

Geesteswetenschappen revised?

Jurrian Akkerman

16 januari 2005

Inhoudsopgave

1	Inleiding	3
2	ICT als ondersteunende factor	4
3	Reflectie culturele invloed van de ICT	4
4	Databases	4
4.1	Het gebruik van databases binnen geesteswetenschappen	4
4.2	De maatschappelijke consequenties	5
4.3	Wisselwerking	5
5	Het WWW	6
5.1	Het gebruik van het WWW binnen geesteswetenschappen	6
5.2	Maatschappelijke consequenties	6
5.3	Wisselwerking	7
6	Grid computing	7
6.1	Het gebruik van grid computing binnen geesteswetenschappen . .	7
6.2	De maatschappelijke consequenties	7
6.3	Wisselwerking	8
7	Information Retrieval	8
7.1	Het gebruik van IR binnen geesteswetenschappen	8
7.2	De maatschappelijke consequenties	9
7.3	Wisselwerking	9
8	Conclusie	10

1 Inleiding

De ICT heeft de afgelopen tien jaar enorme ontwikkelingen meegemaakt. Deze ontwikkelingen zijn te merken op alle vlakken in de maatschappij, dus ook op het terrein van de geesteswetenschappen. De afdeling Informatiekunde aan de RuG houdt zich vooral bezig met een ondersteunende rol voor de ICT binnen de geesteswetenschappen te vinden. Tijdens colleges voor de cursus Inleiding Informatiekunde hebben we dan ook gezocht naar toepassingen die voor de geesteswetenschappen een aanvulling kunnen zijn. In het stuk 'Geëngageerde Geesteswetenschappen' [1] komt echter een andere visie over het gebruik van ICT binnen de geesteswetenschappen aan bod, namelijk dat er voornamelijk aandacht moet zijn op de gevolgen van de steeds sterkere invloed van de ICT binnen de geesteswetenschappen. Deze visie zal ik verderop in deze paper uitgebreider uiteenzetten. In deze paper zal ik allebei de visies uiteenzetten, waarna ik door middel van een aantal voorbeelden een conclusie probeer te trekken over welke visie in mijn ogen het beste is. De voorbeelden splits ik steeds op in twee gedeeltes, één met de vraag: 'Hoe valt dit voorbeeld te gebruiken binnen de geesteswetenschappen?' de tweede vraag 'Wat zijn de maatschappelijke consequenties van dit voorbeeld?' en als laatste vraag: 'Wat is de wisselwerking tussen de twee verschillende visies?'

2 ICT als ondersteunende factor

Zoals al eerder gezegd probeert de afdeling Informatiekunde van de RuG een ondersteunende rol te vinden voor de ICT binnen de geesteswetenschappen. Enkele passend voorbeelden hiervan zijn: Databases die makkelijk te gebruiken zijn met Letterenstudies, het Internet om een snelle verspreiding van bronteksten te realiseren en het automatisch onttrekken van informatie uit historische teksten. In de ogen van deze wetenschappers zijn bovengenoemde toepassingen verlengstukken van het onderzoek dat bij letteren centraal staat. Onderzoek als hoe informatie dragers in het verleden functioneerden, en hoe deze te verbeteren om ze beter in gebruik te maken voor Letterenstudies. Hun mening is kort gezegd: De ICT blijft vooral een ondersteunend middel om geesteswetenschappelijk werk te bevorderen, en om toepassing van geesteswetenschappelijk onderzoek mogelijk te maken.

3 Reflectie culturele invloed van de ICT

Tegenover de visie van wetenschappers als die in Groningen staat de visie die beschreven wordt in 'Geëngageerde Geesteswetenschappen'. Ziet de afdeling Informatiekunde de mogelijkheden voor een ondersteunende rol van de ICT, dan zien Bijker en Peperkamp juist de gevolgen en maatschappelijke consequenties van het gebruik van de ICT. Ze zijn van mening dat er meer gezocht moet worden naar maatschappelijke consequenties en antwoorden op vragen die de steeds grotere rol van de ICT binnen geesteswetenschappelijke gebied met zich mee brengt. Om het simpel te zeggen, Bijker en Peperkamp zijn van mening dat onderzoek naar gevolgen van de grotere rol van de ICT, belangrijker is dan een ondersteunende rol te zoeken voor de ICT.

