen vinden, wordt dat vaak in het commentaar aangegeven met frases als "Ik kan niet vinden waar...". Omdat de proefpersonen bij Infocus niet zozeer zelfstandig naar iets op zoek zijn, maar gevraagd zijn de pagina te bekritisieren wordt het commentaar vooral beschreven als "Ik vind dit..." of "Ik zou hier...". Bij de feedbackbutton kan om die reden vooral gesproken worden van 'een klacht' wanneer een websitebezoeker een reactie plaatst, terwijl het commentaar bij Infocus voornamelijk louter 'opmerkingen' zijn. Om terug te komen op de vorige deelvraag, waarbij de mate van ernst van het commentaar van beide meetmethoden bepaald werd, kan ook hier gesteld worden dat de mate van ernst bij de feedbackbutton een stuk hoger ligt, aangezien het hier daadwerkelijk klachten betreffen in plaats van opmerkingen of tips.

Een ander punt wat opviel met betrekking tot verbalisering is het verschil in nauwkeurigheid. Uit de geanalyseerde data bleek de nauwkeurigheid van de verbalisering gerelateerd te zijn aan de relevantie voor de website. Een klein tweederde deel van al het commentaar op de feedbackbutton is irrelevant (5090 reacties). Dit commentaar bevat vooral automatisch ingevoerde reclame (voornamelijk voor Viagra en porno) (18,7 %), non-commentaar (15,3 %), testcommentaar van de webmaster (14,3 %) en klachten die niet op de website gericht zijn en reacties die niet via de feedbackbutton ingevoerd moeten worden, zoals leningaanvragen of sollicitaties (51,7 %).

De verbalisering van laatstgenoemde irrelevante reacties bevatten veel spelfouten en de zinnen zijn grammaticaal vaak onvoldoende (33 %). Bij de relevante reacties is dit een stuk minder te zien (10,8 %). De linguïstische kennis van de websitebezoekers voor wie het onduidelijk is waar de informatie die zij invoeren via de feedbackbutton eigenlijk geplaatst moet worden, ligt volgens deze resultaten lager dan die van de bezoekers die relevante reacties plaatsen. Als mensen al niet weten waar ze terecht moeten voor bijvoorbeeld een lening, blijken zij zulke problemen ook niet goed te kunnen verwoorden. Voor vervolgonderzoek zou bekeken kunnen worden of de intelligentie van websitebezoekers gerelateerd is aan de relevantie en nauwkeurigheid van hun reacties. Meer hierover in de Discussie.

Wanneer nog kritischer gekeken wordt naar het relevante commentaar, kan bepaald worden of de reacties volledige goede reacties zijn. Een goede reactie op een feedbackbutton bevat namelijk drie vaste elementen:

Het object – Is de plaats van het commentaar duidelijk?

De inhoud – Is de strekking van de boodschap duidelijk?

De oplossing – Wordt er een mogelijke verbetering aangekaart?

Op de pagina Parkeren is van het afgebakende materiaal (zonder de dubbele reacties en het commentaar van medewerkers) vastgelegd in hoeverre deze reacties voldoen aan bovenstaande eisen. De reacties zijn relevant, maar betekent dat ook dat het volledig correcte reacties zijn?

Wanneer elke reactie gemeten wordt op bovenstaande drie punten kan elke reactie dus maximaal drie punten 'verdienen'. Onderstaande tabel geeft de samenvatting van de analyse over de 93 reacties op de pagina's Parkeren en de gemiddelde score per reactie:

Element	Aantal akkoord (% van de 93)	
Object	92	(98,9 %)
Inhoud	84	(90,3 %)
Oplossing	30	(32,6 %)
Totaal (gemiddelde score per reactie)	206	(206 / 93 = 2,2)

Tabel 10: Score van de 93 reacties per element en gemiddelde score per reactie op de pagina's Parkeren bij de feedbackbutton

