

Naar een sociologie van de verbeelding
Crowdsourcing als hermeneutisch experiment

RMA Scriptie Nederlandse letterkunde

Naam Lucas van der Deijl
Nummer 3689115
Begeleiders Dr. Saskia Pieterse
Prof. dr. Els Stronks
Datum 28 juni 2016



Universiteit Utrecht



Inhoud

1.	Inleiding	2
2.	Diversiteit in de Nederlandstalige literatuur	4
	2.1 Genderdiversiteit	
	2.2 Etnische diversiteit	
	2.3 Het ‘zuivere’ literatuurideaal	
	2.4 Probleemstelling	
3.	Methoden	11
	3.1 Web 2.0 en crowdsourcing	
	3.2 Technologie	
4.	Theoretisch kader	17
	4.1 Een hermeneutiek van crowdsourcing	
	4.2 Een epistemologie van crowdsourcing	
5.	Analyse	27
	5.1 Grafieken, kaarten, tabellen	
	5.2 <i>Het jongensuur</i>	
	5.3 <i>De asielzoeker</i>	
	5.4 <i>Het gym</i>	
6.	Receptie en procesevaluatie	48
	6.1 Reacties in de media	
	6.2 Reacties op blogs en fora	
7.	Conclusie	53
8.	Literatuur	55
Bijlagen		
1.	Voorbeeldlemma database	
2.	Formulierinstructies	
3.	Media-overzicht Personagebank	
4.	Informatieschema na ISBN-input	

1. Inleiding

In 2015 stelde Karin Amatmoekrim in *De Groene Amsterdammer* de vraag: ‘mag literatuur neutraal zijn?’ In dit onderzoek wil ik haar overdenking aanvullen met de wedervraag: ‘*kan* literatuur neutraal zijn?’ Iedereen die de afgelopen veertig jaar een inleidende cursus literatuurwetenschap gevolgd heeft, moet zo’n basale vraag in de oren klinken als een platitude, een open deur. Want natuurlijk kan ze dat niet. Literatuur maakt immers deel uit van uiteenlopende niet-literaire machtsdiscoursen, zegt Foucault, of van economieën van waardering, zegt Bourdieu. Literatuur ontstaat altijd in een bepaalde setting die door politiek, cultuur en geschiedenis bepaald is, volgens Greenblatt. Literatuur bestaat bovendien uitsluitend in de interactie tussen de subjectieve lezer en de tekst, volgens de empirische literatuurwetenschap. Neutraliteit, kortom, bestaat niet.

Toch is het deze neutraliteit die de Nederlandse literatuur achtervolgt en die eens in de zoveel jaren critici en literatuurwetenschappers tot felle discussies weet te verleiden. ‘Neutraal’ kent daarin meerdere betekenissen maar altijd is het predicaat belast met een poëtische lading. Amatmoekrim lijkt het woord te gebruiken om te benoemen dat de Nederlandse letteren ‘niet uitgesproken’ is, ‘geen sterke sprekende waarde’ heeft (Van Dale). In haar stuk, dat de ‘ondraaglijke witheid van de Nederlandse letteren’ aan de kaak stelt, schrijft ze: ‘De praktijk van de letteren is weinig volks, weinig van buiten in het algemeen en daardoor schrikbarend neutraal.’ In mijn wedervraag – ‘kan literatuur neutraal zijn?’ – stel ik eerder een eigenschap van literatuur aan de orde die in de letterkunde met de inmiddels wat sleetse term ‘autonomie’ wordt geassocieerd. Van Dale definieert die betekenis als ‘vrij van mogelijke (ongunstige) beïnvloeding’. Een ‘neutrale’ literatuur zou dan er dan een zijn die niet toegeeft aan invloeden vanuit het politieke, religieuze, economische of welk domein dan ook. Het is een literatuur die gewaardeerd wordt vanwege een zekere eigenstandigheid; die haar bestaansrecht niet ontleent aan extra-literaire invloeden en die zich daardoor veel kan permitteren voordat lezers en critici zich ertegen verzetten. Sterker nog, de ongebondenheid van literatuur wordt met hand en tand verdedigd, en lezers die zich wél kritisch opstellen – zoals Amatmoekrim – wordt onbegrip verweten.

Zoals veel publieke debatten zijn ook recente poëtische discussies over de Nederlandse literatuur niet zozeer een weging van argumenten maar een botsing van wereldbeelden, in dit geval over de literaire wereld. De ‘autonomistische’ literatuuropvatting van de literaire elite botst met de cultuuranalytische en historiserende houding die letterenfaculteiten uitdragen sinds het poststructuralisme, een houding die ook Amatmoekrim meekreeg in de collegebanken. Treffend voor deze academische distantie was een profiel uit 2009 van de destijds recent aangestelde hoogleraren moderne Nederlandse literatuur Thomas Vaessens, Geert Buelens en Jos Joosten. ‘Ontmaskeraars van de goede smaak’ werden ze genoemd (De Boer 2009). Wanneer academici uit deze school zich mengen in het poëtische debat, verandert het literaire veld in een linie die hen met veel retorisch geweld wil terugdringen naar de academie. Toen Thomas Vaessens in *De revanche van de roman* opriep tot een meer geëngageerde lectuur van hedendaagse romans, verweet Joke Hermsen hem een ‘moord op de literatuur’ en Arie Storm noemde de Amsterdamse hoogleraar ‘tendentieus en leugenachtig’. Ook toen Saskia Pieterse in 2014 pleitte voor meer oog voor de ethische en kritische houding van de lezer in een stuk met de veelzeggende titel ‘De emancipatie van de lezer’, achtte *Vrij Nederland*-criticus Carel Peeters haar rijp voor een Politbureau (Pieterse 2014; Peeters 2014). De doctrine van neutraliteit en autonomie biedt de gevestigde critici een stok om pleitbezorgers van andere leeshoudingen weg te slaan uit de vrije proeftuin die zij voor literatuur aanzien.

Dit onderzoek is een poging om tegelijkertijd een brug te slaan én een confrontatie te ensceneren tussen de academische leeshouding en de talloze andere attitudes die lezers aanwenden tijdens het lezen. Vanuit de overtuiging dat de academische en de niet-academische (voor het gemak schets ik hier een dichotomie) van elkaar kunnen leren maar ook meer gemeen hebben dan beide partijen misschien willen toegeven, hoop ik een middel te presenteren waarmee we een complex probleem kunnen oplossen zoals

een letterkundige alleen dat nooit zou kunnen. Ik wil dat doen aan de hand van een thema dat ik hier als casus beschouw, maar dat evengoed in een andere context een zelfstandige, intrinsieke status verdient: de diversiteit van romanpersonages in recente Nederlandstalige literatuur (1945 tot heden). Dat thema is hier relevant omdat het de inzet was van een recente poëtische discussie die zowel kritische als ‘autonomistische’ lezers mobiliseerde.

Het middel dat ik voor ogen heb is een crowdsourced, open database van romanpersonages die de diversiteit onder Nederlandse romanpersonages moet blootleggen. Via een web-interface worden lezers uitgenodigd om informatie over personages uit door hen gelezen romans in te vullen en kunnen zij actuele verhoudingen in de database inzien en opvragen. Voor zover ik weet werd er niet eerder geëxperimenteerd met het verzamelen van lezersinterpretaties om kwantitatieve uitspraken te doen over *de gelezen literatuur* (in plaats van over lezers, zoals de meeste studies binnen de receptie-esthetica beogen, of over formele eigenschappen van teksten, zoals de computationele stylometrie doet). Dit onderzoek is dan ook een methodologisch experiment en een kritisch-theoretische reflectie op dat experiment ineen. De webapplicatie die ik ervoor ontwikkelde biedt in potentie de basis voor een kwantitatieve en transhistorische vorm van literatuurwetenschap die hypothesen nadrukkelijk beschouwt als toetsbaar. In werkelijkheid levert zo’n digitale synthese van geformaliseerde interpretaties uiteenlopende hermeneutische, epistemologische en ideologische problemen op. Over die problematiek gaat deze scriptie, met als hoofdvraag:

In hoeverre kan crowdsourcing via geformaliseerde romaninterpretaties valide en generaliseerbare kennis voortbrengen over een specifieke aspect (representaties van personages) van deze romans en in welke mate bemoeilijkt een verschil in leeshouding onder de participanten deze generaliseerbaarheid?