4 Databases

4.1 Het gebruik van databases binnen geesteswetenschappen

Iedereen krijgt te maken databases, of je het nou wil of niet. Denk hier bijvoorbeeld aan de gegevens van elke student aan de RuG, die staan bij de universiteit zelf in een database, maar ook nog eens bij de Informatie Beheer Groep. Ook bij de gemeente waar je staat ingeschreven staan je gegevens in een database, niet alleen je naam, telefoonnummer, adres etc. maar ook waar je verzekert bent, of je lid bent van een kerkgenootschap, de mogelijkheden zijn legio.

Bij een gemeente als Groningen is een database bijna noodzakelijk, met zo veel en inwoners en zo veel informatie. Ditzelfde geldt voor onderzoeken binnen de geesteswetenschappen, hierbij is er meestal sprake van grote hoeveelheden informatie, en is het belangrijk dat deze informatie makkelijk toegankelijk is en goed te bewerken. Een database biedt hierin dus uitkomst. Een prima voorbeeld hiervan is het onderzoek van dr. Marten Jan Bok¹ aan de UvA. Hij deed onderzoek naar de kunsthandel in de Gouden Eeuw. Hij deed dat niet aan de

¹<http://www.uva.nl/onderzoek>

van een studie van de artiesten of ateliers zelf, maar zijn bronnen zijn veilinglijsten, bankwissels, notariële akten en handelscorrespondentie. Hij richtte zich met zijn onderzoek op een nieuw aspect van de kunstgeschiedenis, namelijk de economische kant. In de Gouden Eeuw ging het economisch zo goed met de Republiek, dat er veel vraag was naar schilderijen. Jan van Goyen schilderde op een gegeven moment iedere dag een nieuw doek. Schilderijen verzamelen werd een rage onder de sterk groeiende middenklasse, een goed schilderij kocht je dan ook al voor tien gulden. Halverwege de zeventiende eeuw onstond ook een 'tweedehands' kunstmarkt, met veilingen uit nalatenschappen. Agenten van Europese koningshuizen kochten hele scheepsladingen op.

Dit onderzoek is een uitstekend voorbeeld van hoe databases geesteswetenschappelijk onderzoek kan ondersteunen. Veilinglijsten, bankwissels, officiële akten en handelscorrespondentie bevatten grote hoeveelheden informatie. Deze informatie is pas goed te overzien als het goed opgeslagen is, bijvoorbeeld in een database. Veilinglijsten met namen van kopers en hun afkomst, kan bijvoorbeeld een beeld geven welk land het meeste kunst uit Nederland kocht. Bankwissels kunnen dan weer een beeld geven in welke mate de prijs van een bepaald schilderij is veranderd. Werken gekoppeld aan de maker kan weer een beeld geven wat de schilder gemaakt heeft. U ziet het al uit bovengenoemde voorbeelden, databases kunnen een uitkomst bieden bij geesteswetenschappelijke onderzoeken met grote hoeveelheden informatie.

4.2 De maatschappelijke consequenties

Eerder genoemd voorbeeld van de gegevens die een gemeente van je heeft, was niet overdreven. Ze kunnen ongeveer je hele levensgedrag bepalen door middel van persoonlijke gegevens in een database. Zo kan het met alles gaan. Een klassiek voorbeeld hiervan is spam. Vul een e-mailadres in op een site, en binnen een paar dagen puilt je inbox uit van ongewenste mail. Afmelden heeft geen zin, als je dit wel doet weten de verzenders van de spam dat het e-mailadres echt wordt gebruikt. Koppel dit aan de voorkeuren van reclame en interesses op uiteenlopende gebieden die je ooit hebt aangegeven bij het aanmaken van een hotmail account, en zie: Je bent het perfecte slachtoffer voor spam. Alle persoonlijke gegevens zijn opgeslagen in databases; e-mailadres, het in werking zijn van dat adres, voorkeuren etc. Uiteraard kan dit worden doorverkocht aan andere bedrijven die hier ook belang bij hebben.