Uit bovenstaande tabel blijkt dat vrijwel alle reacties duidelijk commentaar bevatten (92 van de 93). De zinnen zijn dus te begrijpen. De strekking van de reacties is helaas niet altijd volledig begrijpelijk (84 van de 93). Een voorbeeld van een onduidelijke reactie is "Ik krijg een andere gebiedskaart te zien dan verstrekt door Parkeerbeheer". Hierbij is niet direct duidelijk wat het probleem, de vraag of de klacht van de bezoeker

is. De boodschap is helder, dus krijgt deze reactie wel een positieve score wat het object betreft, maar webmasters tasten wellicht in het duister als ze de strekking van deze reactie proberen te achterhalen. Slechts 30 van de 93 reacties bevatten een duidelijke oplossing. Een voorbeeld van een concrete oplossing is “Ik kan nergens vinden of ik in gebied 1, 2 of 3 woon. Misschien een idee om bij veelgestelde vragen te zetten, of een linkje met een kaartje toevoegen”. Verder wordt in veel gevallen impliciet een oplossing gegeven. Wanneer een bezoeker schrijft “Op deze site mis ik een overzichtskaart met de gebiedsindeling” geeft deze impliciet aan dat deze kaart wel erg gewenst is.

Verder geeft de tabel aan dat de gemiddelde score per reactie een 2,2 is. Dit is geen slechte score, aangezien de hoogst mogelijke score een 3 is. Maar wanneer meer bezoekers een directe oplossing vermelden in hun reactie, is het commentaar vollediger en dus beter bruikbaar voor de webmasters.

Bij Infocus komt de volgende tabel naar voren, wanneer gekeken wordt naar de 140 berichten op de pagina Parkeren:

Element	Aantal akkoord (% van de 140)
Object	120 (85,7 %)
Inhoud	140 (100 %)
Oplossing	62 (44,3 %)
Totaal (gemiddelde score per reactie)	322 (322 / 140 = 2,2)

Tabel 11: Score van de 140 reacties per element en gemiddelde score per reactie op de pagina's Parkeren bij Infocus

Uit bovenstaande tabel blijkt dat bij een groot deel van alle reacties duidelijk is waarop het commentaar gericht is (85,7 %). Dit is echter een stuk lager dan bij de feedbackbutton. Een reden hiervoor kan zijn, dat de proefpersonen bij Infocus visueel kunnen markeren waarop zij hun commentaar richten. Bij de feedbackbutton is dit niet het geval. Wanneer de proefpersonen van Infocus al visueel op de wegpagina gemarkeerd hebben waar hun commentaar over gaat, hoeft dit niet meer tekstueel vermeld te worden. Bij de feedbackbutton zijn daarom de reacties wat uitgebreider als het gaat om de plaats van het commentaar.

De strekking van de reacties is juist wel in alle 140 reacties helder. Vergeleken met het commentaar van de feedbackbutton is het commentaar van Infocus dus een stuk beter te begrijpen. Dit heeft waarschijnlijk te maken met het verschil in gebruikers. De proefpersonen van Infocus waren hoogopgeleide studenten. De echte gebruikers van de website die reageren op pagina's hebben verschillende leeftijden en opleidingsniveaus.

Bij Infocus worden iets meer oplossingen gegeven in de reacties (44,3 % tegenover 32,6 %). Ook dit komt waarschijnlijk door het opleidingsverschil tussen de gebruikers. Websitegebruikers met een hoog opleidingsniveau zijn eerder geneigd oplossingen te bedenken en te vermelden dan laagopgeleiden. Verder

geeft de tabel aan dat de gemiddelde score per reactie een 2,2 is. Dit is dezelfde score als bij Infocus.

Opvallend is ook dat van de 140 reacties via Infocus slechts 7 spel- of grammaticale fouten bevatten (5 % van 140 reacties). Bij de relevante reacties op de feedbackbutton is dit dubbel keer zoveel (10,8 %). Hier komt wederom naar voren dat de hoogopgeleide proefpersonen van Infocus duidelijkere en linguïstisch betere reacties plaatsen dan de echte websitegebruikers.

Kortom, de reacties bij Infocus zijn minder duidelijk wat betreft de plaats van het commentaar, maar zijn juist begrijpelijker qua inhoud en bevatten vaker oplossingen dan de reacties op de feedbackbutton. Hierbij moet wel vermeld worden dat de plaats van het commentaar bij Infocus alsnog duidelijk is, omdat dit visueel gemarkeerd kan worden.

