

Eindopdracht Geëngageerde wetenschappen

W.W.Beukers
sn:1516051

11 januari 2005

1 Inleiding

1.1 Invloed van de ICT op de maatschappij

De invloed van de ICT op de maatschappij is enorm groot geworden. Tientallen jaren geleden had men niet kunnen voorspellen wat de ICT teweeg zou brengen in de maatschappij. Het is nu zelfs zo ver gekomen dat de maatschappij niet meer zonder de ICT kan. Er zijn ontelbare voorbeelden van waar computers in de hedendaagse maatschappij worden toegepast. Denk bijvoorbeeld aan het betalingsverkeer, bedrijfsadministraties, etc. Deze ontwikkeling zijn begin jaren '90 in een stroomversnelling geraakt. Door de snelle toename van de rekenkracht van computers en de opkomst van het internet is de ICT een essentieel onderdeel van de maatschappij geworden.

1.2 Invloed van de ICT op de wetenschap

Ook voor de wetenschap heeft deze ontwikkeling zeer veel veranderd. Door de opkomst van de ICT hebben onderzoekers nu de mogelijkheden gekregen om bepaalde onderzoeken uit te voeren die voorheen nog niet mogelijk waren. Denk bijvoorbeeld aan het onderzoek naar het DNA van de mens. Ook deze ontwikkeling zijn ontstaan door de toenames van de rekenkracht en het internet.

1.3 Toekomst van de ICT binnen de geesteswetenschap

Er is onlangs een rapport verschenen, namelijk: Geëengageerde Geesteswetenschappen: Perspectieven op cultuurveranderingen in een digitaliserend tijdperk[2]. Dit rapport is geschreven door het AWT. Het AWT is een adviesorgaan van de Regering en de Staten-Generaal. In dit rapport staat beschreven wat de rol moet zijn van de ICT bij de geesteswetenschappen. De centrale boodschap van het rapport is, een geesteswetenschappelijk perspectief kan belangwekkende inzichten opleveren wanneer men deze maatschappelijke ontwikkelingen probeert te begrijpen en/of probeert bij te sturen. Daarom pleiten ze voor de ontwikkeling van een vorm van geëengageerde geesteswetenschappen. Deze studie richting moet dus vooral onderzoeken wat voor invloed ICT heeft op de maatschappij. Dus ze vinden het minder van belang hoe de ICT kan worden toegepast binnen de geesteswetenschappelijke disciplines maar dat de aandacht juist gevestigd moet worden op de vragen die de maatschappelijke ontwikkelingen en de geesteswetenschappelijke oproepen. De reden dat ze het toepassen van ICT minder van belang vinden is dat ze het alleen toepassen van de ICT te beperkt vinden.

De reden waarom de de auteurs deze mening hebben wordt hieronder toegelicht. Aan de hand van een vijftal samenhangende deelgebieden wordt besproken welke ICT-gerelateerde vragen er zijn ontstaan en onderzocht moeten worden door de geesteswetenschappen. Hieronder worden de vijf deelgebieden besproken en nader toegelicht.

1.3.1 Ethiek

Door de opkomst van de ICT zijn er vele ethische vraagstukken ontstaan. Dit wordt mede veroorzaakt door de snelle ontwikkeling van de ICT. Door de snelle ontwikkeling hiervan loopt de wetgeving hierdoor flink achter. Omdat de wetgeving flink achterloopt komt de verantwoordelijk over bepaalde ethische

vraagstukken bij de gebruikers en de bedrijven te liggen, daardoor is er een enorm grijs gebied ontstaan over wat wel en wat niet is toegestaan bij het toepassen van ICT. Een andere oorzaak van het ontstaan van ethische dilemma's is de globalisering door het internet. Het internet is niet specifiek gebonden aan een land en dus ook niet gebonden aan de wetgeving van een land. Hierdoor ontstaan ook allerlei ethische dilemma's.