Nu is spam relatief onschuldig, maar hetzelfde kan ook gebeuren met bijvoorbeeld de database van een toekomstige regering met irisscans en vingerafdrukken. Dit is uiterst gevoelige informatie, waar ongetwijfeld veel bedrijven een kapitaal voor willen betalen.

Mijn grootste zorg op dit punt is dus de privacy van de bevolking. Hier zou de visie van Bijker en Peperkamp uitstekend op zijn plaats zijn, zij willen graag dat er meer onderzoek wordt gedaan naar gevolgen. Stel dat hier een onderzoek naar komt, dan kan dat onderzoek bijdragen tot een discussie over privacy en beveiliging van persoonlijke informatie.

4.3 Wisselwerking

Het is duidelijk dat databases op maatschappelijk vlak een grote rol spelen, en dat dit verschijnsel een sterke ontwikkeling zal doormaken. Het werk van

wetenschappers als die in de afdeling Informatiekunde zal dan zeker onmisbaar zijn; het is van groot belang dat informatie makkelijk is te doorzoeken en te gebruiken. Zoals al eerder gezegd is de beveiling van informatie van groot belang, iets waar men zich in de toekomst zich dan ook zeker veel mee bezig zal gaan moeten houden. Persoonlijk vind ik dit van groter belang dan de gevolgen van het toenemende database gebruik te onderzoeken. Dit is uiteraard wel belangrijk, maar meestal zal het een ieder wel duidelijk zijn wat de gevolgen hiervan zijn. Onderzoek naar de gevolgen zal hopelijk een discussie doen oplaaien, waarin het duidelijk zal worden dat Alfa-Informatici nog een grote uitdaging zal wachten. Wetenschappers die gevolgen willen gaan onderzoeken moeten uitkijken dat ze wel bij de primaire taak van hun vakgebied blijven, namelijk een taal of literair werk beter te begrijpen. Op dit punt vind ik dus dat de visie van Bijker en Peperkamp is ondergeschikt aan die van de Alfa-Informatici.

5 Het WWW

5.1 Het gebruik van het WWW binnen geesteswetenschappen

Praktisch iedereen maakt tegenwoordig gebruik van het internet en iedereen kan dan ook beamen dat informatie via het internet ontzettend snel en makkelijk verspreid kan worden. Dit geldt uiteraard voor teksten en boeken die binnen de geesteswetenschappen worden gebruikt. Voor de opkomst van het WWW was het vaak noodzakelijk om voor een zeldzaam handschrift een al dan niet lange reis te maken. Tegenwoordig kunnen dit soort teksten in een digitale versie makkelijk worden verspreid via het WWW.

De snelheid van het WWW springt ook snel in het oog. Deze snelheid kan goed worden benut door wetenschappers. Als je vroeger een onderzoek wilde delen met iemand anders, moest je dit per post versturen of op een traditionele gegevensdrager overhandigen. Traditionele gegevensdragers, zoals een diskette of een USB-stick, hebben genoeg opslagruimte voor tekst, maar deze moeten persoonlijk overhandigt worden, wat tijd kost. Via het internet kun je het makkelijker e-mailen of online zetten.

Ook het opzoeken van een onderwerp in een encyclopedie of een woordenboek is aanzienlijk makkelijker geworden, de toegankelijkheid is sterk verbeterd door het WWW. Je kunt jezelf makkelijker toegang verschaffen tot digitale informatie, dan tot traditionele informatie in bijvoorbeeld een bibliotheek.