Conclusie

In dit onderzoek is achterhaald op welke fronten de feedbackbutton en Infocus van elkaar verschillen. Aan de hand van deelvragen is grondig nagegaan of deze verschillen aanwezig zijn, wat de oorzaken daarvan kunnen zijn en wat deze verschillen teweeg brengen. In deze Conclusie zullen de resultaten van deze deelvragen bondig herhaald worden wat zal resulteren in een antwoord op de hoofdvraag.

Op welke dimensies richt het commentaar zich?

Uit de resultaten van deze eerste deelvraag blijkt dat de feedbackbutton en Infocus beide vooral reacties oproepen over de inhoud van een pagina. Echter, als de reacties grondiger geassocieerd worden in de dertien eerder genoemde subcategorieën, blijkt dat deze niet in dezelfde subcategorieën geplaatst kunnen worden. Via Infocus blijken gebruikers zich veel meer te richten op de vormgeving van de pagina dan via de feedbackbutton.

Wanneer gekeken wordt naar de categorieën van de WEQ, blijken deze geen gehoor te geven aan alle problemen van een website. De WEQ is onderverdeeld in drie hoofdcategorieën, namelijk *Inhoud*, *Bereikbaarheid* en *Vormgeving*. De categorie *Inhoud* is vervolgens onderverdeeld in Relevantie, Begrijpelijkheid en Volledigheid, de *Bereikbaarheid* in Gebruikersgemak, Hyperlinks, Structuur, Snelheid en Zoekmachine. Maar uit bovenstaande resultaten blijkt dat én de problemen van de proefpersonen van Infocus én de problemen van de echte websitegebruikers niet allemaal in te delen zijn in deze categorieën. De categorie *Bereikbaarheid* in de WEQ bevat te veel irrelevante subcategorieën, zoals Zoekmachine, terwijl *Vormgeving* geen enkele subcategorieën bevat. Daardoor blijven er altijd problemen over die nergens exact en helder in te delen zijn. Hierover meer in de Discussie.

Verder verschilt ook de hoeveelheid aan informatie bij beide meetmethoden. Infocus geeft gehoor aan veel problemen, die vaak allemaal eenmalig genoemd worden. De feedbackbutton geeft juist grote gebruikersproblemen aan die veelvuldig gemeld worden door de gebruikers. Deze hoeveelheid aan reacties is in relatie te brengen met de relevantie. Hierover meer bij de volgende deelvraag.

In hoeverre is het commentaar relevant?

De relevantie van het commentaar is ingeschaald aan de hand van de severity rating van Nielsen. Elk commentaar op de pagina Parkeren is bij beide evaluatiemethoden geassocieerd met een schaalwaarde. Een hogere waarde betekent een grotere mate van ernst. De reacties bij Infocus bleken een lagere mate van ernst te hebben dan de klachten bij de feedbackbutton. Ook de mate van kans verschilde bij beide

meetmethodes. Bij de feedbackbutton werd het grootste probleem (de onduidelijkheid over de parkeergebieden) in bijna de helft (47 %) van alle reacties benoemd, terwijl bij Infocus in slechts 7 van de 140 reacties (5 %) dit punt aan de orde kwam. De relevantie van een probleem komt bij Infocus dus in mindere mate naar voren dan bij de feedbackbutton, omdat bij laatstgenoemde methode de bezoekers alleen grote gebruikersproblemen melden. Bij Infocus worden kleinere oneffenheden gemeld, die uiteraard ook wel van belang zijn, maar om grote problemen te traceren kan het woord beter gegeven worden aan de echte gebruikers.

Wat zijn de kosten en opbrengsten van beide meetmethodes?

Ook qua kosten verschillen beide evaluatiemethoden van elkaar. Infocus is wellicht duurder in opzet en toepassing, aangezien hiervoor een ruimte en proefpersonen geregeld en betaald worden moeten. Dit is bij de feedbackbutton niet het geval.

