

STEPHAN BESSER & THOMAS VAESSENS

Digital Humanities: the Next Big Thing?

Enkele notities bij een ontluikend debat*

Abstract – In the form of provisional notes, the authors offer suggestions for an intensification of the theoretical debate on the digital humanities and computational literary studies in particular. From the perspective of poststructuralist theory, they address some of the epistemological underpinnings of current projects in computational literary studies. Against the backdrop of international debates in the digital humanities, they focus on the Dutch situation and argue for a further integration of various forms of (digital and non-digital) literary studies.

Zoals veel literatuurwetenschappers zijn ook wij van mening dat een doordachte integratie van digitale werkwijzen in het onderzoek van groot belang kan zijn, onder meer voor de noodzakelijke kwantificering daarvan en voor de empirische onderbouwing van geesteswetenschappelijke bevindingen en intuïties. De beschikbaarheid van grote elektronische tekst- of databestanden en van nieuwe technieken waarmee ze kunnen worden onderzocht, zal ertoe leiden dat ook in de literatuurwetenschap en de Nederlandse letterkunde oude vragen op nieuwe manieren beantwoord kunnen worden en dat nieuwe vragen kunnen worden gesteld. Tal van collega's zijn met gedrevenheid en enthousiasme bezig nieuwe mogelijkheden van de informatie- en communicatietechnologie te verkennen en te benutten voor hun onderzoek, zowel in Nederland als daarbuiten. Franco Moretti's concept van 'distant reading' was ongetwijfeld een van de belangrijkste inspiraties voor het literatuurwetenschappelijke debat in de afgelopen 15 jaar en onderzoeksprojecten zoals het *Stanford Literary Lab* in de vs en *The Riddle of Literary Quality* in Nederland hebben de discussies over nut, beloftes en implicaties van de computationele literatuurwetenschap aangewakkerd.

Wij willen met deze notities bijdragen aan de ontluikende discussie en pleiten voor een intensivering van het theoretisch debat over de *Digital Humanities* (DH). Daarmee sluiten we aan bij de oproep tot kritische reflectie op DH in twee recente Amsterdamse oraties, *Het einde van de Geesteswetenschappen 1.0* van Rens Bod en *De stijl van R* van Karina van Dalen-Oskam, en ook bij een belangrijk artikel uit 2011 van mediawetenschapper David M. Berry over 'The Computational Turn'.¹ Hoe veelbelovend de nieuwe technieken ook zijn, we moeten inderdaad ook kritisch reflecteren, niet alleen op de voorlopige resultaten, maar ook op de vooronderstellingen ervan. We doen dit vanuit de positie van literatuurwetenschappers die hechten aan centrale inzichten van de poststructuralistische theorievorming en daarom sceptisch aankijken tegen bepaalde suggesties van es-

* Deze bijdrage is gedeeltelijk gebaseerd op een lezing bij gelegenheid van het symposium 'Handgemaakte inzichten: digitale technieken in de letterkunde' (UvA, 1-2-2013) en op bijdragen aan het openbare debat 'Digital Humanities: the Next Big Thing?' (UvA, 1-6-2013).

1 Bod 2012, Van Dalen-Oskam 2013, Berry 2011.

sentialisme, *fantasies of completeness*, formalistische tendensen en eenvoudige ontologieën van ‘literaire kwaliteit’ in de DH. Anders dan Bod en Van Dalen-Oskam zijn wij echter buitenstaanders: als literatuurwetenschappers die vooralsnog met traditionele middelen en corpora werken, behoren wij niet tot de avant-garde van pleitbezorgers – en kenners – van nieuwe digitale technieken. We zijn ons ervan bewust dat we in de huidige situatie, waarin een nieuw instrumentarium ontwikkeld wordt, de ‘vragende partij’ zijn. In het debat, waartoe de hieronder bijeengebrachte observaties, discussiepunten, vragen en aantekeningen moeten bijdragen, hopen en verwachten wij dan ook onze vooringenomenheden te zullen moeten bijstellen.

1 Sötemanns zoekmachines

In 1997, het jaar van de lancering van Google, verscheen in het tijdschrift *Nederlandse Letterkunde* een bijdrage over ‘Internet en de letterkundige neerlandistiek’. Daarin werd een aantal voorbeelden genoemd van nieuwe onderzoeksmogelijkheden die vooral met de studie naar beïnvloeding en allusies te maken hadden. Zo merkten de auteurs op dat iedere student zich, met behulp van de ook toen al beschikbare digitale bibliotheek van Duitse, Franse en Engelse dichters uit de negentiende eeuw, binnen een paar uur een beeld kon vormen van de sporen die de historische romantiek had nagelaten in het dichtwerk van Willem Kloos. Dit beeld, opperden de auteurs, zou beter zijn – want op meer feitelijke tekstovereenkomsten gebaseerd – dan dat van zelfs de meest belezen onderzoeker zónder digitale hulpmiddelen.²

Er zat uiteraard wat provocatie in die laatste opmerking, en deze bleef niet onopgemerkt. Enige tijd na publicatie van het artikel had een van de auteurs ervan de eer tijdens een receptie aangesproken te worden door A.L. Sötemann. Sötemann had zich de inhoud van het artikel aangetrokken, en was het absoluut niet eens met de pesterige suggestie die erin vervat lag, namelijk dat het literatuurwetenschappelijk handwerk, het verzamelen en verwerken van onderzoeksmateriaal, voortaan wel door machines kon worden overgenomen.

Sötemann had natuurlijk gelijk: literatuurwetenschap is in elk geval óók mensenwerk. Toch had de aarzeling bij juist deze wetenschapper ook iets paradoxaals. Digitale technieken houden namelijk juist voor het type onderzoek waarin Sötemann excelleerde, enorme beloften in. Sötemann was een van de Nederlandse vertegenwoordigers van die eerbiedwaardige geesteswetenschappelijke traditie waarin acribie op de vierkante centimeter van de tekst wordt gekoppeld aan een duizelingwekkende eruditie. Die eruditie stelde zulke geleerden in staat om toch, ondanks hun soms ten onrechte als bijziendheid veroordeelde tekstgerichtheid, brede historische verbanden te zien. Juist Sötemanns onderzoek naar de verspreiding van bepaalde poëtische metaforen in de Westerse literatuur of zijn zoektocht naar de bronnen van dichters als Boutens en Leopold zou bij digitale technieken en tekstcorpora gebaat zijn geweest.