Zoals al gezegd is de drempel van het publiceren van onderzoeken en artikelen lager geworden. Dit heeft als gevolg dat er meer informatie beschikbaar is, gepubliceerd door mensen met verschillende meningen die voor de komst van het WWW niet de kans hadden om die mening te publiceren, dat zou in die tijd door middel van een boek of artikel moeten gebeuren.

5.2 Maatschappelijke consequenties

Doordat er zoveel informatie beschikbaar komt, is er ook de vraag: 'Wie heeft gelijk?' Zoveel informatie, en iedereen wil zijn mening door laten klinken. Dit hoeft niet altijd verkeerd te zijn, meningen zijn immers altijd goed. Maar als je op zoek bent naar feiten voor een onderzoek, paper etc. dan wil je niet altijd

meningen hebben, maar vaak ook harde feiten. Grote leugens worden dan ook zeer makkelijk verspreid via het WWW, waar tegenwoordig de meeste geruchten beginnen. Mensen kunnen ook compleet worden afgebroken op het internet, denk bijvoorbeeld aan Bush, die tijdens een toespraak zei (hoe toepasselijk): 'I hear there's rumors on the uh Internets'. De toen toch al niet bepaald populaire Bush werd in alle mogelijke forums afgekraakt.

5.3 Wisselwerking

Opnieuw zie ik hier mogelijkheden voor geesteswetenschappers om het ethische aspect wat het internet met zich mee brengt te onderzoeken. Alfa-Informatici zouden met die informatie bijvoorbeeld kunnen kijken of ze wel een goede informatiestroom kunnen maken, die voornamelijk uit harde feiten bestaat. Verder valt er natuurlijk heel veel te onderzoeken met betrekking tot invloed van het WWW op de maatschappij, maar of het daadwerkelijk ontzettend veel nut heeft durf ik niet te zeggen. Alleen als ze een aanvullende rol op de Alfa-Informatica zullen spelen, zie ik het nut van zulke onderzoeken in.

6 Grid computing

6.1 Het gebruik van grid computing binnen geesteswetenschappen

Internet is ondertussen al jaren ingeburgerd in de maatschappij en in de wetenschap. Het world wide web is vooral gericht op het delen van informatie, grid computing maakt het mogelijk om ook de apparatuur te delen². Denk hierbij bijvoorbeeld aan rekenkracht en meet-apparaten. 'Grid computing' of 'distributed computing' is een techniek gebaseerd op het koppelen van computers voor het verwerken van grote hoeveelheden data. Het principe van grid computing bestaat erin netwerken tot stand te brengen die geografisch verspreide reken capaciteit en opslag capaciteit aan elkaar koppelen. Binnen de geesteswetenschappen kun je deze rekenkracht bijvoorbeeld gebruiken om de eerder genoemde grote hoeveelheden informatie te onderzoeken. Één van de grote voordelen is dat het geld en tijd scheelt. Zelf een supercomputer bouwen kost dat namelijk, en door middel van een grid ben je een stuk goedkoper uit, de hardware die je anders moest laten bouwen, bestaat nu immers al.

6.2 De maatschappelijke consequenties

Grid computing doet bij veel mensen niet gelijk een bel rinkelen. Maar programma's als Kazaa, bittorent en eDonkey kent bijna iedereen. Dit zijn eigenlijk datagrids. Duizende gigabyte's aan data op duizenden pc's zijn via een handige interface te doorzoeken. Ontwikkelaars zeggen dat dit slechts een schijntje is vergeleken bij de mogelijkheden die grid computing ons in de toekomst gaat bieden. Verder valt er nu niet veel over de maatschappelijke consequenties te zeggen, aangezien grid computing nog sterk in de opmars is.

Als het verspreiden van muziek en films in de toekomst door grid computing zal toenemen, valt daar wel degelijk wat over te zeggen. Nu al wordt file sharing

²<http://www.gridforum.nl>

door miljoenen mensen gebruikt. Of met de woorden van Fred von Lohmann³, ontwikkelaar van het peer-to-peer programma Morpheus,: “Fifty million people voted for Bush in the presidential election,” he said. “Sixty-one million Americans use Kazaa.” Met de opkomst van grid computing kan dit aantal zelfs nog verder groeien. In mijn ogen is dit weer een prima onderzoekspunt voor mensen die de mening van Bijker en Peperkamp delen.