Deze vraag heeft een analyserend en een evaluerend karakter. Gesteund door de theoretische expertises van de hermeneutiek en de digital humanities hoop ik allereerst de problemen van deze vorm van crowdsourcen te ontleden en tot een werkbare – dat wil zeggen hypothese-stellende – theorie van het crowdsourcen van interpretaties te komen. Vervolgens wil ik de waarde en relevantie van deze methode evalueren door zowel het potentieel als de beperkingen te onderzoeken, respectievelijk met een interpretatie van de data uit de database en via een drietal roman-analyses.

Met dit onderzoek heb ik twee producten voor ogen:

1. Een technologie die een vorm van crowdsourcen voor letterkundig onderzoek mogelijk maakt.
2. Een analyse van de waarde, relevantie en problematiek van die technologie.

Omdat de methode, onderdeel 1, zo’n centrale rol speelt in dit onderzoek en omdat de theorie een referentiekader biedt voor een evaluatie van die methode eerder dan voor de operationalisering zelf, wijkt de structuur van deze scriptie enigszins af van de gebruikelijke vorm. Hieronder volgt allereerst een korte analyse van de poëtische discussie over diversiteit in de Nederlandstalige literatuur, omdat het voorgestelde middel zich positioneert in dat debat. Daarna beschrijf ik de technologie zelf onder de kop ‘Methode’ en daarna volgt pas een theoretische reflectie. Ik draai deze twee onderdelen om, omdat het in mijn ogen noodzakelijk is te begrijpen wat het platform precies doet en inhoudt voordat een theoretische reflectie daarop mogelijk is. Na de theorie problematiseer ik de methode verder door met close readings van enkele door lezers ingevulde romans de hermeneutische en epistemologische implicaties van een geformaliseerde interpretatie te demonstreren. Ook evalueer ik het proces en de reacties op het project in verschillende media.

2. Diversiteit in de Nederlandstalige literatuur

In wat volgt hoop ik te illustreren hoe leeshoudingen en literatuuropvattingen kunnen botsen aan de hand van een recente discussie over een vermeend gebrek aan diversiteit in literaire instituties en de Nederlandstalige literatuur als geheel. Ik wil deze leeshoudingen contextualiseren en historiseren, om duidelijk te maken welke belangen en overtuigingen hier precies wringen. Deze beschrijving maakt hopelijk inzichtelijk dat er een hiërarchie bestaat van literaturopvattingen, die een uitwisseling van interpretaties of argumenten tussen leeswijzes en critici bemoeilijkt. Dat probleem is relevant voor dit onderzoek, omdat het mogelijk ook succesvolle crowdsourcing van romaninterpretaties in de weg staat. In de beschrijving van de discussie maak ik een onderscheid tussen genderdiversiteit en etnische verscheidenheid.

2.1 Genderdiversiteit

De recente discussie over genderdiversiteit op het literaire veld is terug te leiden tot een bijdrage van Daphne van Paassen in *Opzij* (2013). Van Paassen telde een half jaar lang de recensies in Nederlandse boekenbijlagen en concludeerde dat vrouwelijke auteurs beduidend minder vaak besproken worden dan hun mannelijke collega's (27 tegenover 73 procent). Volgens Van Paassen duidt dat verschil op een kritische norm. Ze parafraseert letterkundigen Thomas Vaessens en Jacqueline Bel, die in *Schrijvende vrouwen* beweerden dat de verhouding tussen man en vrouw in het boekenaanbod 'sinds de jaren zeventig steeds evenwichtiger geworden is' (geciteerd in Van Paassen 2013). De critici selecteren dus in het voordeel van de mannen. Die vaststelling vindt bijval van schrijver Pauline Slot. In *Trouw* wijst ze op de ongelijke verdeling onder schrijvers van het jaarlijkse Boekenweekgeschenk en onder de genomineerden op de Libris shortlists. Net als Van Paassen situeert Slot het probleem bij de critici in plaats van de uitgevers: het 'is niet omdat er geen vrouwen zijn. Ze dringen alleen niet door tot de eregalerijen' (Slot 2013). Slot signaleert bovendien dat critici en juryleden zich vaak beroepen op hun onfeilbare oog voor kwaliteit, dat zij in twijfel trekt: 'Maar kan iemand ooit van zichzelf beweren dat hij of zij door geen enkele vooringenomenheid wordt geplaagd?'

Op recensieplatform RecensieWeb reageert John Hermse sceptisch op het geturf van *Opzij*. Hermse ging zelf op onderzoek uit en concludeert op basis van de groslijsten van de Academia-debutantenprijs en de AKO Literatuurprijs dat we de ongelijkheid misschien toch aan de uitgevers moeten wijten: slechts een derde van auteurs uit die selecties was vrouwelijk (Hermse 2013). Volgens Hermse is het dan ook voorbarig om de literatuurkritiek vooringenomenheid te verwijten (hoewel hij vergeet te vermelden dat de 1/4-verhouding in de kritiek nog altijd geen representatieve weerspiegeling is van de 1/3-verhouding in de groep die hij gebruikt als indicator voor de productie). Tot slot schermt de internetcriticus met het beroep op kwaliteit dat Slot al zo vaak hoorde in deze discussie. *Opzij* bijt zichzelf in de staart met zijn prijs voor uitsluitend vrouwelijke auteurs, zo vindt Hermse: Manon Uphoff bijvoorbeeld, hun laatste laureate, werd immers ook geshortlist voor de AKO Literatuurprijs en de Libris Literatuurprijs. Een auteur die zowel door mannelijke als vrouwelijke critici tot een belangrijk schrijver wordt bekroond, bewijst voor Hermse dat er van vooringenomenheid geen sprake is. Uphoff is gewoon een goede schrijver.

In *Trouw* biedt hoogleraar Jos Joosten een aanvulling op de bevindingen van Slot. Net als Hermse vermoedt Joosten dat het probleem niet (alleen) ligt bij de critici. Op basis van alle boeken die sinds 1945 bij de grote literaire uitgeverijen gepubliceerd zijn, stelt hij vast dat de man/vrouwverhoudingen onder de gepubliceerde auteurs net zo scheef zijn als die binnen de groep die het blikveld van de critici bereikt (Joosten 2013). Zijn verdeling benadert die van Hermse (65 versus 35 procent). Joosten wijst erop dat vrouwelijke auteurs beter verkopen dan mannelijke: op de top-100 bestverkochte auteurs in 2010 was de

verhouding al beter in balans (55 versus 45 procent). Hier raakt de hoogleraar dus aan een verschil in voorkeuren tussen verschillende lezersgroepen: institutionele versus commerciële groepen; professionele versus recreatieve lezers.

Het is precies dat verschil tussen institutionele normen en de culturele praktijk dat drie jaar later de discussie weer doet oplaaien. Onder de schuilnaam 'Lezeres des Vaderlands' verwoordt een anonieme critica in *De Groene Amsterdammer* haar frustratie over de structurele miskenning van de verscheidenheid aan lezersgroepen op de literaire en culturele afzetmarkt (2016). Op haar blog slaat zij aan het tellen van recensies in boekenbijlagen. Het aantal besproken of besprekende vrouwen komt vaak niet boven de 25 procent uit – critici blijken sinds de *Opzjj*-telling in ieder geval consequent. En ook De Lezeres laakt het beroep op kwaliteit dat de kritische molen de eigen criticasters voorhoudt:

'Het argument luidt altijd: dat vrouwen minder goed vertegenwoordigd zijn is geen opzet, we selecteren enkel op basis van Kwaliteit – of iemand man, vrouw, wit, zwart of wat ook ertussenin is, dat doet er niet toe. Dit argument kan mij intussen niet meer overtuigen. Kwaliteit is namelijk geen objectief gegeven' (Lezeres des Vaderlands 2016).

Daar denkt Jamal Ouariachi anders over. In een stuk in *Vrij Nederland* uit 2014 kaatst hij de bal terug naar Saskia de Coster en Fleur Speet, die in *De Morgen* de masculiene blik in de literatuurkritiek aan de kaak stelden (Speet 2014; De Coster 2014). Volgens hem ligt de desinteresse en lage waardering voor vrouwenliteratuur aan de vrouwen zelf (Ouariachi 2014). Ook hij speelt de kwaliteit-kaart: boeken van mannen zijn nu eenmaal 'opvallender' (want dikker), en daarom moeten vrouwen de lat hoger leggen. Opdat zij opgemerkt worden, masculiene blik of niet.