Hieronder staan een aantal voorbeelden van ethische vraagstukken. De auteursrechten van muziek, films en software worden door de opkomst van het internet en vooral de share-programma's geschonden. Op dit moment is het vrij gemakkelijk om de nieuwste muziek en films gratis te downloaden.

Een ander voorbeeld hiervan zijn de persoonsgegevens. Steeds meer bedrijven houden allerlei gegevens van de mensen bij, dit gebeurt vaak zonder dat de persoon er weet van heeft. Je kunt hierbij denken dat een bedrijf je ingevulde gegevens op een website doorverkoopt aan derden. Bijvoorbeeld je e-mail adres wordt verkocht aan een spam-verstuurder. Je surfgedrag wordt bijgehouden doormiddel van spyware en aan de hand daarvan krijg je reclameboodschappen. In de toekomst kan dit nog veel verder gaan. Bijvoorbeeld de medische gegevens, je wilt een ziektekostenverzekering afsluiten maar dat wordt je geweigerd omdat er in je familie een erfelijke ziekte zicht voordoet. Dit gebeurt nu nog niet maar het is al wel mogelijk. Wyber en Byker zijn van mening dat deze vraagstukken onderzocht moet worden door de geesteswetenschap. En omdat het niet echt bij een geesteswetenschappelijke richting hoort pleiten zij ervoor om dit onder te brengen bij de door hun voorgestelde geëngageerde wetenschappen.

1.3.2 Identiteit en zingeving

Bij dit deelgebied bespreken ze welke maatschappelijke vragen er zijn ontstaan door de invloed van ICT op het individu. Hierbij behandelen ze vooral wat de invloed is van het internet op het individu. De mening van Byker en Peperkamp over het internet is als volgt: *Duidelijk is in ieder geval dat we het internet niet alleen moeten waarderen in technische zin, dat wil zeggen: als een open netwerksysteem dat bestaat uit vele miljoenen computers die met elkaar verbonden zijn. Het internet moet ook worden opgevat als een complex sociaal netwerk dat is samengesteld uit vele miljoenen individuen en instituties uit inmiddels meer dan tweehonderd landen die dat netwerksysteem gebruiken om te kunnen communiceren en te kunnen handelen.*¹

Doordat een individu redelijk anoniem van het internet gebruik kan maken kan een individu zichzelf een andere identiteit aanmeten. Het gevolg hiervan is dat er een algemeen wantrouwen heerst tegen een digitale identiteit ongeacht of dit overeenkomt met de werkelijk identiteit van een individu. Een andere ontwikkeling is dat mensen met een zelfde mening over een bepaald onderwerp zich gemakkelijk kunnen verenigen via het internet. Een voorbeeld hiervan is de antiglobalisering beweging.

Het internet heeft ook zijn weerslag op de levensbeschouwing van een individu. Dit is veroorzaakt door het internet. De toegang tot religieus en levensbeschouwend materiaal is enorm vergemakkelijkt door het internet. Ook de hoeveelheid van religieus en levensbeschouwen is enorm toegenomen door het internet. Al deze ontwikkeling hebben een invloed op het individu daarom vin-

¹Definitie van het internet door Byker en Peperkamp op pag. 62 [2].

den Byker en Peperkamp het belangrijk dat ook naar deze gevolgen onderzoek wordt gedaan.

1.3.3 Taal

Doordat er ook op een digitale manier wordt gecommuniceerd heeft de ontwikkeling van de ICT ook een direct effect op de taal. Het belangrijkste gevolg van de digitalisering van de taal is dat er grensvervaging optreedt. Je kunt hierbij denken aan grensvervaging tussen: spreektaal en schijftaal, talen onderling.

Doordat er op het internet grotendeels engels wordt gesproken zullen er steeds meer engelse termen in de Nederlandse taal worden opgenomen. Bedrijven profileren zich sneller als internationaal bedrijf en bieden hun website meertalig aan. Al deze ontwikkelingen kunnen gezien worden als een bedreiging of een verrijking van de Nederlandse taal. Wat wel vaststaat is dat de Nederlandse taal door deze ontwikkelingen grote veranderingen zal ondergaan. Byker en Peperkamp van belang is om onderzoek te verrichten naar wat de gevolgen hiervan zijn voor de maatschappij.