Hetzelfde kun je beweren voor de op enorme eruditie gestoelde studies van

2 Mourits & Vaessens 1997: 78.

comparatisten als Leo Spitzer en Erich Auerbach en hun op bibliotheken, kaartenbakken, papier en (vooral) een enorm geheugen gebaseerde onderzoekstechnologie. Het is dan ook geen toeval dat het ‘probleem’³ van de ‘wereldliteratuur’ – de duizelingwekkende hoeveelheid literaire teksten uit verschillende periodes en culturen – via het werk van Franco Moretti en anderen recentelijk belangrijke impulsen voor de ontwikkeling van de computationele literatuurwetenschap heeft gegeven. Maar ook van Sötemanns dissertatie over *De structuur van Max Havelaar*, waarin hij op zoek was naar objectieve en empirische, aan de bestudering van tekstcompositie ontleende fundering voor uitspraken over literaire kwaliteit, zou je je een ‘gedigitaliseerde’ versie 2.0 kunnen voorstellen – of een onderzoeksproject met het doel om Sötemanns stellingen nu eens met de middelen van de DH te onderbouwen of te weerleggen.

De aarzeling van Sötemann was dus misschien wel een misverstand. Maar wellicht is het juist daarom *ook* een misverstand om te denken dat de DH en de computationele literatuurwetenschap zelf per definitie vooruitstrevend, grensverleggend en innovatief zijn. Wat betekent het dat de vooronderstellingen van het onderzoek dat Sötemann en anderen in de jaren 60, 70 en 80 deden, zo goed aansluiten bij bepaalde claims van DH vandaag?⁴

2 Vooruit naar het verleden?

Dezelfde eeuwenoude vraag die Sötemann bezig hield – de vraag naar literaire kwaliteit – staat ook centraal in het genoemde DH-project *The Riddle of Literary Quality*, en net als Sötemann destijds belooft dit project deze vraag eindelijk eens echt wetenschappelijk aan te pakken. Destijds zocht Sötemann die wetenschappelijkheid in een methode die precieze en controleerbare uitspraken over structuur opleverde; nu voegt de techniek daaraan de mogelijkheid toe niet slechts één tekst, maar een enorm elektronisch corpus van moderne romans op meetbare elementen te analyseren. Het project is gericht op de vragen: ‘Wat maakt dat een boek door een groot publiek goed of juist slecht wordt bevonden? Waarom noemen mensen sommige boeken “literair” en andere boeken niet?’ Het vraagstuk van de ‘literaire kwaliteit’ wordt hier dus nadrukkelijk aan de receptie gekoppeld en krijgt op deze manier een sociologische dimensie. Maar hoe deze dimensie in de onderzoeksopzet – gebaseerd op een *online*-enquête van lezers – verwerkt moet worden lijkt nog een open vraag. De bemiddelende instanties en ideologieën, die smaakoordelen en *verschillende* definities van ‘literatuur’ beïnvloeden, spelen in de onderzoeksopzet geen rol. Door de kortsluiting van smaakoordelen en tekstenmerken (zoals lengte van zinnen, omvang van de woordenschat etc.) lijkt ‘literaire kwaliteit’ uiteindelijk wel tot een eigenschap van *teksten* gereduceerd te worden; haar culturele inbedding blijft op deze manier immers buiten beeld.

Je zou je, in het verlengde hiervan, kunnen afvragen waar de hernieuwde belangstelling vandaan komt voor de al dan niet gemedieerde toekenning van ‘lite-

3 Moretti 2013: 46.

4 Ook Berry (2011: 4) roept op tot het bevragen van de ‘assumptions implicit in such research, e.g. close reading, canon formation, periodization, liberal humanism, etc.’

5 Cf. de website van het *Riddle*-project: <http://literaryquality.huygens.knaw.nl> (9-13-2013).

raire kwaliteit', voor processen van canonvorming⁶ en de ontwikkeling van genres. Waarom opnieuw zoveel aandacht voor auteursherkenning, de persoonlijke stijl van schrijvers, directe vormen van intertekstualiteit (zoals de 'paper trail van Vico naar Hegel' waarover Rens Bod sprak in een interview),⁷ en literaire 'invloed'?⁸ Matthew Jockers, een van de vooraanstaande computationele literatuurwetenschappers in de vs, komt in het hoofdstuk 'Influence' van zijn studie *Macroanalysis* tot de empirisch onderbouwde conclusie dat Melvilles beroemde roman *Moby Dick* in een lange traditie van 'whaling- and seafaring-themed fiction' stond. Volgens Jockers vond deze traditie uiteindelijk haar 'full expression' (Jockers 2013: 154) in Melvilles roman. Waarop deze positieve kwalificatie – 'full expression' – precies gebaseerd is blijft onduidelijk; de digitale methodiek lijkt door deze bewering wel direct compatibel met een conventionele notie als die van het literaire 'meesterwerk'. Dit is zeker geen uitzondering: Jockers koppelt esthetische waardeoordelen herhaaldelijk aan stilistische tekstkenmerken – 'in the quarry of great literature, it is style, or "technique," that ultimately separates the ore from the tailings' (Jockers 2013: 63) – en beschrijft de 'creativiteit' van schrijvers als 'kern' van hun literaire identiteit. Deze kern moet volgens Jockers niet met externe invloeden worden verward: 'A consistent problem for authorship researchers [...] is the possibility that [...] external factors (for example, linguistic register, genre, nationality, gender, ethnicity, and so on) may influence or even overpower the latent authorial signal' (Jockers 2013: 63; cf. 156). Gender en etniciteit dus als 'externe' invloeden op een ergens 'binnen' de schrijver gesitueerde literaire stem: als dit geen essentialistische definitie van literatuur – en van subjectiviteit! – is, zal het moeilijk worden om er überhaupt een aan te wijzen.