Ouariachi's provocatieve bijdrage ontlokt vele reacties. Een daarvan is van Lieke Marsman, in *Tirade*. Ze ergert zich aan Ouariachi's betuttelende oproep en hekelt het feit dat een toonaangevend medium als *Vrij Nederland* een ongenueanceerde stem verkiest boven een genuanceerder en correcter verhaal (Marsman 2014). Voor Marsman is het een teken dat het mannenbolwerk van de literaire kritiek de status quo wel best vindt, hoewel ze gelooft dat die mannen niet een bewuste misogyne houding aannemen. Maar, zo benadrukt ze, het feit dat het om 'onbewuste processen' gaat, maakt duidelijk dat er iets veranderen moet.

Wat de oorzaak ook moge zijn, de discussie levert genoeg argumenten om aan te nemen dat er inderdaad sprake is van een achterstelling van vrouwelijke auteurs op het Nederlandstalige literaire veld. Er lijkt een geautomatiseerde, veelal onbewuste, 'vooringenomen' reactie te bestaan ten aanzien van zowel mannelijke als vrouwelijke auteurs. Die reactie leidt ertoe dat vrouwen minder vaak besproken worden (Van Paassen 2013; Lezeres des Vaderlands 2016; Speet 2014); minder vaak in de prijzen vallen (Slot 2013; De Coster 2014; Speet 2014) en minder gelezen worden op middelbare scholen (Lezeres des Vaderlands 2016). Ook heeft deze geijkte reactie tot gevolg dat vrouwen vaak met mannen vergeleken worden maar niet andersom (Weijers 2014); dat het beeld van de Grote Schrijver mannelijk geconnoteerd is (Weijers 2016; De Coster 2014; Van der Veer 2014) en dat literaire prijzen uitsluitend naar mannen vernoemd worden (Weijers 2014).

In antwoord op deze argumenten zijn er critici of schrijvers zoals John Hermse en Jamal Ouariachi die het bestaan van zo'n geautomatiseerde lezersreactie ontkennen. Voor hen doen ogenschijnlijk contextuele factoren zoals het gender van de auteur of het personage niet ter zake. Critici zijn volgens hen 'neutraal' en in staat om literaire kwaliteit onvoorwaardelijk te beoordelen. De enig denkbare conclusie voor hen is dat het aan de kwaliteit van de vrouwelijke schrijvers ligt. En wie die neutraliteit in twijfel trekt, begrijpt niet waar de werkelijke oorzaken liggen (de uitgevers, de vrouwen) of maakt zich schuldig aan politieke correctheid.

2.2 Etnische diversiteit

De discussie over etnische diversiteit binnen de Nederlandstalige literatuur voltrekt zich via vergelijkbare patronen en stellingnames. Een eerste bijdrage is het geruchtmakende essay van Anil Ramdas getiteld ‘Moedwil en kwade trouw bij blanke schrijvers. Niemand heeft oog voor het vreemde’, gepubliceerd in *NRC Handelsblad* op 14 maart 1997. Ramdas beschouwt het als een ‘formidabele wanprestatie’ dat witte schrijvers geen oog hebben voor identiteiten die niet slechts ‘projecties van zichzelf’ zijn, identiteiten die bovendien in een multiculturele samenleving verre van schaars zijn. Als voorbeeld bespreekt hij *Zionoco* (1995) van Leon de Winter en *De buitenvrouw* (1994) van Joost Zwagerman, waarin de zwarte Ander volgens Ramdas stereotypisch en uitsluitend door een witte bril gerepresenteerd wordt.

Joost Zwagerman reageert twee weken later in dezelfde krant. Hij verwijt zijn collega literaire met morele oordelen te vermengen. Daarnaast vindt Zwagerman dat Ramdas de belevingswereld van personages verwart met die van hun auteur, een denkfout die ‘zelfs een eerstejaars student Nederlands wel uit zijn hoofd zou laten’ (Zwagerman 1997). Ramdas begrijpt literatuur niet, volgens Zwagerman. Daarnaast vindt hij dat je niet van de auteur kunt verlangen dat zij/ hij buiten de grenzen van de eigen verbeelding treedt:

‘Iedere schrijver kiest een eigen vorm, een eigen stijl, en volgt vanzelfsprekend de eigen obsessies en preoccupaties, of hij nu de voorkeur geeft aan “straatruoer” en tijdsbeeld of aan literatuur als autonoom taalbouwsel en een verschansing in de aloude ivoren toren. Sociaal realisme of l’art pour l’art: de ene schrijver kan zich met polemische middelen teweerstellen tegen de praktische positionering van de ander, maar nóóit zal men elkaar het vanzelfsprekende recht ontzeggen om de plaats te kiezen die hem goeddunkt’ (Zwagerman 1997).

Een deel van deze discussie herhaalt zich acht jaar later, nadat Saskia Pieterse in NY een essay publiceert getiteld ‘De emancipatie van de lezer’. Ze signaleert dat sommige Nederlandse romans gelezen en bekroond worden vanwege hun engagement met de multiculturele samenleving, maar daarvoor altijd de lens van een Westers personage kiezen. Daarbij wordt het recht om zich te engageren voorbehouden aan de (witte) schrijver en ontzegd aan de lezer, zo laat Pieterse zien. Ze bespreekt een incident waarin enkele lezers bezwaar maakten tegen de representatie van zwarte vrouwen in Robert Vuijsjes *Alleen maar nette mensen* (2008). Deze groep kritische lezers werd verweten geen verstand te hebben van de literatuur, die slechts ‘een proeftuin van de verbeelding is’, niet te verwarren met de werkelijkheid (Pieterse 2014). De paradox die Pieterse aansnijdt is dus dat critici romans waarderen vanwege hun vermogen om zich te mengen in het maatschappelijke domein maar tegelijkertijd van lezers verwachten dat ze die zogenaamd geëngageerde literatuur als autonome, vrijblijvende objecten lezen.

Naar aanleiding van het stuk van Pieterse verdedigt Carel Peeters in *Vrij Nederland* het klassieke standpunt van de literatuurkritiek. De criticus interpreteert haar aanklacht als een poging om de literatuur te onderwerpen aan een eigen poëtica: ‘Saskia Pieterse wil dat de geëmancipeerde lezer iets te zeggen krijgt over de verbeeldingswereld van de schrijver’ (Peeters 2014). Dat zal leiden tot ‘ideologische en tendentieuze romans’. De enigszins retorische parafrase duidt op een fundamentele botsing van literatuuropvattingen. Peeters geeft met zijn weerlegging zijn autonomistische poëtica bloot: lezers hebben zich te voegen naar de wetten van het literaire werk, want dat werk stelt nu eenmaal zijn eigen wetten. Vanuit die overtuiging is het zinloos om literaire representaties kritisch te beschouwen, want schrijvers en romans opereren in het vacuüm van de verbeelding.

Een half jaar later kiest columnist Ebissé Rouw in het *NRC Handelsblad* positie aan de kant van het sociaal realisme, volgens de oppositie van Zwagermans. Ze ondergraaft het enthousiasme waarmee de literaire wereld het tijdschrift *Das Magazin* begroette: ‘Onlangs las ik in *Volkskrant Magazine* een lofzang op *Das Magazin* waarin een cruciaal punt over het hoofd wordt gezien: het is wit’, in tegenstelling tot de

hedendaagse samenleving, zoals ze net als Ramdas benadrukt (Rouw 2015). Rouw verschuift het accent in de discussie van Ramdas, Peeters en Pieterse naar etnische diversiteit *buiten* het literaire werk, op het literaire veld. Ze ziet Das Magazin als een voortzetting van de culturele segregatie die Nederland al zo lang kenmerkt.