1.3.4 Kunst

Door de ontwikkeling van de ICT is er veel veranderd met betrekking tot het publiceren van kunsts. Elk gerenommeerde kunstinstelling heeft op dit moment een website waar op een belangrijk deel van hun kunstcollectie te bezichtigen valt.

Een andere ontwikkeling die in opkomst is de technologisering van de kunst. Door de technologisering is het mogelijk om verschillende vormen van kunst te combineren hierdoor vervagen de klassieke genregrenzen tussen de verschillende kunstvormen. Een bekend begrip wat hier mee te maken heeft is multimedia.

Door het grote aanbod van kunst, mogelijk gemaakt door het internet, bestaat er nu de mogelijkheid om specifiek een bepaalde kunstvorm te bezichtigen. Je kunt hierbij denken aan het 'wandeling' door een virtueel museum. En doordat het internet een wereldwijd karakter heeft, is het nu vrij gemakkelijk om kunst van over de hele wereld te bezichtigen.

Al deze ontwikkeling hebben weer gevolgen voor de maatschappij en zijn dus weer onderzoeksgebieden voor de geëngageerde geesteswetenschappen.

1.3.5 Informatie en kennis

Door de opkomst van de ICT is er veel veranderd met betrekking tot informatie en kennis. Informatie kan nu heel gemakkelijk via het internet worden verspreid. Allerlei wetenschappelijke tijdschriften en wetenschappers publiceren hun wetenschappelijke werken op het internet. Doordat er zoveel informatie beschikbaar is en de kwaliteit van de informatie soms moeilijk te beoordelen is kan het zeer moeilijk worden om juiste informatie over een bepaalde onderwerp te vinden. Dit leidt tot de vraag wat is er vandaag de dag belangrijker voor een student. Moet een student kennis hebben van bepaalde zaken of moet een student het beheersen om die kennis te kunnen vergaren via de nieuwe mogelijkheden die de ICT biedt. Al deze gevolgen voor de maatschappij zijn onderzoeksgebieden voor de geëngageerde geesteswetenschappen.

De algemene strekking van dit stuk is zoals blijkt uit bovenstaande is dat de huidige geesteswetenschap meer aandacht moet besteden aan wat de effecten van de ICT op de geesteswetenschappen. Byker en Peperkamp willen dit bereiken door de opleiding geëngageerde geesteswetenschappen op te starten. Deze opleiding heeft specifiek het doel het bestuderen van wat de effecten zijn van ICT op de maatschappij.

Deze mening is de Groningse Alfa-informatici niet toebedeeld. Zij vinden het juist van belang dat de ICT puur en alleen wordt ingezet als een ondersteunend middel om geesteswetenschappelijk onderzoek te verrichten. De Groningse Alfa-informatici vinden het vooral interessant welke mogelijkheden de ICT biedt voor geesteswetenschappelijk onderzoek. Er ontstaan dus nieuwe mogelijkheden door de ICT maar dat vloeit voort uit het onderzoek en niet door de reflectie van de toenemende integratie van de ICT in de maatschappij.

Aan de hand van een aantal praktijkvoorbeelden laat ik zien dat het toepassen van ICT als ondersteunende rol binnen een onderzoek veel interessanter is dan het onderzoeken van wat voor invloeden de ICT heeft op de maatschappij. Deze praktijkvoorbeelden zijn mogelijk geworden door de ontwikkeling van een groot aantal nieuwe technologieën in de ICT. Een aantal technologieën worden in dit stuk behandeld. De technologieën worden eerst uitgelegd en daarna zullen een aantal geesteswetenschappelijke onderzoeken behandeld worden die mogelijk zijn geworden door deze nieuwe ICT-technologieën.