6.3 Wisselwerking

Het zal zeker zo zijn dat grid computing steeds invloedrijker zal zijn, de rekenkracht die voortvloeit uit een grid zal zeer geschikt zijn als ondersteunende factor bij bijvoorbeeld een tekstonderzoek. De ontwikkelaars[2] van grid computing zeggen zelf dat grid computing groter zal zijn dan het internet zoals wij dat nu kennen. De vraag is of dat ook daadwerkelijk zo zal zijn. Een visie van een wetenschapper over zijn eigen werk is immers erg subjectief. De complete gevolgen van grid computing vind ik op dit tijdstip erg onduidelijk. Hier zie ik dan een grote rol weggelegd voor onderzoek op de effecten van grid computing binnen de geesteswetenschappen. De resultaten van het onderzoek zal een grote factor zijn in het ondersteunen van de geesteswetenschappen, zoals de afdeling Informatiekunde dat wil. Maar zoals eerder gezegd kan het ‘file-sharen’ ook een grote stap voorwaards maken, iets wat ethisch minder verantwoord is. Op dit punt vind ik dus dat onderzoek vooreerst een grotere rol moet spelen dan de ontwikkeling zelf.

7 Information Retrieval

7.1 Het gebruik van IR binnen geesteswetenschappen

In deze paper heeft informatie een grote rol. Dit kan ook niet anders aangezien we in een ‘informatietijdperk’ leven. Overal waar we komen krijgen we grote hoeveelheden informatie te verwerken. Met de computer is het niet anders, in wetenschappelijk onderzoek is informatie verwerken de grootste rol van de computer.

Vóór de komst van snelle netwerken bestonden informatiebronnen veelal uit databases van teksten en specifiek hiervoor ontwikkelde zoekmethodes. In veel gevallen werd het vinden van informatie gedaan door informatici, aangezien zij de syntax van de methode begrepen. Voor onderzoekers zonder kennis van de zoekmethode was het moeilijk zoeken. Grofweg gezegd kun je de onderzoekers die zich bezig houden met IR opdelen in twee groepen: Onderzoekers die van mening zijn dat de beste zoekresultaten te verkrijgen zijn als de zoekmachine natuurlijke taal beheerst, en de andere groep gaat uit van statistische analyses van woordgebruik. Tot nu toe lijkt het erop dat de tweede groep de beste zoekresultaten krijgt. Een stuk software of een computer is praktisch niet in staat om een natuurlijke taal onder de knie te krijgen. In de praktijk komt het er dan ook op neer dat de zoekmachines resultaten aan de gebruiker geven, na opgave van kernwoorden, en niet door complexe vragen van gebruiker.

³<http://www.collegian.psu.edu/archive/2003/04/04-15-03tdc/04-15-03dnews-10.asp>

Voor de komst van het Internet stond informatie op één gegevensdrager, dit vergemakkelijkte het zoeken vergeleken bij het zoeken op het Internet. Daar staan alle documenten over de hele wereld, niet gecatalogiseerd en soms zelfs dubbel. Dit wordt dan nog eens moeilijker door de niet altijd betrouwbare computers waar de informatie op staat. Een zoekmachine is dus van groot belang om alle informatie te doorzoeken. Een zoekmachine maakt gebruik van 'spiders', deze spiders maken gebruik van links die op webpagina's staan. Deze links volgen ze, en brengen het resultaat terug naar een 'index-database', beter bekend als zoekmachine. De taak van een zoekmachine is zeer complex, denk bijvoorbeeld aan de hoeveelheid talen op het web, zoekopdrachten die niet specifiek genoeg zijn en dubbele betekenissen van woorden.