Die kritiek krijgt bijval in een nummer van *De Groene Amsterdammer* in augustus 2015. Socioloog Thomas Franssen en recensent Daan Stoffelsen berekenden dat er inderdaad sprake is van een achterstelling van ‘allochtone’ schrijvers in de literaire wereld. Ze worden minder geboekt op literaire evenementen, ontvangen minder prijzen en publiceren ook verhoudingsgewijs minder, gemeten naar het aandeel allochtonen in de Nederlandse samenleving (Franssen & Stoffelsen 2015). Karin Amatmoekrim stelt in hetzelfde nummer dat het getuigt van luiheid om niet te zien dat er wel degelijk veel gebeurt in de niet-witte echelons van de literaire pikorde. Tegelijkertijd verwijt ze de Nederlandse literatuur haar ‘neutraliteit’, haar onvermogen om uitgesproken te zijn en om een ander etnisch of ethisch perspectief te kiezen dan het gangbare.

Twee weken na het stuk van Rouw toont in dezelfde krant Jan Kuitenbrouwer zich sceptisch over haar suggestie dat de literaire wereld lijdt aan een voorkeur voor witte schrijvers. Hij vermoedt dat allochtone schrijvers stomweg minder geïnteresseerd zijn in de literaire wereld (Kuitenbrouwer 2015). Verder vindt hij dat er nog geen overtuigend bewijs is dat deze groep tegengewerkt wordt in hun mogelijke wens om te publiceren. Ook Jamal Ouariachi mengt zich weer in de discussie en steunt het standpunt van Kuitenbrouwer. Zoals hij destijds de onderwaardering van vrouwelijke schrijvers aan de vrouwen toeschreef, zo wijt hij nu het probleem aan de allochtonen en hun gemiddelde lagere opleiding: ‘Misschien moeten we de intellectuele achterstand onder allochtonen dus eerst eens wegwerken – een taakje voor de politiek, lijkt me – voordat we de literatuur xenofob noemen’. (Ouariachi 2015)

Waar de discussie over genderdiversiteit vooral de institutionele achterstelling van vrouwelijke schrijvers betrof, gaat het in het debat over etnische diversiteit ook over de scheve verhoudingen van literaire representaties. Anil Ramdas, Saskia Pieterse, de lezers die zich verzetten tegen stereotyperingen in *Alleen maar nette mensen*; ze accepteren het niet dat de literatuur structureel de sociale werkelijkheid verdraait of negeert. Deze groep lezers verwacht dat de roman pas vernieuwend, vervreemdend, interessant of relevant wordt zodra die zich rekenschap geeft van de ideologie achter representatie. Te vaak zien zij, bij Zwagerman, De Winter, Vuijsje, Wieringa, dat schrijvers slechts een beperkte set persoonlijkheden aan het woord laten en andere identiteiten reduceren tot de overbekende stereotypen. Daarnaast ziet deze categorie critici de institutionele drempels waar behalve vrouwelijke ook niet-witte schrijvers voor gesteld worden.

Tegenover deze groep verwoorden critici als Carel Peeters, Joost Zwagerman, Jan Kuitenbrouwer en Jamal Ouariachi een autonomistische literatuuropvatting. Ze gaan ervan uit dat literatuur zich niet laat sturen en ontkennen het bestaan van discriminatie of representatienormen. Schrijvers zijn nu eenmaal subject aan hun eigen ‘preoccupaties’ en aan de sociale realiteit waarin zij leven. Als dat betekent dat ‘de Ander’ een secundaire positie krijgt of stemloos blijft, dan hebben wij dat als lezers te accepteren, en moeten we dat eerder toeschrijven aan het Nederlandse integratieprobleem dan aan de blinde vlekken van schrijvers (Peeters). Op het institutionele niveau werkt het argument als volgt: als niet-witte schrijvers minder gepubliceerd en gewaardeerd worden, dan zal dat wel te maken hebben met hun desinteresse in de literatuur (Kuitenbrouwer) of hun ‘intellectuele achterstand’ (Ouariachi). Zo legitimeert deze poëtische school de literatuur als een autonoom domein waarop niet-literaire mechanismen zoals discriminatie of structurele achterstelling op welk niveau dan ook, geen vat krijgen.

2.3 Het ‘zuivere’ literatuurideaal

Uit de hierboven beschreven discussie blijkt een zekere hiërarchie van literatuuropvattingen en lezersreacties. Zodra lezers de representatie van personages en daarmee de keuzes van schrijvers

bekritisieren, wordt hen verweten dat zij de literatuur niet begrijpen. Vertegenwoordigers van de zuivere literatuuropvatting diskwalificeren zo andere leesstrategieën; zij bezitten het patent op legitieme interpretaties. Waar komt die reflex vandaan?

Thomas Vaessens onderzocht de dominantie van de autonomenorm onder critici en schrijvers in zijn boek *De revanche van de roman* (2009). Bekend in dit verband is Mulisch' expliciete afzwering van de roman in 1969, omdat hij vond dat de problemen in de wereld om zijn aandacht vroegen maar zich moeilijk in de romanvorm aan de kaak lieten stellen. Vaessens stelt dat deze opstelling al decennia gemeengoed is onder literaire smaakmakers en schrijvers (Vaessens 2009: 11-12). Het is een standpunt:

‘dat de literaire connaisseurs sinds het einde van de negentiende eeuw steevast innemen. “Echte” literatuur is universeel en eeuwig: zij ontstijgt altijd de tijd en de plaats van haar ontstaan. Om het in inmiddels wel heel belegen geworden poëtische termen te zeggen: de roman is autonoom en bij voorkeur “zuiver” literair. Wat “zuiver” is, laat zich eigenlijk alleen definiëren in relatie tot het tegendeel ervan: “onzuiver”. Onzuiver is de roman die zichzelf al te nadrukkelijk verbindt aan het voorbijaande of de auteur die zich door andere dan literaire motieven laat leiden, bijvoorbeeld door “wat de gemoederen bezighoudt” en door het verlangen daarop in te spelen’ (Vaessens 2009: 11).

Ik beschouw het ‘zuivere’ literatuurideaal hier dus als synoniem voor wat Amatoekrim ‘neutraal’ noemde, of wat in andere contexten ‘autonoom’ heet. Het is hier niet mijn doel om de semantische nuances en betekenisverschillen tussen die termen te beschrijven; Vaessens definitie van ‘zuiver’ is voor mij werkbaar.

Ook Frans Ruiter en Wilbert Smulders dateren de oorsprong van deze houding in de late negentiende eeuw. Ze zien dat kunstenaars in deze periode zoeken naar nieuwe vormen van ‘transcendentie’ in een wereld die door de Verlichting en de industrialisering haar mythische glans verloren heeft, een proces dat Max Weber de ‘onttovering van de wereld’ noemde (Ruiter & Smulders 1996: 22). De Tachtigers introduceerden in reactie op het moralistische en brave realisme van de vroegere negentiende eeuw een vorm van autonome kunst, die dus voortkwam uit hun afkeer van de ontzielde, dagelijkse werkelijkheid: ‘Het opeisen van het recht om in volkomen vrijheid idiosyncrasieën uit te leven was een radicale poging tot zingeving binnen de nieuwe context van de moderne cultuur’ (Ruiter & Smulders 1996: 24).

De poëtica van deze kunstbeweging zou in korte tijd de dominante positie opeisen in een artistiek klimaat dat door avantgardistische en modernistische vernieuwingen werd beheerst. Die omwenteling ging ten koste van de artistieke status van vertegenwoordigers van een ‘heteronome’ kunstopvatting. De heteronome positie gold niet langer vanzelfsprekend als legitiem, maar dat betekende niet dat er geen animo meer was voor sociaal-realistische kunst. Er ontstond kortom een hiërarchische tweedeling tussen ‘transcendente’, universele en tekst-immanente literatuur enerzijds en teksten die een politieke, realistische of religieuze lading droegen anderzijds.

Die tweedeling was bovendien van meet af aan verankerd in de sociale stratificatie die vanaf de late negentiende eeuw langzaam geïnstitutionaliseerd zou raken tijdens de verzuiling. Ruiter en Smulders signaleren de interessante paradox tussen de artistieke, modernistische pretenties van de Tachtigers (en de avant-garde als geheel) en de processen van democratisering en emancipatie die volgens geschiedkundige definities de ‘moderne’ tijd aankondigden:

‘Op het moment dat de artistieke elite zich met nogal onmoderne, aristocratische allure als profeten en verkenners van het moderne levensgevoel opwierp, slaagden de confessionelen erin op een moderne wijze een massale achterban te organiseren rond een antimoderne ideologische kern’ (Ruiter & Smulders 1996: 25).