- Database
- Spraakherkenning
- Classificeren van teksten

2 Nieuwe technolgieen

2.1 Database

In de jaren 60 was Dr. E. F. Codd een onderzoeker voor IBM. Als wiskundige was hij ontevreden over de destijds gangbare manieren van gegevensopslag. Hij kwam tot de conclusie dat alle beschreven modellen gevoelig waren voor fouten en redundantie. Nadat hij enige jaren aan dit onderwerp heeft besteed, publiceerde hij in 1970 een onderzoek met de titel *A Relational Model of Data for Large Shared Databanks*[3]. In dit onderzoek wordt vooral besproken hoe op een gestructureerde manier gegevens opgeslagen kunnen worden en hoe de onderlinge relaties tussen de verschillende gegevens vastgelegd kunnen worden. Door deze publicatie is er veel veranderd in de wereld van gegevensopslag met behulp van databases.

Door de ontwikkeling van de relationele database is er de mogelijkheid ontstaan om grote hoeveelheden data op een gestructureerde manier op te slaan. Maar wat nog veel interessanter hieraan is dat de er allerlei vragen, de zogenaamde queries, over de data gesteld kunnen worden. Door deze mogelijkheden is het dus gemakkelijk om grote hoeveelheden data te onderzoeken en te analyseren.

2.2 Spraakherkenning

De spraakherkenning door computers is in de jaren '70 in een stroomversnelling geraakt. Dit werd veroorzaakt door steeds krachtiger computers en doordat de gesproken tekst op een nieuwe manier werd geanalyseerd[5]. De gesproken tekst werd nu op een grammaticaal niveau geanalyseerd. Dit werd op de volgende manier bereikt. De kans wordt berekend wat het volgende woord zou zijn aan de hand van het laatst gesproken woord. Omdat dit nogal veel rekenkracht vraagt van de ICT is deze ontwikkeling pas eind jaren '80 goed in gang gezet doordat de rekenkracht van de computer heel snel heel veel groter is geworden. Vandaag de dag zijn er tal van toepassingen waarbij spraakherkenning wordt toegepast. Bijvoorbeeld mensen met RSI-klachten dicteren hun teksten in plaats van te typen.

2.3 Classificeren van teksten

Met tekst classificatie worden op een geautomatiseerde manier teksten geanalyseerd en geclassificeerd. Tekst classificatie wordt meestal op de volgende manier uitgevoerd. De belangrijkste woorden worden uit een tekst gehaald en dan vergeleken met een index. Aan de hand van deze vergelijking kan een tekst worden ingedeeld in een bepaalde groep[1]. Een voorbeeld van tekst classificeren is het herkennen van spam-mailtjes.

3 Praktijkvoorbeelden

Door bovenstaande technologieën als ondersteuning toe te passen binnen een geesteswetenschappelijk onderzoek kunnen er baanbrekende resultaten bereikt worden. Hieronder staan een aantal geesteswetenschappelijke onderzoeken die zijn uitgevoerd door de huidige geesteswetenschappelijke disciplines. Al deze geesteswetenschappelijke onderzoeken zijn door vragen uit de maatschappij opgestart.

3.1 MALACH

MALACH² is een project wat is gestart door een aantal wetenschappelijke instellingen³ en door het bedrijfsleven⁴. Het doel van MALACH is het ontwikkelen van methodes om toegang te krijgen tot een grote meertallige database. Deze database is opgezet door de organisatie Survivors of the Shoah Visual History Foundation's⁵. Deze database bevat allerlei interviews, zowel geluid als video, van mensen die de 2e wereld oorlog hebben overleefd. De database bevat op dit moment 52.000 interviews.

Om in de database te kunnen zoeken moet er eerst een soort van index worden gemaakt. De index bevat dan allerlei gegevens van het interview je kunt hierbij denken aan het onderwerp, wat voor taal, plaats, etc. Omdat de database op dit moment 52.000 interviews bevat is het ondoenlijk om dit met de hand te gaan doen.