Ik denk dat de grote droom van veel Alfa-Informatici is dat mensen met computers kunnen communiceren door middel van natuurlijk taalgebruik. Maar ik denk dat we dan ook al heel wat jaren verder zijn, als dat zou gebeuren.

7.2 De maatschappelijke consequenties

Voor iemand zonder kennis van zoeksystemen is het best lastig zoeken naar een document op het web. Veelal moet je toch enige kennis hebben, om goede zoekopdrachten te geven, en zo het gewenste resultaat te krijgen. Dit kan ook voor een wetenschapper gelden, die informatie op het web zoekt, die zoveel zoekresultaten krijgt die niet relevant zijn, of juist allemaal op het eerste gezicht, dat hij door de bomen het bos niet meer ziet. Het is dus de taak aan de Alfa-Informatici om het zoeken naar informatie gebruikersvriendelijk te houden. Het beste zou uiteraard zijn om een zoekmachine te creëren die door middel van natuurlijk taal een zoekopdracht verwerkt. Dit zal helaas de komende jaren nog geen optie zijn denk ik. Zoals al eerder aangegeven werkte het zoeken door middel van steekwoorden beter dan zoeken met natuurlijk taal.

7.3 Wisselwerking

De wisselwerking tussen deze twee verschillende groepen vind ik op dit punt erg klein. Er zou onderzoek gedaan kunnen worden naar hoe de mensen zoekmachines gebruiken, en wat er verbeterd zou kunnen worden. Maar dit is iets wat de Alfa-Informatici zelf ook doen. Op dit punt vind ik dus dat de visie van Bijker en Peperkamp absoluut niet op zou kunnen gaan.

8 Conclusie

In deze paper heb ik geprobeerd een eigen mening te vormen over de ICT binnen de geesteswetenschappen. Wat mij vooral opviel is dat de ethische kant van het verhaal een stuk zwaarder weegt dan ik had verwacht. Vooral bij databases en grid computing vond ik dat het geval. In mijn ogen is het dan ook het belangrijkste dat de Alfa-Informatici doorgaan met hun ondersteunende rol binnen de geesteswetenschappen, maar dan moeten ze wel de eventuele gevolgen van hun werk goed in hun achterhoofd houden. Dit zal misschien behoorlijk pessimistisch klinken, maar dat is in het geheel niet zo. Integendeel, nieuwe ontwikkelingen die geesteswetenschappelijk onderzoek makkelijker maken juich ik persoonlijk toe! Maar helaas zijn er altijd mensen die misbruik maken van werk, wat niet voor die 'slechte' doeleinden was gebruikt.

Onderzoek zoals dat gedaan zou moeten worden volgens Bijker en Peperkamp kan uiteraard erg nuttig zijn, bijvoorbeeld ook om gevolgen van nieuwe ontwikkelingen te voorspellen. Dit kan bijvoorbeeld worden gedaan door de gevolgen van een soortgelijk project te onderzoeken. Maar om puur alleen gevolgen op de maatschappij te onderzoeken, als geesteswetenschapper, vind ik persoonlijk minder nuttig. Maar als ze kunnen helpen met de ethische kant, zal ik dat uiteraard toejuichen.

Om het nog eens samen te vatten: het werk van Alfa-Informatici is erg nuttig, maar men moet goed in het achterhoofd houden dat er mensen zijn die hun werk uit eigenbelang gaan gebruiken, hier dienen Alfa-Informatici rekening mee houden. Als onderzoek op de gevolgen van de ICT binnen de geesteswetenschappen kan helpen met dit probleem op te lossen, vind ik dit een erg goede samenwerking tussen deze twee groepen, met een toch aanzienlijk verschil in mening.

Referenties

- [1] Wybe Bijker en Ben Peperkamp. *Geengageerde Geesteswetenschappen*. AWT, 2002.
- [2] Ian Foster. *The Grid: Blueprint for a New Computing Infrastructure*. Morgan Kaufmann Publishers Inc, 1999.