Passages als deze dragen bij aan onze indruk dat de DH traditionele literatuurwetenschappelijke concepten – auteur, literariteit, invloed, maar ook de toekenning van literaire waarde – opnieuw agenderen zonder op de theoretische implicaties en tekortkomingen van deze concepten te reflecteren. Gechargeerd geformuleerd: de theorievorming blijft achter bij de digitale middelen. Thomas Rommel heeft deze indruk in de *Companion to Digital Humanities* als volgt samengevat: 'By focusing primarily on empirical textual data, computer studies of literature tend to treat text in a way that some literary critics see as a reapplication of dated theoretical models' (Rommel 2004).

Misschien heeft deze ontwikkeling ook gewoon te maken met de niet te onderschatten technische beperkingen van de huidige DH: categorieën als auteur, stijl en genre kunnen nuttige handvatten bieden in een nieuw en technisch complex onderzoeksgebied. Met mediawetenschapper Gary Hall denken we echter dat het debat over DH er ook nu al baat bij heeft wanneer het, naast de *praktijk* van DH, ook de kritische theoretische zelfreflectie op de soms wat ouderwets aandoende, *commonsensical* vooronderstellingen ervan centraal stelt. Hall constateert dat er in het debat in de vs tegenwoordig sprake is van een 'shift from a concern with theory and ideology to a concern with methodology' en dat het 'vroege stadium' van de ontwikkeling van de DH in dit verband als legitimatie voor een post-theoreti-

6 Cf. de projectbeschrijving 'Between Canon and Corpus. The 20th Century Novel in English' op de website van het *Stanford Literary Lab* (<http://litlab.stanford.edu/>).

7 Nieuwsbrief FGW, Universiteit van Amsterdam, nr.142, oktober 2011: 1.

8 Cf. het hoofdstuk 'Influence' in Jockers 2013.

sche houding wordt aangehaald: nu even aandacht voor software en technologie, *critical theory* hadden we in de afgelopen decennia al genoeg ...⁹ Ons lijkt deze bewuste of onbewuste onverschilligheid ten opzichte van de cultuur- en literatuurtheoretische discussies van de afgelopen veertig jaar geen stap vooruit te zijn. De computationele literatuurwetenschap heeft naar onze inschatting juist veel te winnen als ze aansluiting zou zoeken bij de verschillende vormen van *problematizing* van alledaagse aannames over literatuur – volgens Antoine Compagnon een van de belangrijkste projecten van de literatuurwetenschap in de afgelopen decennia.¹⁰ Als deze (wederzijdse) toenadering achterwege blijft, lopen we het risico dat DH het volgende kortstondige literatuurwetenschappelijke modeverschijnsel wordt.

3 Digitale discoursanalyse

Een ander punt is de – naar ons idee te beperkte – opvatting van contextualisering in de computationele literatuurwetenschap. De bijna uitsluitend *littéraire* onderzoekscorpora van projecten als *The Riddle of Literary Quality* of de *Stanford Literary Lab* doen denken aan de verzuchting van Jonathan Culler in het debat over wereldliteratuur: ‘Comparative Literature, at Last!’ (Culler 2006: 237) Cullers reflex richtte zich tegen de tendens in de literatuurwetenschap van de afgelopen decennia om het literatuurwetenschappelijk instrumentarium in te zetten in de bestudering van een steeds breder en heterogener empirisch domein: niet meer alleen *littéraire* teksten, maar ook ‘teksten’ van allerlei andere aard. De verzuchting zou dus nu kunnen zijn: ook al brengt DH met zich mee dat we als literatuurwetenschappers nu enorm veel teksten moeten bestuderen, we kunnen het in ieder geval weer bij *littérature* houden ...

Bestaande projecten op het gebied van de computationele literatuurwetenschap concentreren zich inderdaad vaak op een tamelijk *disciplinaire* manier op de ‘littéraire’ tekst, alsof er een duidelijk omljnd of vanzelfsprekend corpus zou zijn van zulke teksten. Zo’n nieuwe beperking tot literatuur ‘op zich’ – ook al worden ‘contextuele’ factoren ter verklaring van observaties op de gebieden van genre, literaire invloed en stijl aangehaald – zou ons inziens voor de literatuurwetenschap geen goede ontwikkeling zijn. In de afgelopen decennia heeft juist de belangstelling voor literatuur in relatie tot *andere* vertogen, media, kunsten en kennisvormen tot boeiende en vernieuwende onderzoeksprojecten en vraagstellingen geleid. Ook al worden de databases van literaire teksten nog zo groot, zonder de technische mogelijkheden en de bereidheid om ze aan niet-literaire datasets te koppelen, dreigt de computationele literatuurwetenschap een enigszins zelfgenoegzame oefening te worden. Gechargeerd gesteld: Een petabyte literatuur kan

9 Hall vat deze gedachtegang als volgt samen: ‘[C]ritical and self-reflexive theoretical questions about the use of digital tools and data-led methodologies should be deferred for the time being, lest they have the effect of strangling at birth what could turn out to be a very different form of humanities research before it has had a chance to properly develop and shape’ (Hall 2012). Als voorbeeld van deze houding haalt Hall onder anderen uitspraken van Tobias Scheinfeldt, directeur van het *Center for History and New Media* aan de George Mason University (Virginia), aan.

10 Compagnon 2004: 193.

cultureel minder significant zijn dan één biologische verhandeling waarin een literaire tekst geciteerd wordt.

Juist op het gebied van de digitale discoursanalyse liggen enorme kansen voor DH. Een klein voorbeeld hiervan geeft Rodney Jones met een exemplarische analyse van liedteksten van popster Lady Gaga. Met het programma AntConc vergeleek Jones de teksten van 59 liedjes van Lady Gaga met een *reference corpus* van honderd recente chart-hits. Hij constateert dat de term ‘love’ in de teksten van Lady Gaga veel vaker in combinatie met het persoonlijk voornaamwoord ‘me’ in de accusatief verschijnt dan in de andere songs. Voor Jones is dat een aanwijzing dat er sprake is van een specifieke ‘discursieve constructie van liefde’ in de teksten van Lady Gaga waarbij het lyrisch subject vooral als lijdend voorwerp van bewondering en verlangen figureert.¹¹ Op zich is dat misschien geen opzienbarend onderzoeksresultaat, maar het geeft wel aan dat verschillende vormen van discoursanalyse flink kunnen profiteren van een uitbreiding van hun onderzoeksrepertoires en methodische middelen met digitale technologieën.