Echter, via Infocus is direct waar te nemen waar het commentaar zich op focust. Daarnaast is elk commentaar gericht op de website omdat deze evaluatiemethode volgens een opgezet onderzoek uitgevoerd wordt. Bij de feedbackbutton is een overgroot deel van het commentaar irrelevant en moet er nog goed nagezocht worden waar het commentaar zich precies op focust, aangezien dit niet exact aangegeven kan worden door de bezoekers. Hierdoor gaan er behoorlijk wat uren (dus kosten) zitten in het werk de website te verbeteren aan de hand van de reacties op de feedbackbutton. Dit is bij Infocus in mindere mate het geval.

Is er een verschil in kwaliteit van de verbalisering van de gebruikers?

Wat direct opvalt bij het vergelijken van de verbalisering van de reacties op de feedbackbutton en van het commentaar van Infocus is dat de echte gebruikers van de website hun commentaar vooral als klacht formuleren en de proefpersonen bij Infocus enkel opmerkingen plaatsen. De reden hiervoor is dat de echte gebruikers daadwerkelijk de voor hen relevante informatie willen vinden. Lukt dit niet, dan is er reden tot klagen. De proefpersonen bij Infocus hoeven niks te vinden uit eigenbelang, maar hebben enkel de opdracht een opgestelde zoektaak te voltooien en opmerkingen te leveren op de website. Wanneer iets in de ogen van de proefpersoon tegenvalt of verbeterd kan worden, is dat nog geen reden om erover te klagen.

Wat betreft de nauwkeurigheid van de reacties is deze bij Infocus minder duidelijk wat de plaats van het commentaar betreft, maar zijn deze reacties juist begrijpelijker qua inhoud en bevatten vaker oplossingen dan de reacties op de feedbackbutton. Hierbij moet wel vermeld worden dat de plaats van het commentaar bij Infocus alsnog duidelijk is, omdat dit visueel gemarkeerd kan worden.

Kortom, om inzicht te krijgen in de gebruiksvriendelijkheid van een website is het verstandig een feedbackfaciliteit op elke pagina te plaatsen. Via deze evaluatiemethode plaatsen echte gebruikers klachten en opmerkingen over de website die voornamelijk betrekking hebben op grote gebruikersproblemen zoals ontbrekende of onduidelijke informatie en technische problemen.

Wanneer je meer wilt weten over hoe een website optimaal verbeterd kan worden is een onderzoek met Infocus van belang. Via deze evaluatiemethode worden enorm veel verschillende oneffenheden benoemd door daarvoor opgeroepen proefpersonen. Deze problemen blijken van mindere ernst dan die van de feedbackbutton, maar richten zich wel op meerdere dimensies.

Om een website op alle dimensies te optimaliseren én optimaal gebruiksvriendelijk te maken zouden beide evaluatiemethoden ingezet kunnen worden. Uit dit onderzoek blijkt namelijk dat beide methoden verschillende inzichten bieden en tot verschillende resultaten leiden. De ene methode is om die reden niet beter dan de ander. Ze verschillen van elkaar, maar zijn beide van belang voor het verbeteren van websites.

Discussie

In bovenstaande Conclusie komt naar voren dat beide evaluatiemethoden op verschillende wijze inzicht bieden in de kwaliteit van gemeentelijke websites. Echter, om een nog duidelijkere vergelijking te kunnen maken tussen de feedbackbutton en Infocus zouden enkele punten uit dit huidige onderzoek meegenomen moeten worden voor eventueel vervolgonderzoek. Deze komen hier aan bod.

In de Resultaten kwam naar voren dat de meest bezochte pagina's niet direct de pagina's zijn waar het meest op gereageerd wordt. De pagina's waar het meest op gereageerd wordt door echte websitegebruikers zijn de pagina's waar het meest over te klagen valt. Om deze pagina's optimaal te kunnen verbeteren zouden juist deze pagina's ook deel uit moeten maken van een onderzoek met Infocus. Het is immers van belang juist de meest gebrekkige pagina's te verbeteren in het belang van de gebruiker. In een vervolgonderzoek zouden dus juist de pagina's waar websitegebruikers veel commentaar op leveren via de feedbackbutton tevens geëvalueerd moeten worden met Infocus om een nog uitgebreidere analyse van te kunnen maken.