Maar tegelijkertijd

‘werd binnen de invloedssfeer van de zuilen door kunstenaars juist gepleit voor een heteronome kunst, een kunst dus die niet autonoom is, maar juist dienstbaar wil zijn aan een levensbeschouwelijke of politieke richting. De dominee-dichters hadden in de negentiende eeuw nog relatief onbekommerd hun stichtelijke dichtkunst kunnen beoefenen voor een min of meer gelovig publiek. Dertig jaar later was de toestand drastisch veranderd. De verzuilde erfgenamen van de domineedichters moesten hard vechten voor de culturele legitimatie van hun heteronome kunstopvatting’ (Ruiter & Smulders 1996: 25).

Deze artistieke tweedeling voltrekt zich dus langs sociale lijnen die via distinctiemechanismen vorm krijgen. Dat sociologische proces kan de sterke hiërarchie verklaren in de tegenstelling tussen ‘sociaalrealistische kunst’ en ‘l’art pour l’art’ zoals die bijvoorbeeld doorklinkt in Zwagermans antwoord op Anil Ramdas of Peeters’ repliek aan Pieters. In een land dat nog steeds de sporen draagt van de verzuiling, is het niet ondenkbaar dat literatuur- en cultuuropvattingen onderdeel geworden zijn van de groepsidentiteiten die door sociaal-economische segregatie jarenlang probleemloos konden voortbestaan. Kritiek op de dominante literatuuropvatting betekent dat men de inherente *appeasement*-strategieën van de verzuiling loslaat en dat de status die de literaire elite eraan ontleent, niet langer vanzelfsprekend is.

Hoe zeer het zuivere ideaal geïnternaliseerd is door de Nederlandse literaire kritiek blijkt uit de reactie van de journalistieke en academische literatuurbeschouwing op het postmoderne denken. In de jaren zeventig raakte dat gedachtegoed in zwang op universiteiten in Frankrijk en de Verenigde Staten, maar het Nederlandse literaire establishment toonde zich doof voor deze kritische herziening van het modernistische project. De journalisten van de *Haagse Post* waren bij wijze van uitzondering kort gecharmeerd van de beweging en droegen in een themanummer een soort polderpostmodernisme uit, een veilig afgeslankte vorm die de poëtische zuilen in stand hield. Vaessens typeert dit als een ‘op de (Amerikaanse) massacultuur georiënteerd, ludiek postmodernisme, met het relativistische *anything goes* als devies’ (Vaessens 2009: 56). Het kritische potentieel van de deconstructivistische methode en de postmoderne aanval op kolonialisme, racisme en seksisme die de radicale afrekening met modernistische termen als ‘universeel’, ‘schoonheid’ en ‘waarheid’ mogelijk maakte, werd niet gezien. Critici oarmden liever een radicale vorm van het relativisme dat de postmodernisten uitdroegen. Dat relativisme kreeg bijvoorbeeld vorm in de postmodernistische ontmaskering van de term ‘universeel’ als een retorische constructie die verhult dat wat universeel heet vooral mannelijk, wit en Westers is. Het doorgetrokken polderpostmodernisme besloot daarop dat alle ethische of politieke posities geconstrueerd waren en dus verdacht. Dat maakte de kleine postmoderne erfenis in Nederland behalve ludiek ook vrijblijvend, waardoor het onverhoopt een strategie werd om het zuivere literatuurideaal te behouden:

‘Bij diezelfde struisvogelpolitiek hoort ook dat de humanisten hun op het literaire modernisme georiënteerde poëtica niet heroverwegen. De steekwoorden van die poëtica – zoals ‘autonomie’ en ‘zuiverheid’ – kunnen nog altijd worden gehoord’ (Vaessens 2009: 60).

Ondanks de accentverschuivingen bleef en blijft het zuivere standpunt dus de norm. Meer dan een eeuw na de Tachtigers schieten critici of schrijvers bij sociaal-politieke of ideologische lezersreacties nog altijd in de reflex om een ‘sociaal-realistische’ of ideologische poëtica als niet-legitiem ter zijde te schuiven terwijl zij zich zelf automatisch beroepen op een zuiver, universeel literatuurideaal.

2.4 Probleemstelling

De zuivere literatuuropvatting kent dus een lange geschiedenis in de Nederlandse kritiek en heeft zich diep verankerd in de institutionele omgang met literatuur. In de marges klinkt er kritiek die voortkomt uit de

ideologiekritische leesstrategieën, maar die vinden weinig weerklank op de redacties van toonaangevende media. Er lijkt een segregatie van leeshoudingen of literaturopvattingen te bestaan, waarin een selecte kring van academische, critici en feministische of activistische lezers gedistantieerd is van een leescultuur die literatuur uitsluitend om haar artistieke kwaliteiten wenst te lezen. In een klimaat waarin zulke verschillen bestaan leidt een literair of poëticaal debat tot spraakverwarring en zijn argumenten al snel eerder ideologisch gemotiveerd dan wetenschappelijk geïnformeerd. In dat krachtenveld begeeft deze scriptie zich, met als centrale probleemstelling:

In hoeverre kan crowdsourcing via geformaliseerde romaninterpretaties valide en generaliseerbare kennis voortbrengen over een specifieke aspect (representaties van personages) van deze romans en in welke mate bemoeilijkt een verschil in leeshouding onder de participanten deze generaliseerbaarheid?

Relevante deelvragen zijn: welke aspecten van de roman zijn leidend bij de interpretatie en beoordeling van die roman door verschillende lezers en door ‘de computer’, en in hoeverre zijn lezers bereid om hun lezing op andere kenmerken te beoordelen? In wat volgt doe ik een voorstel voor een methode en een technologie die antwoorden kunnen geven op deze vragen en probleemstelling. Vervolgens problematiseer ik deze methode op basis van de theoretische traditie van de hermeneutiek en de digital humanities. Die theoretische reflectie levert enkele hypothesen op, die ik zal testen in een analyse van drie romans. De theoretische reflectie en de analyse geven uiteindelijk antwoord op de algemene probleemstelling.

3. Methode

De wending naar een meer kwantitatieve literatuurbenadering is niet nieuw en kent inmiddels een eigen traditie. Een van de bronnen van deze methode ligt in het boekje *Graphs, Maps, Trees. Abstract Models for Literary History* (2005) van Franco Moretti. Moretti werkt zijn methodologische positie nauwelijks uit, maar erkent wel dat vergelijkbare vormen van wetenschappelijke vernieuwing al eerder plaatsvonden. In tegenstelling tot collega's die zich lieten inspireren door de Franse en Duitse 'metaphysics' (die woordkeuze verwijst ook naar Derrida's 'metafysica', maar de retoriek zal weinigen ontgaan), vindt Moretti dat de literatuurwetenschap veel kan leren van de sociale- en natuurwetenschappen (Moretti 2009: 2). Het is daarom niet toevallig dat hij zich positioneert met verschillende verwijzingen naar de *Annales*-historici, die in de jaren dertig furore maakten met een sociaaleconomische geschiedwetenschap die zich toelegde op meetbaarheid, objectiviteit en die zich afkeerde van de Grote Mannen-geschiedenis zoals die daarvoor werd bedreven (Moretti 2009: 3-4). De beperkte literatuurwetenschappelijke blik op slechts enkele honderden canonieke romans, een fractie van de totale boekproductie, frustrereert Moretti, zoals de historici van *Annales* de consensus loslieten over wie of wat het bestuderen waard was in een bepaalde periode. Moretti propageert dus niet alleen een methode zelf, maar ook een richtlijn voor een meer inclusieve materiaalselectie: voorbij de impliciete poëtische normen over literaire kwaliteit.

De waarde van Moretti's betoog is niet zozeer dat hij een nieuwe, vruchtbare methode introduceerde, maar bovenal maakt hij kernachtig duidelijk wat de verschillen zijn tussen kwantitatieve en kwalitatieve informatie voor de boekwetenschap in het bijzonder en de literatuurwetenschap in het algemeen. 'Distance' als het tegendeel van 'close reading' is niet een obstakel, maar 'a specific form of knowledge: fewer elements, hence a sharper sense of their overall interconnection' (Moretti 2009: 2). Daarnaast beschouwt Moretti de literatuur niet als 'a sum of individual cases', maar als 'collective system', en dat vraagt om een andere benadering: gegevens over het eerste zeggen weinig over het laatste, en vice versa. Tot slot is kwantitatieve data idealiter onafhankelijk van interpretatie volgens Moretti (2009: 9). Dat betekent dat dezelfde data aanleiding zouden kunnen geven tot verschillende interpretaties door verschillende onderzoekers. Het betekent ook dat 'data' neutraal worden en niet langer slechts betekenis hebben in de specifieke uitleg van één individuele onderzoeker, met alle mogelijke vormen van vooringenomenheid van dien.