Een ander, nog niet gerealiseerd, project zou een onderzoek kunnen zijn naar het ontstaan en de culturele verspreiding van de discursieve praktijk van *coming out/uit-de-kast* komen in Nederland. Als een toereikende hoeveelheid bronnen uit verschillende vertogen gedigitaliseerd zou zijn – denk bijvoorbeeld aan het tijdschrift *Sextant*, psychologische studies en romans over homoseksuele identiteiten – zou het mogelijk moeten zijn om uit te zoeken sinds wanneer de uitdrukking ‘coming out’ in Nederland gebruikelijk is, op welke manieren ze tussen verschillende discoursen is gaan circuleren en welke invloed dat op de identiteits-effecten van het concept had. Zo’n onderzoek naar de verspreidingspatronen van bepaalde begrippen en concepten zou trouwens nauw aansluiten bij de positivistische oriëntatie van Michel Foucaults *Archeologie van het weten* (1969). Ook voor Foucault ging het immers in eerste instantie niet om de ‘achterliggende’ betekenissen van vertogen, maar om hun regelmatigheden, genealogische interacties en concrete manifestaties in de vorm van linguïstisch en ander bewijsmateriaal.

Een belangrijke vraag die wij voor de DH en de computationele literatuurwetenschap hebben, is daarom ook wat zij kunnen bijdragen aan onderzoek dat vertogen, discursieve verschuivingen en interdiscursieve circulatie van termen in kaart brengt. In Nederland lijkt het project *E-Humanity Approaches to Reference Cultures: The Emergence of the United States in Public Discourse in the Netherlands, 1890-1990* de meest ambitieuze en grootschalige onderneming in deze richting te zijn. Het feit dat dit project zich vooral concentreert op krantenberichten als bronnenmateriaal – wat ongetwijfeld al genoeg technische en conceptuele problemen oplevert – wijst er echter ook op, dat de mogelijkheden voor digitale discoursanalyses tot nu toe beperkt zijn. Niettemin zien we grote kansen voor een toekomstige samenwerking op dit gebied. Deze zal ongetwijfeld duidelijk maken dat ‘kwalitatief’ en ‘kwantitatief’ onderzoek ook bij uitgesproken positivistisch en empirisch georiënteerde vormen van discoursanalyse niet te scheiden zijn. Discursieve landschappen zijn nooit ‘vlak’ maar krijgen reliëf in de complexe wisselwerking van verschillende en vaak incommensurabele vertogen. Door die omstandigheid worden wederom lastige methodische en theoretische vragen naar de

11 Jones 2012: 127-131.

afbakening van discoursen en de ‘weging’ van uitspraken in verschillende contexten opgeworpen: een medische term in een medische context lijkt bijvoorbeeld cultureel minder significant dan een medische term in een politieke context. Maar kun je deze verschillen überhaupt kwantitatief vatten?

4 Tellen, maar *wat*?

Een van de grote uitdagingen – technisch én conceptueel – van de computationele literatuurwetenschap is het definiëren en afbakenen van *units of observation*, dus de kleinste eenheden van informatie (ofwel tekstkenmerken) in de datasets. Zoals David Berry opmerkt: ‘The key point is that without the possibility of discrete encoding, there is no object for the computational device to process’ (Berry 2011: 2). In het geval van stilistisch en semantisch onderzoek zijn deze objecten meestal woorden, en die zijn relatief makkelijk te isoleren en met programma’s op verschillende kenmerken te onderzoeken (bijv. frequenties en combinaties van woorden in een corpus). Een inspirerend voorbeeld van zo’n onderzoek laat Franco Moretti zien in zijn essay ‘Style, Inc.’ (2009) over de titels van zeventuizend Britse romans in de periode 1740–1850. Van een belangrijke beperking van zijn onderzoek – het gebrek aan gedigitaliseerde versies van veel van de 7000 teksten – weet Moretti op creatieve en ingenieuze manier een voordeel te maken: een focus op alleen de titels van de romans vanuit het perspectief van wat hij ‘quantitative formalism’ (Moretti 2013: 180) noemt. Het lukt hem om bepaalde veranderingen in zijn corpus – zoals het afnemen van de gemiddelde lengte van de titels – door verschuivingen in het culturele en mediale ‘ecosysteem’¹² te verklaren (het ontstaan van tijdschriften, formele eisen van catalogi van bibliotheken, de rol van roman-titels in de algemene economie van de aandacht etc.). Voor veel van de stilistische kenmerken van zijn corpus – de toename van abstracties (*Sense and sensibility*), de ongelijke spreiding van bepaalde en onbepaalde lidwoorden in de onderzochte periode etc. – en hun semantische en ideologische implicaties komt Moretti met originele interpretaties. Anders dan andere onderzoekers op dit gebied verliest hij daarbij de ‘So what?’-vraag¹³ – waarom moeten we als lezers al deze statistieken bekijken en interessant vinden? – nooit uit het oog.

Veel minder overtuigend is daarentegen Moretti’s poging om ook de narratologische dimensie van literaire teksten aan een kwantitatieve en computationele analyse te onderwerpen. Hij doet dit exemplarisch in ‘Network Theory, Plot Analysis’ (2011) met Shakespeare’s *Hamlet*. Als antwoord op de vraag hoe je een plot kwantitatief kunt vatten, presenteert Moretti een visualisatie van de tekst als netwerk waarin de handeling van *Hamlet* gereduceerd is tot twee soorten elementen: *nodes* (de personages) en *edges* (woordenwisselingen/dialogoog tussen de personages, zonder kwantitatieve meting). Deze procedure levert interessante observaties op, bijvoorbeeld over de structurele relevantie van personages voor de plot en de correlaties tussen de dichtheid van het netwerk (met zijn centrum rond Hamlet en Claudius) en stilistische tekstkenmerken (meer figuratief taalgebruik in deze

12 Moretti 2013: 186.

13 Moretti 2013: 202.

zones). Deze benadering lijkt ook interessante mogelijkheden te bieden voor heel ander onderzoek zoals dat van Eve Sedgwick in haar studie *Between Men* naar homosocialiteit, gender-asymmetriën en *love triangles* in de Engelse literatuur van de zeventiende tot negentiende eeuw. Een netwerkanalyse à la Moretti had haar conclusies misschien wel kunnen onderbouwen (of weerleggen).