Daarom moeten de interviews op een automatische manier geïndexeerd worden. Het eerste probleem wat men tegenkomt is dat om video's te kunnen indexeren dat de gesproken woorden uit de video gehaald moet worden met behulp van spraakherkenning. Het volgende probleem is dat de interviews meertalig zijn. Dus al deze interviews moeten vertaald worden naar een hoofdtal, in dat geval is dat engels. Als men al deze problemen heeft opgelost met behulp van ICT-middelen is het mogelijk om een catalogi van de interviews te maken en om specifiek te zoeken naar een bepaald onderwerp.

Door het inzetten van ICT-hulpmiddelen heeft men het doel bereikt van dit geesteswetenschappelijk onderzoek. Door de verschillende aard van de problemen die tijdens dit project naar voren kwamen zijn er wetenschappers ingezet die geschoold zijn in de verschillende geesteswetenschappelijke disciplines.

3.2 Smart-email

Samen met de Universiteit van Groningen ontwikkelde callcenter BSC smart e-mail: digitale post die zelfstandig antwoorden geeft op vragen van klanten. Het scheelt BSC een hoop mankracht en de klant wordt vaak nog beter bediend ook. Alleen moet de computer nog wel wat bijleren[4].

BSC is een call-center die voor bedrijven de klachten en vragen van klanten behandelt. Om de klanten beter tot hun dienst te kunnen zijn proberen ze een select deel van de vragen per e-mail automatisch te beantwoorden. Dit wordt nu

²Multilingual Access to Large spoken ArCHives

³Survivors of the Shoah Visual History Foundation, Johns Hopkins University, University of Maryland, Charles University, University of West Bohemia

⁴IBM Research

⁵Meer informatie over deze organisatie is hier te vinden: <http://www.vhf.org/>

alleen nog maar toegepast voor simpele vragen, bijvoorbeeld: "Ik heb drie weken geleden een folder aangevraagd en ik heb hem nog steeds niet binnen". Door deze tekst te classificeren en daarbij een correct antwoord uit de database te genereren wordt een antwoord op gestelde vraag gegenereerd. Het gegenereerde antwoord wordt nog wel door een persoon beoordeeld omdat het systeem nog niet 100% correct werkt.

Dit doelstellingen van dit project zijn behaalt door ICT-middelen in te zetten ter ondersteuning van het project.

3.3 OVIS

OVIS⁶ is een project dat is uitgevoerd door een aantal onderwijsinstellingen⁷. Het doel van dit project is om op een automatische manier reisinformatie te geven aan reizigers die bellen[6]. Dus een reiziger belt het nummer stelt de vraag: ik wil van Amsterdam naar Rotterdam op 4 februari. OVIS analyseert wat de reiziger heeft gezegd en geeft dan als respons een gesproken bericht met de mogelijke antwoorden. Op dit moment wordt op 83% van de vragen een correct antwoord gegeven. In de toekomst zal dit percentage verhogen omdat er op het gebied van spraakanalyse nog veel verbeteringen mogelijk zijn.

3.4 Question Answering for Dutch using Dependency Relations

Dit project wordt uitgevoerd door een aantal mensen⁸ die verbonden zijn aan de RUG. Daarnaast is er een samenwerkingsverband met de uitgeverij Het Spectrum. Het doel van dit project is om een Nederlands vraag&antwoord systeem te ontwikkelen. Dus een gebruiker kan een vraag stellen aan het systeem en krijgt dan een antwoord terug. Doormiddel van complexe analyse van de vraag wil men het exacte antwoord terug geven.

Dit project is doormiddel van het inzetten van ICT-middelen uitvoerbaar.

Zoals blijkt uit bovenstaande praktijkvoorbeelden kan men door het inzetten van ICT-middelen ter ondersteuning van een geesteswetenschappelijk project baanbrekende resultaten bereikt worden. Dit onderschrijft de mening van de Groningse alfa-informatici.