Tegelijkertijd ontstaat er een paradox wanneer distant reading, het oppervlakkig lezen van teksten, ingezet wordt bij ideologiekritische lezingen of interpretaties die rekening willen houden met verborgen lagen van de tekst. De dominante opvatting binnen verschillende ideologiekritische tradities is namelijk dat die lagen zich bevinden *onder* het oppervlak van de tekst: 'Broadly speaking, this practice encompasses an interpretive method that argues that the most interesting aspect of a tekst is what it represses' (Best en Marcus 2009: 3). Het is de aloude Freudiaanse tegenstelling tussen latent en manifest die leidend is voor dergelijke critici. De aanname van Marxisten zoals Fredric Jameson of postkoloniale critici zoals Toni Morrison en Edward Said is evenwel dat het latente, het ideologische, zich openbaart op het manifeste, materiële niveau. Die manifeste oppervlakte wordt zowel gevormd door wat de tekst zegt als wat de tekst open laat – de stiltes, de afwezigheid van tekst, zijn feitelijk ook materieel. Dat rechtvaardigt dus het lezen van teksten voor dit doeleinde: ze verschaffen toegang tot waar het deze lezers om gaat, de sub-tekst. Het is echter nog maar de vraag of deze aanpak ook werkt wanneer er honderden of duizenden teksten gelezen worden. Alsof ze allemaal sleutels zijn van één en dezelfde deur die leidt naar een groot, amorf domein dat we 'ideologie' noemen. We gooien daarmee niet alleen alle teksten op één hoop, maar ook de uiteenlopende wereldbeelden die ze voortbrachten.

Ik ga dit probleem in dit onderzoek niet oplossen, maar wil het in plaats daarvan omarmen als een methodische vraag waarmee mijn scriptie experimenteert. Het onderzoek is een uitwerking van het toekomstbeeld dat Best & Marcus voor ogen stond toen zij schreven over de methodologische uitdaging

van wat zij ‘surface reading’ noemen, een manier van lezen die in haar afkeer van selectie (van zinnen, van teksten) lijkt op Moretti’s distant reading (Best & Marcus 2009: 16-17). Surface reading is de ideologiekritische pendant van distant reading. En de computer kan bij de inzet van surface reading een centrale rol spelen. Ik citeer uitgebreider:

‘We have in mind here, first, the recent turn toward computers, databases, and other forms of machine intelligence across a range of fields and practices, from book history to distant reading. Where the heroic critic corrects the text, a nonheroic critic might aim instead to correct for her critical subjectivity, by using machines to bypass it, in the hopes that doing so will produce more accurate knowledge about texts. Where the heroic text commands admiration of its unique features, computers can help us to find features that texts have in common in ways that our brains alone cannot. Computers are weak interpreters but potent describers, anatomizers, taxonomists.’ (Best & Marcus 2009: 17)

Hoe ziet zo’n inzet van computertechnologie of distant reading er in de praktijk uit? Hoe elimineer je ‘subjectiviteit’ in een onderzoek om zo de ideologiekritische claims overtuigender te maken? Een vaak beproefde strategie om verschillende ideologische interpretaties te verenigen, is om ongelijkheid of representatienormen te kwantificeren. Zo’n aanpak kiest bijvoorbeeld de eerdergenoemde *Lezeres des Vaderlands*, die wekelijks alle gerecenseerde en recenserende vrouwen in de belangrijkste boekenbijlagen telt en zo op haar blog de achtergestelde positie van vrouwen in de literaire wereld aantoont met cijfers. Een ander voorbeeld is de website voor datavisualisaties en data-narratieven Polygraph, waarop een analyse verscheen van de verhouding tussen mannen en vrouwen in alle dialogen van 2000 Hollywood-films opgesplitst per genre (Anderson & Daniels 2016), opnieuw om zo genderongelijkheid op het witte doek aan de kaak te stellen. Een derde voorbeeld is een studie van Van der Deijl et al. (2016), waarin representatienormen in 170 Nederlandstalige romans werden aangetoond door alle personages te categoriseren volgens enkele sociale variabelen. Al deze initiatieven zetten cijfers in om ideologiekritische analyses van de literaire wereld, de filmwereld of de literatuur zelf te staven. Ze impliceren daarmee dus dat kwantitatieve informatie een andere epistemologische validiteit en mogelijk een grotere retorische waarde heeft dan kwalitatieve informatie.

Een probleem van de bovengenoemde benaderingen is echter dat de data nog altijd voortkomen uit de leeshouding van slechts één of enkele onderzoekers. In de gegeven voorbeelden gaat het steeds om getrainde lezers, studenten of intellectuelen wier leeshouding duidelijk beïnvloed is door een academische scholing. De geschiedenis van het autonomistische literatuurideaal in Nederland leert dat critici de academische leeshouding met argusogen bekijken en haar argumenten gemakkelijk terzijde schuiven. Idealiter zouden de gegevens over diversiteit in de literatuur dus behalve kwantificeerbaar ook afkomstig moeten zijn van grotere groepen (niet-academische) lezers. Daarom leent dit probleem zich voor ‘crowdsourcing’: het uitbesteden van bepaalde taken aan een (online) gemeenschap via een vrije oproep (Brabham 2013: 2; Howe 2006). Ik beperk me daarmee tot onderzoek naar diversiteit in de literatuur zelf, als aanvulling op de pilotstudie van Van der Deijl et al. 2016.

3.1 Crowdsourcing en Web 2.0

De wetenschap ontdekt de laatste jaren in toenemende mate de ongekende mogelijkheden die crowdsourcing biedt, ook in Nederland.¹ Wie zich met crowdsourcing inlaat, moet zich echter bewust zijn van de historische achtergrond en de ethiek van deze methode. In deze paragraaf wil ik kort ingaan op de bredere technologische ontwikkeling waar crowdsourcing een onderdeel van is. Op basis daarvan zal ik

¹ Zo organiseerde KNAW op 11 juni 2016 een symposium over Citizen Science en verzamelde het instituut de wensen en ervaringen van de vele Nederlandse onderzoekers die met crowdsourcing werken.

enkele ethische en technologische criteria formuleren die leidend zijn bij het ontwerp van de applicatie die de daadwerkelijke interactie met de ‘crowd’ moet faciliteren.

De crowdsourcing-beweging speelt in op een ontwikkeling van het internet naar een stadium dat met ‘Web 2.0’ wordt aangeduid. De term verwijst naar een historische fase van het world wide web, dat een fundamentele transformatie onderging rond de dotcombubbel. In plaats van een statisch communicatiemedium werd het web een platform dat gebruikers uitnodigt te participeren (Facebook, Twitter en andere social media) en gebruikersgedrag inzet om de eigen diensten te verbeteren (Google, Amazon en feitelijk alle diensten die gebruikmaken van ‘cookies’):

‘The bottom line is this: The immutable, standalone Web site is dead. Say hello to a Web that increasingly looks like a library full of chatty components that interact and talk to one another. Increasingly, people are engineering software, databases, and Web sites so that they not only meet private objectives, but so that they can be used in ways the originators did not know or intend.’
(Tapscott & Williams 2008: 38; vgl. O’Reilly 2005)

In de geest van Web 2.0 zal dit onderzoek een dergelijke database inzetten, die niet alleen mijn eigen vragen beantwoordt, maar aanleiding zou moeten geven tot een interactieve dialoog met de data.