Tegelijk maakt Moretti's artikel echter ook duidelijk hoe moeilijk het is om narratieve tekstkenmerken te formaliseren en verwerkbaar te maken. De enige actoren (*nodes*) waar hij rekening mee houdt, zijn personages. Maar hoe zit het met *nodes* die 'groter' of 'kleiner' zijn dan mensen, met niet-menselijke actoren en dingen, entiteiten zonder naam etc.? Volgens Moretti's analyseschema zou de 'stille kracht' geen actor in de gelijknamige roman van Couperus zijn en het kolonialisme geen factor in de plot. En hoe zit het met vormen van interactie (in netwerk-jargon: *edges*) die geen *paper trail* in de vorm van woordenwisselingen achterlaten? Kunnen die überhaupt statistisch worden vastgelegd zonder dat een menselijke lezer erbij te pas komt?¹⁴

Dit maakt ook duidelijk hoe groot de rol van vooronderstellingen over subjectiviteit en *agency* bij narratologische analyses al bij de *samenstelling* van de datasets is. Dit probleem zou nog exponentieel groeien als je niet één tekst onderzoekt (zoals Moretti) maar duizenden, want dan zouden kleine veranderingen in de parameters en definities enorme gevolgen voor het eindresultaat kunnen hebben.

Ook een ander aspect van de narratologische dimensie van teksten betekent een grote uitdaging voor de computationele literatuurwetenschap, namelijk het vertelperspectief (en, daarmee verbonden, focalisatie). Fenomenen zoals de niet duidelijk aan één spreker toe te schrijven vrije indirecte rede in *Madame Bovary* en de complexe meerstemmigheid van veel uitspraken (ook in niet-literaire teksten) roepen de stelling van Michail Bakhtin in herinnering dat linguïstische en stilistische benaderingen per definitie niet in staat zouden zijn om de dialogische dimensie van taal adequaat te beschrijven.¹⁵ We vragen ons af of de computationele literatuurwetenschap op termijn wél instrumenten kan ontwikkelen die ook niet-menselijke lezers in staat stellen om de dialogische gelaagdheid van uitspraken in haar verschillende vormen – in Bakhtins terminologie: *stilisering*, *double-voiced discourse*, *internal polemic*, *hidden dialogue* etc. – te herkennen en statistisch verwerkbaar te maken. Dat zou een belangrijke innovatie voor alle vormen van literatuurwetenschappelijk onderzoek betekenen die zich voor de *interne* splitsing en gelaagdheid van woorden en het fundamentele dialogisme van taal interesseren.

5 'There is no raw data!'

DH kunnen (en moeten) bijdragen aan de onttovering van de producten van de menselijke 'geest', maar data en digitale technieken hebben ook hun eigen magie en toverkracht. Het is opvallend dat een theoretisch enorm goed onderlegd onderzoeker als Moretti af en toe de kritische en reflecterende afstand tot de door

¹⁴ In de inmiddels enorm gevarieerde subdiscipline van de computationele narratologie lijken annotaties door menselijke lezers nog steeds een belangrijke rol te spelen, cf. bijv. Mani 2010.

¹⁵ Bakhtin 1981: 181-185.

hem zelf samengestelde datasets verliest. In de programmatische laatste alinea van zijn artikel ‘Network Theory, Plot Analysis’ – tevens de slotzinnen van zijn bundel *Distant Reading* – roept Moretti zijn literatuurwetenschappelijke vakgenoten in de retorische vorm van een vraag op om in de toekomst data als ‘grondstof’ van het onderzoek met elkaar te delen: ‘Will we, as a discipline, be capable of sharing raw materials, evidence – *facts* – with each other?’ (Moretti 2013: 240) Hij lijkt hier bijna te vergeten dat data *nooit* onbewerkte (‘rauwe’) gegevens zijn maar altijd op de een of andere manier geproduceerd en dus ‘voorgekookt’. “*Raw Data*” is an *Oxymoron*’ luidt dan ook de treffende titel van een recente verzameling van artikelen over DH van de mediawetenschapper Lisa Gitelman.

Het blijft opvallend dat het kennelijk nodig is om dit simpele inzicht – data zijn culturele en geen natuurlijke resources – zo nadrukkelijk in herinnering te roepen. Het lijkt dan ook niet overbodig dat literatuurwetenschapper James English, zeker geen tegenstander van DH, Moretti en anderen op hun ouderwets en bijna onschuldig geloof in data wijst: ‘Naïve ontologies of the dataset – equating quantification with objectivity, data with facts, and bigger data with better access to reality – have been thoroughly critiqued by historians and sociologists of science and, more lately, by digital humanists. [...] From the standpoint of such critiques, every kind of literary data is, by definition, already cooked according to some recipe of literariness, which in turn derived from some history of reading and, yes of interpretation. However annoyingly half-baked our concepts, there really is no escape from the kitchen’ (English 2012). Het is een oud, maar wel belangrijk, deconstructivistisch inzicht: net zoals teksten kennen data-sets een eigen *drift*, een constante verschuiving van sporen en inscripties die elke ultieme vastlegging van betekenis onmogelijk maakt. ‘Feiten’ zijn altijd gerelateerd aan de kaders en afbakeningen waarmee ze – tijdelijk – worden voortgebracht.

We denken daarom dat de reflectie op de betekenis-effecten van visualiserings-technieken en de ornamentale en symbolische functie van grafieken in wetenschappelijke teksten – ‘kijk, dit is vooruitstrevend onderzoek!’ – veel nadrukkelijker dan dit nu gebeurt onderdelen zouden moeten zijn van DH in het algemeen en van de computationele literatuurwetenschap in het bijzonder. Hierbij valt ongetwijfeld veel te leren van de *iconic turn* in de geesteswetenschappen, maar ook van het onderzoek naar de poëtica’s van het weten in de *cultural studies of science*, de studies van *database aesthetics* (Lev Manovich) en het onderzoek naar de vorming van culturele en discursieve werkelijkheden – ‘werelden’ – door algoritmes. Als data zelf een cultureel product zijn, kunnen ze juist in DH niet gewoon als gegeven worden verondersteld.