⁶Openbaar Vervoer Informatie Systeem

⁷IPO Eindhoven, Rijksuniversiteit Groningen, Universiteit Utrecht, Universiteit van Amsterdam

⁸dr. G Bouma, dr. G. van Noord, L. van der Plas, J. Mur, dr. J. Tiedemann

4 Conclusie

De stelling van Byker en Peperkamp dat de geesteswetenschappen meer aandacht moet besteden aan het onderzoeken van wat voor invloed de ICT op de maatschappij heeft vind ik niet helemaal terecht. Ik vind wel dat het onderzoek naar deze invloeden wel gedaan moet worden, maar ik vindt dat ICT als ondersteuning in te zetten voor geesteswetenschappelijk onderzoek veel interessanter is. Want zoals uit de voorgaande praktijkvoorbeelden blijkt dat de ondersteuning van ICT voor geesteswetenschappelijke onderzoeken veel meer mogelijkheden en winst opleveren.

Ook het voorstel van Byker en Peperkamp om de opleiding geëngageerde geesteswetenschappen op te richten vind ik niet nodig. De belangrijkste reden die zij hiervoor geven is dat de huidige geesteswetenschappelijke disciplines niet in staat zijn om de invloeden van ICT op de maatschappij te onderzoeken. Mijn mening is dat de huidige geesteswetenschappelijke disciplines wel in staat zijn om dit onderzoeken. Zoals ook blijkt uit voorgaande praktijkvoorbeelden. Want als een geesteswetenschappelijk onderzoek vraagt om verschillende geesteswetenschappelijke disciplines kunnen er meerdere wetenschappers ingezet worden die geschoold zijn in de verschillende benodigde geesteswetenschappelijke disciplines. Een andere vraag die men kan stellen is, wordt de opleiding geëngageerde geesteswetenschappen een niet te brede en te algemene opleiding? In mijn ogen wel omdat het gebied wat Bijker en Peperkamp aangeven als het onderzoeksterrein voor de geëngageerde geesteswetenschappen is een te breed gebied wat veel verschillende geesteswetenschappelijke disciplines bestrijkt. En daardoor bereikt men niet de diepgang die de huidige verschillende geesteswetenschappelijke disciplines wel bieden.

De vijf thema's die worden behandeld door Byker en Peperkamp leveren inderdaad een aantal interessante geesteswetenschappelijke onderzoeken op. En in mijn ogen kunnen deze onderzoeken uitstekend worden uitgevoerd door de huidige geesteswetenschappelijke disciplines. En waar nodig moeten dus wel multidisciplinaire onderzoeksteams worden samengesteld om de discipline overschrijdend geesteswetenschappelijke onderzoeken te kunnen uitvoeren.

Referenties

- [1] B Baeza-Yates, R. Ribeiro-Neto. *Modern Infromation Retrieval*. ACM Press Books, 1999.
- [2] B. Bijker, W. Peperkamp. Geëngageerde geesteswetenschappen. mei 2002.
- [3] E.F. Codd. A relational model of data for large shared data banks. januari 1970.
- [4] C. Hospes. Smart e-mail: u vraagt, wij mailen. *Management team*, 4, 2001.
- [5] T. Robinson. Speech analysis. Final project report, Machine Intelligence Laboratory, University of Cambridge, 1998. Beschikbaar op <http://svr-www.eng.cam.ac.uk/~ajr/SA95/SpeechAnalysis.html>.
- [6] G. Sima'an K. Noord G. Bonnema R. Veldhuijzen van Zanten, G. Bouma. Evaluation of the nlp components of the ovis2 spoken dialogue system. Evaluation, IPO Eindhoven, Rijksuniversiteit Groningen, Universiteit Utrecht, Universiteit van Amsterdam, 2000. Beschikbaar op <http://http://odur.let.rug.nl/~van Noord/papers/oviseval/oviseval.html>.