Een veelgehoorde analyse is dat het Web 2.0 niet alleen de menselijke omgang met technologie heeft veranderd, maar ook ingrijpende gevolgen had voor de hedendaagse economie en politiek. Tapscott en Williams schreven een boek met de veelzeggende titel *Wikinomics. How Mass Collaboration Changes Everything* (2008). Zij spreken van een ‘collaboration economy’ in het tijdperk na de dotcombubbel, waarin bedrijven hun waardevolle data gaan delen met het publiek om tot nieuwe oplossingen te komen. (Tapscott & Williams 2008: 55) Een bekend voorbeeld is het Canadese mijnbouwbedrijf Goldcorp, dat in 2000 met overweldigend succes besloot zijn waardevolle geologische data vrij te geven en een prijs beloofde aan degene die hen kon vertellen op welke plek zij nieuwe goudbronnen konden vinden. Het is een beeldend voorbeeld van een economie waarin de grenzen tussen producerende industrie en een consumerende markt vervagen omdat het intellectuele kapitaal van de consument onderdeel wordt van die industrie.

Het risico bestaat echter dat de industrie altijd meer profiteert van deze transactie dan het individu. Crowdsourcing en citizen science maken ook deel uit van deze collaboration economy. Een manier om die ongelijkheid deels weg te nemen is om vrijwilligers in crowdsourcing projecten een (gedeelde) eigendomsrecht toe te kennen, tenzij zij daarvoor gecompenseerd worden (vgl. Riesch & Potter 2014: 8). Daarnaast moet de wetenschap die data vrij toegankelijk maken en is het onetisch om de data te verkopen zonder de vrijwilligers te laten delen in de winst.

Het potentieel van een gedecentreerde inzet van intellectueel kapitaal had een bedrijf als Goldcorp afgekeken bij de open source beweging. In de jaren 1980 en 1990 keerde deze gemeenschap van ontwikkelaars zich om idealistische of antikapitalistische motieven tegen de hegemonie van IBM en Microsoft. Gezamenlijk en zonder winstoogmerk ontwikkelden zij Linux, Ubuntu en andere open source software (Benkler 2002). Wikipedia was vervolgens een van de eerste organisaties die liet zien hoe krachtig de ‘wisdom of the crowd’ ook buiten de IT-sector kan zijn. Dergelijke bewegingen gaven aanvankelijk aanleiding tot een optimisme over de disruptieve en zelfs revolutionaire potentie van het internet. Het world wide web kon de kennis, het kapitaal en de technologie die voorheen in handen waren van overheden, universiteiten en technologiebedrijven, decentreren en toegankelijk maken voor iedereen die beschikt over een internetverbinding (de geografische en sociaaleconomische begrenzing van dit idealistische narratief moge daarmee duidelijk zijn). Deels volgt mijn onderzoek dit idealistische narratief omdat ook ik geloof dat het internet in brede zin een alternatief kan bieden aan het symbolische kapitaal en de vanzelfsprekende autoriteit van literaire instituties in de vorm van blogs, crowdfunding en goedkope promotiekanalen via social media. Specifiek voor dit onderzoek ligt het potentieel in de collectieve interpretatie van lezers die dankzij het internet mogelijk kan concurreren met die van de gevestigde critici.

Tegelijkertijd moet en wil ik kritisch zijn op de technologie die dit onderzoek inzet, want ook crowdsourcing is geen heilige graal. Het democratische ideaal van het internet ten spijt is er in de laatste jaren opnieuw een centrering van technologische en intellectuele macht zichtbaar. Recentelijk hebben bedrijven een business model ontdekt waarbij consumenten op grote schaal geld kunnen verdienen met hun materiële kapitaal, waartoe de industrie slechts nog de infrastructuur levert. Dergelijke bedrijven stimuleren de zogenaamde deeleconomie, met Über en Airbnb als bekendste succesverhalen. De keerzijde van die economie is echter dat ze een vorm van ongeregistreerd ondernemerschap faciliteert dat door middel van prijsverlaging moet concurreren met de gevestigde hotel- en taxisector. Die concurrentie gaat niet alleen ten koste van de omzet en werkgelegenheid bij de geregistreerde bedrijven, maar is bovendien uitsluitend mogelijk omdat zelfstandige ondernemers geen sociale verzekering (kunnen) afsluiten en geen belasting betalen. In zo'n economie functioneert de term 'delen' vooral als een retorische constructie, want het grootste stuk van de taart gaat uiteindelijk naar de IT'ers in Silicon Valley, ten koste van zowel Überchauffeurs als de officiële taxibestuurders. Deze werkelijkheid staat ogenschijnlijk ver af van zogenaamde *citizen science* en de inzet van crowdsourcing in de wetenschap, maar wie crowdsourcing als een veelbelovende methode presenteert, moet zich ook bewust tonen van het feit dat crowdsourcing niets meer is dan het uitbesteden van onderzoekswerk aan onbetaalde krachten (Riesch & Potter 2014: 11).

Op basis van deze historische en ethische reflectie stel ik een zestal technische en ethische criteria op die het ontwerp van de applicatie zullen bepalen:

1. De technologie volgt een compromisloos open data-beleid, waarbij de database op ieder moment opvraagbaar moet zijn in .xml (leesbaar in Excel) en voorzien van een duidelijke documentatie.
2. De applicatie moet de data voor de gebruiker interpreteren (*explanatory*) en tegelijkertijd ruimte laten voor eigen interpretaties en analyses van de gebruiker (*exploratory*).
3. Participanten bezitten een aandeel in de eigendomsrechten van de data.
4. De tijdsinvestering van de participant moet tot een minimum beperkt worden.
5. Participeren is volledig vrijwillig en uit eigen beweging, conform het 'open call'-principe in de crowdsourcing-definitie van Howe.
6. Persoonlijke gegevens of metadata over gebruikersgedrag (cookies) worden op geen enkele manier opgeslagen, laat staan verkocht.

3.2 Technologie

Ik heb een web-applicatie ontwikkeld waarmee lezers gegevens over personages in een gestandaardiseerde vorm kunnen aanleveren, de database kunnen bevragen en ordenen, en actuele verhoudingen in de database kunnen zien. Deze applicatie heb ik 'Personagebank' gedoopt en draait op een server van de faculteit Geesteswetenschappen van de Universiteit Utrecht (www.personagebank.hum.uu.nl). De applicatie heeft in overeenstemming met de zes criteria twee kernfuncties in de interactie met gebruikers:

1. Dataharvesting (lezers leveren gegevens aan)
2. Datapublicatie (lezers kunnen de database bevragen en interpreteren)

De applicatie valt uiteen in drie componenten:

1. Een database met daarin gegevens over personages en metadata over romans.
2. Een web-interface die de informatie uit de database visueel weergeeft, gebruikers toegang verschaft tot die informatie en hen informatie laat opslaan.
3. Een softwareprogramma dat draait op de webserver en dat de interactie tussen de web-interface en de database reguleert.

Hieronder licht ik beknopt de functie van deze onderdelen toe.

De database

De database is niets meer dan één XML-bestand dat de informatie organiseert en markeert via een vaststaande, hiërarchische structuur. XML staat voor ‘Extensible Markup Language’ en kent als belangrijk voordeel dat de gebruiker zelf de semantische tags kan definiëren. Dit aspect onderscheidt het van de verwante basistaal van het internet, HTML, die steunt op een vaste set tags waaronder bijvoorbeeld ‘<body>’ en ‘<div>’. Voor de Personagebank betekent het bijvoorbeeld dat auteursgegevens als volgt gemarkeerd worden:

```
<auteur num="1">
  <auteur_naam>Haase, Hella</auteur_naam>
  <auteur_gender>Vrouw</auteur_gender>
  <auteur_geboortejaar>1918</auteur_geboortejaar>
</auteur>
```

Door deze gegevens in te kapselen in html-tags die een browser wel kan interpreteren, is het mogelijk om in XML gemarkeerde informatie als html-content te tonen. Deze mogelijkheid om gegevens te markeren en semantisch te verrijken biedt naar mijn idee voordelen ten opzichte van andere database-designs, hoewel het meer gangbare ontwerp van een ‘relational database’ (georganiseerd in rijen en kolommen in plaats van de hiërarchische structuur van XML) voor dit doeleinde evengoed werkbaar was geweest. De keuze voor XML en de bijbehorende technieken XPath en XQuery in plaats van een relational database zoals MySQL in combinatie met PHP en SQL is in zekere mate ook pragmatisch geweest: mijn technische expertise beperkte zich tot de eerstgenoemde technieken. Zie bijlage 1 voor een voorbeeldlemma uit de database.