6 *Wahlverwandschaften*

Het feit dat DH in Nederland, naar onze indruk, deze uitdagingen nog niet echt zijn aangegaan, heeft misschien ook te maken met de specifieke socio-economische en politieke context waarin ze hier terecht zijn gekomen. We hadden al geconstateerd dat projecten zoals *The Riddle of Literary Quality* maar ook *Jockers Macroanalysis* goed aansluiten bij – vanuit ons perspectief – relatief traditionele concepten als ‘canon’ en ‘literariteit’ en het *commonsensical* begrip ervan in het

huidige publieke debat over literatuur in de literatuurkritiek en culturele journalistiek (het adjectief ‘literair’ als kwaliteitslabel). Er is niet al te veel ideologiekritisch wantrouwen nodig om te vermoeden dat de onmiskenbare *appeal* van het fenomeen DH in de samenleving vandaag voor een deel te maken heeft met een humanistisch gekleurde onvrede van het bredere publiek met een verondersteld gebrek aan respect van hedendaagse geesteswetenschappers voor canonieke kunstwerken en teksten.

De huidige belangstelling voor ‘culturele identiteit’, ‘erfgoed’ en canons in alle vormen en maten biedt de DH in Nederland veel aanknopingspunten en strategische opties om hun maatschappelijk belang te onderstrepen en onderzoeksfinanciering veilig te stellen. Een andere belangrijke contextuele factor is de tegenwoordige nadruk op kennisvalorisatie, zichtbaar bijvoorbeeld bij de invoering van de zogenaamde ‘kennisbenuttingsparagraaf’ als verplicht onderdeel van alle bij NWO ingediende subsidieaanvragen. Expliciete aandacht voor kennisvalorisatie is zeker geen slechte zaak, in tegendeel, maar een *beperkte definitie* ervan wel. Om een voorbeeld te geven: in de projecten van het *Center for Digital Humanities* van de Universiteit van Amsterdam en het coöpererende *Creative Research Centre Amsterdam* ligt in de presentatie voor de buitenwereld veel nadruk op de ontwikkeling van nieuwe producten en diensten (‘valorisatie’ door bijvoorbeeld apps) en minder op wetenschappelijke vraagstellingen. Daardoor is het soms niet makkelijk in deze projecten de claim terug te zien die op de website van CDH wordt geformuleerd: ‘Each project involves a *research question* from the humanities of which the solution is approached by means of a proof of concept or a concrete product using techniques from DH’.

Het socio-economische en politieke krachtenveld waarin deze onderzoeken zich bewust of onbewust plaatsen komt duidelijk naar voren in de titel van het project *Creative Industries in the Dutch Republic*.¹⁶ Terwijl het project inhoudelijk vooral gericht is op marktstrategieën uit de vroegmoderne geschiedenis van het Nederlandse boekbedrijf haakt de titel aan bij een hedendaags economisch paradigma: de creatieve industrie. Deze retroprojectie van een recent concept op een bloeiperiode van de Nederlandse cultuur heeft een politieke lading: ze verleent indirect historische legitimatie aan actueel beleid dat, geïnspireerd onder andere door Richard Florida’s omstreden studie *The Rise of the Creative Class* (2002), inzet op een socio-economische transformatie van urbane centra zoals de Nederlandse Randstad. Critici als Matteo Pasquinelli houden dit beleid verantwoordelijk voor de gentrificatie van steden en de uitbuiting van natuurlijke resources;¹⁷ de mediawetenschappers Geert Lovink en Ned Rossiter wijzen erop dat de ‘creative industry meme’ inmiddels ook een belangrijke factor in ‘research funding calls in the humanities’ (Lovink en Rossiter 2007: 15) is.

Voor de duidelijkheid: We willen niet betogen dat de DH op zich ‘neo-liberaal’, onkritisch of theoretisch ‘ouderwets’ zijn. Het gaat hier veeleer om bepaalde *Wahlverwandtschaften*, keuzes op basis van affiniteiten, met specifieke voor- en nadelen waarop we willen wijzen. Wij denken bijvoorbeeld dat geestesweten-

¹⁶ Project van de Vrije Universiteit Amsterdam in samenwerking met het *Descartes Centre for the History and Philosophy of the Sciences and the Humanities*, Huygens/ING (KNAW) en het Amsterdams Centrum voor de Studie van de Gouden Eeuw.

¹⁷ Cf. Lovink & Rossiter 2007: 13.

schappers die, in de woorden van Derek Attridge, ‘nostalgisch’ terugverlangen naar een onbekommerde, humanistisch georiënteerde vakbeoefening *before theory* misschien niet de beste bondgenoten zijn in de ontwikkeling van het volle potentieel van DH.¹⁸ Daar staat tegenover dat een slim *gebruik* van de verschillende *Wahlverwandtschaften* die ontwikkeling juist kan stimuleren. Cultuur- en media-wetenschapper Cathy Davidson stelt in dit verband: ‘It is difficult to understand why so many humanists feel irrelevant to a culture that names the “creative class” as one of its defining features. [...] [H]umanists should be addressing the role of humanistic culture and cultural policy in the neoliberal economy instead of wringing their hands over the lack of their role in that economy’ (Davidson 2008: 708).

7 Niches en verbanden

We moeten voorkomen dat met DH gebeurt wat eerder met de empirische literatuursociologie in Nederland is gebeurd, die vorige poging om de klassieke literatuurwetenschap het aanzien te geven van een ‘echte wetenschap’: het ontstaan van een sterke maar tamelijk geïsoleerde onderzoekstraditie die ‘binnen de literatuurwetenschap tot niet veel meer [heeft] geleid dan een al dan niet meewarige of onverschillige acceptatie’ (Keunen e.a. 2009: 86). In de jaren 70 zong de empirische literatuursociologie zich los van de literatuurwetenschap, radicaliseerde, en verdween in een niche met eigen en niet erg geesteswetenschappelijke spelregels – een niche van waaruit de literatuurwetenschap inmiddels met bevreemding, en soms onwelwillendheid, bekeken wordt. Weinig is in de afgelopen dertig jaar onproductiever geweest voor het vakgebied dan deze scheiding.