Een ander opvallend punt is de relatie tussen de relevantie van berichten en de kwaliteit van de verbalisering. Fout geplaatste opmerkingen op de feedbackbutton, over bijvoorbeeld geboorteaanvragen of uitkeringen, zijn regelmatig grammaticaal onvoldoende. Burgers die dus niet lijken te weten waar zij terecht moeten voor zulke zaken, blijken ook taalkundig laag te presteren. Voor vervolgonderzoek zou onderzocht kunnen worden of er inderdaad een relatie is tussen taalvaardigheid en de relevantie van commentaar. Hierbij kan er ook een afbakening gemaakt worden van de irrelevante berichten, zodat bekeken kan worden of de kwaliteit in verbalisering daadwerkelijk significant verschilt met die van de relevante berichten. Daarbij kan meegenomen worden of commentaar van één persoon als relevant beschouwd moet worden. Bij Infocus is dit wel gedaan. Bij de feedbackbutton wordt elke reactie gelezen en wordt de bezoeker die het commentaar heeft geplaatst ervoor bedankt, maar is het onbekend of er naar gehandeld wordt. Als één persoon klaagt over een aspect op een webpagina, is dat dan een punt om aan te passen of ligt dat dan alleen aan die ene persoon?

Verder bleek het lastig te zijn al het commentaar in te delen in de eerder opgestelde categorieën door Henssen (2010). Hij had nieuwe categorieën ontwikkeld, omdat enkel de hoofdcategorieën van de WEQ niet voldeden. Echter, alsnog had Henssen een categorie Overig nodig, omdat de reacties niet allemaal ergens geplaatst konden worden. Ook de reacties van de echte websitegebruikers via de feedbackbutton konden niet allemaal geclassificeerd worden met de opgestelde categorieën vanuit de WEQ. Voor vervolgonderzoek zou dus een nieuwe protocolanalyse ontwikkeld moeten worden, waardoor al het commentaar te classificeren is, met elkaar vergeleken kan worden en meer inzicht in de kwaliteit van een

website biedt.

Wat er wordt gedaan met het commentaar is direct een punt voor een nieuw onderzoek. Wat is de relatie tussen de relevantie van commentaar en de mate waarin er wat mee gedaan wordt? Voor vervolgonderzoek zou gemeten kunnen worden in welke toestand een webpagina er eruitziet en functioneert in de nulsituatie en hoe de pagina eruitziet en functioneert na een bepaalde tijd waarin gebruikers commentaar hebben geleverd. Zo kan gemeten worden welk commentaar tot welke aanpassingen heeft geleid.

Infocus en de feedbackbutton leveren beide veel inzicht in de kwaliteit van gemeentelijke websites. Hetzij op verschillende manieren. Om het commentaar op de twee evaluatiemethoden nog beter met elkaar te kunnen vergelijken zou een nieuwe protocolanalyse ontworpen moeten worden waarbij het commentaar van beide methoden geassocieerd kan worden. Op die wijze is het vergelijken van beide zeer diverse methoden overzichtelijker en geeft het meer duidelijkheid in de methoden aan sich.

Literatuur

De Jong, M., & L. Lentz (2001). Focus: Design and evaluation of a software tool for collecting reader feedback. *Technical Communication Quarterly*, 10, 387-401.

Van den Haak, M.J., M.D.T. de Jong & P.J. Schellens (2007). Evaluation of an informational Web site: Three variants of the think-aloud method compared. *Technical Communication*, 54 (1), 58-71.

Scripties

Henssen, P. (2009). Masterscriptie 'Infocus onder de Loep', Universiteit Utrecht.

Henssen, P. (2010). Stageverslag 'Gebruikersonderzoek met Infocus', Universiteit Utrecht.

Anh, V. (2007). Masterscriptie 'Uw opmerkingen over de website horen wij graag', Universiteit Utrecht

Websites

Stijlgids EGEM, <http://www.egem-iteams.nl/stijlgids-gemeenten-03>, laatst bezocht op 11-04-2010

Evaluatie van gemeentelijke websites, <http://www.gemeentewebsites.nl>, laatst bezocht op 29-06-2010

Web Evaluation Questionnaire (WEQ), <http://www2.hum.uu.nl/onderzoek/weq/index.html>, laatst bezocht op 06-08-2010

Jakob Nielsen's Website, <http://www.useit.com>, laatst bezocht 28-06-2010

Bijlagen