Iets vergelijkbaars mag met DH niet gebeuren. We moeten geen nieuw eigenstandig filiaal van de geesteswetenschappen laten ontstaan, geen aparte tak van sport met eigen regels en eigen vooronderstellingen. DH moeten daarentegen integraal onderdeel worden van het geesteswetenschappelijke repertoire aan methoden en technieken. Een groter probleem dan het humanistische ressentiment tegen empirische en digitale benaderingen – ‘teksten laten zich niet vangen in cijfertjes’; ‘maar wij hebben het over de mens...’ etc. – lijkt ons daarbij juist de groeiende irritatie van poststructuralistisch geïnspireerde onderzoekers en collega’s die inmiddels voor een deel met verveling en desinteresse reageren als we de term ‘digital humanities’ ook maar noemen (‘Jij ook, Brutus?’). Zonder deze houding te willen verontschuldigen, denken we dat ze te maken heeft met de in deze notities beschreven ontwikkelingen van de DH en een struikelblok kan vormen voor toekomstige samenwerking. Dat zou juist ook voor de digi-sceptische literatuurwetenschappers een gemiste kans zijn: nieuwe methoden en technieken stellen ook hen in staat om hun – onze – *eigen* onderzoeksvragen beter te beantwoorden, professionele intuïties beter te onderbouwen en nieuwe vragen te stellen. Van elkaar leren betekent niet elkaars agenda’s overnemen; empirisme is geen kwaadaardige epistemologische ziekte, maar een belangrijk onderdeel van de (zelf)kritische vermogens, ook van geesteswetenschappers.

18 Attridge 2004: 10.

8 Laboratoria van de toekomst

Op het eerste gezicht kan de term ‘Digital Humanities’ vreemd klinken: zijn er dan bijvoorbeeld ook ‘papieren’ geesteswetenschappen? En is het niet dubieus om te suggereren dat het hier om een nieuw onderzoeksparadigma gaat, terwijl er ‘alleen’ nieuwe *tools* en methodes worden gebruikt? Misschien niet. Jan Blommaert parafraseert McLuhan: ‘Wanneer de technologie van de kennis verandert, verandert het wezen van de kennis zelf en verandert de mens’ (Blommaert 2012). Het ligt dus juist voor de hand om van de digitalisering ook fundamentele verschuivingen van het onderzoek, weten en culturele processen te verwachten. De ‘geesteswetenschappen’ zoals we ze sinds de achttiende eeuw in Europa kennen, zijn traditioneel inderdaad een *papieren* en een boekwetenschap, misschien nog meer wat hun *eigen* materialiteit en formats (boeken, handboeken, artikelen etc.) dan die van hun onderzoeksobjecten betreft. Aan beide kanten, die van de onderzoeksobjecten en die van de vraagstellingen en methodes, zal de digitalisering tot grote verschuivingen leiden, en heeft dat in de afgelopen twintig jaar al gedaan (bijvoorbeeld door het ontstaan van nieuwe manieren van lezen, ‘power browsing’, het gebruik van zoekmachines etc.). Maar dan is het misschien op een gegeven moment ook tijd om afscheid te nemen van de ‘geesteswetenschappen’ en op zoek te gaan naar andere, meer productieve en eigentijdse formateringen van ons onderzoek.

Het valt op dat veel *digital humanists* deze consequentie niet echt willen trekken. Cathy Davidson bijvoorbeeld heeft het over ‘geesteswetenschappen 2.0’ en bedoelt daarmee: geesteswetenschappen die digitale methodes niet alleen als instrument gebruiken (‘humanities 1.0’), maar ook zelf door deze technieken veranderd worden: ‘Data transform theory; theory, stated or assumed, transforms data into interpretation’ (Davidson 2008: 710). Toch komt Davidson uiteindelijk uit op een vrij klassieke definitie van de humanities 2.0 als *geesteswetenschappen* met de taak van reflectie en kritiek: ‘The humanistic turn of mind provides the historical perspective, interpretative skill, critical analysis, and narrative form required to articulate the significance of the scientific discoveries of an era’ (Davidson 2008: 707). Geesteswetenschappen als *reflectie op* en zingeving van de natuurwetenschappen dus.

Ook Rens Bod houdt vast aan de term ‘humanities’, maar heeft het in een overtreffende trap wel over de ‘geesteswetenschappen 3.0’. Dat zijn digitale geesteswetenschappen die ‘kritische reflectie over de schitterende producten van wetenschappelijke kennis’ (Bod 2012: 12) met het gebruik van nieuwe technologieën combineren. Opvallend genoeg blijft de term ‘geesteswetenschappen’ hierbij juist bestaan en krijgt met het adjectief ‘digitaal’ zelfs een vitaliserend infuus van eigentijdse betekenis en legitimatie. Maar zou het niet consequenter zijn om het concept van de geesteswetenschappen zelf in twijfel te trekken en actief te werken aan interdisciplinaire denkstijlen en vormen van onderzoek buiten de *comfort zone* van de *humanities*? Misschien zijn de ‘geesteswetenschappen 3.0’ juist nog niet post-humanistisch *genoeg*?

In de literatuurwetenschap zien we in de afgelopen twee decennia een verbaazingwekkende en soms verwarrende hoeveelheid van kleinere en grotere *turns* die vanuit het poststructuralistische paradigma vertrekken, maar daarbij ook nieuwe paden inslaan, zoals de *affective*, *neuro*, *cognitive* en *ecocritical turn*. Op de

lange termijn blijken deze zoekbewegingen misschien het begin van een nieuwe vorm van kennis waarin mens- en biowetenschappen, natuur- en geesteswetenschappelijke denkstijlen op ongekende manieren (meervoud!) in elkaar overgaan. ‘Computational approaches facilitate disciplinary hybridity that leads to a post-disciplinary university’, meent Berry zelfs, waarbij hij aantekent dat die gedachte ‘deeply unsettling’ kan zijn ‘to traditional academic knowledge’ (Berry 2011: 13).

Of Berry’s visioen werkelijkheid wordt, staat te bezien. Maar het lijkt wel zeker dat de ‘computational turn’ een belangrijke rol gaat spelen in het naar elkaar toegroeien van disciplines (en hun methodes en vraagstellingen). Volgens wetenschapshistoricus Hans-Jörg Rheinberger (1994) zijn experimentele systemen in de natuurwetenschappen meer dan toetsmethodes voor hypothesen: ze kennen een eigen dynamiek en brengen, als ze productief zijn, juist nieuwe vraagstellingen en epistemische objecten voort – dingen die niemand van tevoren had verwacht en die zich in het begin nauwelijks laten benoemen. Als we de DH als zo’n laboratorium van de toekomst kunnen beschouwen, zal de kritiek op de ‘ouderwetse’ concepten van de computationele literatuurwetenschap wellicht snel overbodig worden. Aan de andere kant kan het ontstaan van nieuwe kennisobjecten en vraagstellingen misschien ook helpen om de poststructuralistische literatuurwetenschap te vernieuwen en niet als een soort theologie van de eenentwintigste eeuw te conserveren: esoterisch, scholastisch en opgesloten in haar eigen gelijk. Dan zou de uitwisseling die we hier bepleiten een nuttige zijn geweest.

Bibliografie

- Attridge 2004 – Derek Attridge, *The Singularity of Literature*. London/New York: Routledge, 2004.
- Bakhtin 1981 – Mikhail Bakhtin, *Problems of Dostoevsky’s Poetics*. Vertaling: C. Emerson. Minneapolis: University of Minnesota Press, 1981.
- Berry 2011 – David M. Berry (2011), ‘The Computational Turn: Thinking about the Digital Humanities’. In: *Culture Machine* 12 (2011), p. 1-21. <http://www.culturemachine.net/index.php/cm/article/viewdownloadinterstitial/440/470> (geraadpleegd 1-9-2013).
- Blommaert 2012 – Jan Blommaert, *Over open en gesloten geesten: pleidooi voor een geest-rijke wetenschap*. Tilburg: Tilburg University, 2012.
- Bod 2012 – Rens Bod, *Het Einde van de Geesteswetenschappen 1.0*, Oratie UVA, 2012. <http://staff.science.uva.nl/~rens/OratieRens.pdf> (geraadpleegd 9-13-2013).
- Compagnon 2004 – Antoine Compagnon, *Literature, Theory, and Common Sense*. Princeton/Oxford: Princeton University Press, 2004.
- Culler 2006 – Jonathan Culler, ‘Comparative Literature, at Last’. In: Haun Saussy (red.), *Comparative Literature in an Age of Globalization*. Baltimore: Johns Hopkins UP, 2006, p. 237-247.
- Van Dalen-Oskam 2013 – Karina van Dalen-Oskam, *De stijl van R.* Oratie UVA, 2013. http://www.oratiereeks.nl/upload/pdf/PDF-1290Weboratie_Van_Dalen-Oskam.pdf (geraadpleegd 9-13-2013).
- Davidson 2008 – Cathy N. Davidson, ‘Humanities 2.0: Promise, Perils, Predictions’. In: *PMLA* 123 (2008) 3, p. 707-717.
- English 2012 – James English, ‘Morettian Picaresque’. In: *Los Angeles Review of Books* 27 (June 2012). <http://lareviewofbooks.org/essay/franco-morettis-distant-reading-a-symposium> (geraadpleegd 10-4-2013).
- Foucault 2002 – Michel Foucault, *Archaeology of Knowledge*. Vertaling: A. Sheridan. London: Routledge, 2002.
- Gitelman 2013 – Lisa Gitelman (red.), *‘Raw Data’ is an Oxymoron*. Cambridge, MA: MIT Press, 2013.
- Hall 2012 – Gary Hall, ‘Has Critical Theory Run Out of Time for Data-Driven Scholarship?’ Blogpost op *Debates in the Digital Humanities*, 2012. <http://dhdebates.gc.cuny.edu/debates/text/14>

- (geraadpleegd 10-4-2013).
- Jockers 2013 – Matthew L. Jockers, *Macroanalysis. Digital Methods and Literary History*. Urbana et al: University of Illinois Press, 2013.
- Jones 2012 – Rodney H. Jones, *Discourse Analysis. A Resource Book for Students*. London, New York: Routledge, 2012.
- Keunen e.a. 2009 – Bart Keunen & Daan Vandenhoute & Laurence Van Nuijs, 'Literatuursociologie in de Lage Landen'. In: Sascha Bru en Anneleen Masschelein (red.), *Tijding en tendens. Literatuurwetenschap in de Nederlanden*. Gent: Academia Press, 2009, p.77-90.
- Lovink en Rossiter 2007 – Geert Lovink & Ned Rossiter (red.), *My Creativity Reader. A Critique of Creative Industries*. Amsterdam: Institute of Network Cultures, 2007.
- Mani 2010 – Inderjeet Mani, *The Imagined Moment. Time, Narrative and Computation*. Lincoln/Londen: University of Nebraska Press, 2010.
- Manovich 2001 – Lev Manovich, *The Language of New Media*. Cambridge, London: MIT Press, 2001.
- Moretti 2013 – Franco Moretti, *Distant Reading*. London/New York: Verso, 2013.
- Mourits & Vaessens 1997 – Bertram Mourits & Thomas Vaessens, 'This game is shareware. Internet en de letterkundige neerlandistiek'. In: *Nederlandse Letterkunde* 1 (1997), p. 75-84.
- Rheinberger 1994 – Hans-Jörg Rheinberger, 'Experimental Systems: Historiality, Narration, and Deconstruction'. In: *Science in Context* 7 (1994) 1, p. 65-81.
- Rommel 2004 – Thomas Rommel, 'Literary Studies'. In: *A Companion to Digital Humanities*. Oxford: Blackwell, 2004. <http://www.digitalhumanities.org/companion/> (geraadpleegd 10-4-2013).
- Sedgwick 1985 – Eve K. Sedgwick, *Between Men. English Literature and Male Homosocial Desire*. New York: Columbia UP, 1985.

Adres van de auteurs

Departement Neerlandistiek, Moderne Nederlandse Letterkunde
 Spuistraat 134
 1012 VB Amsterdam
 T.L.Vaessens@uva.nl
 S.Besser@uva.nl