De interface

De term ‘interface’ verwijst naar een digitale omgeving waarin de gebruiker achterliggende scripts in gang kan zetten zonder zelf die scripts te hoeven schrijven. Ook toont de interface de gebruiker een (interactieve) representatie van de statische data in de database. De interface is geschreven in de standaard web-combinatie html, CSS en JavaScript.

De eerste functie van het platform (dataharvesting) wordt in gang gezet als de gebruiker een ISBN invoert. De interface checkt allereerst of de database al een lemma met dat ISBN kent. Als dat zo is, verschijnen de metadata en ingevulde personages uit de database op de webpagina en wordt een formulier gegenereerd waarmee de gebruiker extra personages kan toevoegen aan het bestaande lemma. Als het ISBN nog niet bekend is, raadpleegt de interface de catalogus van de Koninklijke Bibliotheek door middel van een zogenaamde API (Application Programming Interface). API’s zijn programma’s die communicatie tussen verschillende software of datasets mogelijk maken aan de serverzijde. Via JSRU, een vraagtaal ontwikkeld voor de API van de KB, wordt het lemma met het ingevulde ISBN uit de catalogus van de KB opgevraagd. Aan dat lemma (in XML) worden via XQuery enkele gegevens onttrokken:

Metadata

1. Naam van de auteur
2. Titel
3. Uitgever
4. Jaar van uitgave
5. Jaar van de eerste druk
6. Originele taal

7. Categorie
8. URL naar het lemma in de KB-catalogus
9. ID-nummer van de roman

Informatie 1 tot en met 5 toont de interface direct op de webpagina bij een succesvolle response. Onderdelen 5, 6 en 7 dienen als checks: als het opgevraagde boek ouder is dan 1945, niet oorspronkelijk in het Nederlands geschreven is of niet valt onder ‘romans en novellen’, dan geeft de interface een foutmelding en verschijnt het invulformulier niet. Onderdeel 8 wordt slechts voor het gemak van de beheerder opgeslagen, zodat informatie gemakkelijk gecheckt kan worden wanneer die eenmaal in de database is opgenomen. Als de catalogus het ingevoerde ISBN niet kent en dus geen respons geeft, krijgt de gebruiker de mogelijkheid om via een klein formulier bibliografische metadata zelf in te voeren.

Indien via de API van de KB gebleken is dat het ingevulde ISBN naar een Nederlandstalige roman van na 1945 verwijst, verschijnt er een invulformulier waarin de gebruiker per personage waardes kan kiezen of invullen voor tien variabelen door middel van invulvelden of selectmenu’s (‘dropdown’). Variabele 11 wordt automatisch gegenereerd door middel van de zogenaamde GeoCoding API van Google Maps, waarmee plaatsnamen worden vertaald naar geografische coördinaten.

Data

1. Naam (invul)
2. Rol (select)
3. Gender (select)
4. Leeftijd (select)
5. Herkomst (select)
6. Woonplaats land (select)
7. Woonplaats (invul)
8. Woonplaats type (select)
9. Beroep (invul)
10. Opleidingsniveau (select)
11. Geografische coördinaten (GeoCoding API)
12. ID-nummer van het personage (automatisch via XQuery)

Als de gebruiker alle velden voor alle personages heeft ingevuld en op ‘Verder’ klikt, wordt de informatie nog niet naar de server verstuurd, maar vindt er eerst een controle plaats aan de kant van de ‘client’. Een script in JavaScript checkt alle velden en geeft een foutmelding als een daarvan leeg is of ongewenste inhoud bevat. Als alle inhoud aan de eisen voldoet, verschijnt er een tabel boven het formulier met daarin de zojuist ingevulde informatie en een ‘Opslaan’-knop. De gebruiker kan de gegevens een laatste keer controleren voordat zij op ‘opslaan’ klikt. Bij het indrukken van die knop wordt zowel de metadata als alle personagedata naar de server verstuurd en opgeslagen in de database. Tijdens dat proces vertaalt de interface voor elk personage de woonplaatsen naar geografische coördinaten via de GeoCoding API van Google Maps en voegt die informatie toe aan elk personagelemma.

De tweede functie van de applicatie (datapublicatie) kent verschillende vormen. Om te beginnen kunnen gebruikers een lijst van alle personages en een lijst van alle romans opvragen, waarop de beschikbare informatie getoond wordt.² Alle pagina’s van de website bevatten rechtsboven een link naar beide lijsten. Door middel van sorteerfuncties is het mogelijk om de lijst te sorteren op alle categorieën (bijvoorbeeld op jaar van publicatie, uitgever, naam van de auteur etc.). Per lemma kan de gebruiker doorklikken naar een totaaloverzicht van alle metadata en personagedata van één roman uit de database.

² <http://personagebank.hum.uu.nl/search-4.xq> en <http://personagebank.hum.uu.nl/search-3.xq>.

Die pagina wordt verrijkt met het boekomslag. Die afbeelding staat niet op de eigen server, maar wordt geleverd vanaf de boekensite Goodreads, via de API van het sociale medium en op basis van het ISBN. Elk totaaloverzicht bevat ook de knop ‘Voeg personage toe’, waarmee de gebruiker meteen doorverwezen wordt naar de invul-interface voor het betreffende boek (de ‘Doe mee’-pagina).

Naast deze totaaloverzichten kent de website een algemene ‘Resultaten’-pagina.³ Deze pagina bevat een zoekfunctie opgesplitst in ‘zoeken in romans’ en ‘zoeken in personages’, waarbij de zoekfunctie voor romans enkele aanvullende filters biedt (titel, uitgever etc.). Onder deze zoekfuncties toont de pagina actuele verhoudingen onder de personages in de database door middel van verschillende interactieve datavisualisaties en korte toelichtingen. Deze pagina is ‘actief’, dat wil zeggen dat de content varieert en automatisch gegenereerd wordt op het moment dat een gebruiker de pagina opvraagt in de browser. Zodra een gebruiker een roman aan de database toevoegt, verspringen de grafieken nadat de pagina vernieuwd wordt. Voor de interactieve grafieken en voor een van de kaarten maak ik gebruik van de JavaScript-library Amchart, een gebruiksvriendelijke collectie van scripts voor datavisualisaties op het web. Voor de gedetailleerdere woonplaatsen-kaart gebruik ik een kaart van Google Maps, die ik op de pagina integreer met de JavaScript library van Google Maps. Tot slot is het blog ook onderdeel van de data-publicatie functie. Ik gebruik het blog om verbanden tussen variabelen te duiden voor participanten en andere geïnteresseerden en illustreer die met kaarten, tabellen en grafieken.⁴

De software

Voor het opvragen, analyseren en transformeren van data uit de database gebruik ik de XML-vraagtaal XQuery. Dit is een scripttaal die een XML-database kan bevragen en op basis van de response XML of HTML-data genereert en naar de browser stuurt. De XQuery-scripts draaien aan de serverzijde via speciale software. Dat betekent dat gebruikers (aan de ‘client-side’) deze scripts niet kunnen zien. Voor het verwerken van de scripts gebruik ik de open source databasesoftware eXist DB. De software is een totaalpakket dat als zelfstandige server kan functioneren (handig voor de ontwikkelfase op een lokale machine), maar ook geïmplementeerd kan worden in een bestaande serveromgeving voor productie-doeleinden, zoals in dit geval op de servers van de universiteit. eXist biedt een handige interface, onder andere voor het handmatig uploaden van bestanden, voor het monitoren van de prestaties (snelheid, geheugenbelasting), voor configuraties van de URL en voor user- en licentiemanagement (de instellingen over wie bestanden kan zien, draaien en bewerken). Daarnaast levert eXist een code editor die toegespitst is op het schrijven van XML en XQuery. Via deze editor, ‘Exide’, is onderhoud en updaten van de webpagina’s mogelijk. De uiteindelijke deployment van de software en domeinnaam werd gedaan door personeel van de Universiteit Utrecht, in het bijzonder door ICT-beheerder Chris Rauws.

³ <http://personagebank.hum.uu.nl/resultaten.xq>

⁴ <http://personagebank.hum.uu.nl/blog.xq>

4. Theoretisch kader

De methode heeft tot doel om via collectieve, geformaliseerde interpretaties van een specifiek narratologisch aspect van de roman, uitspraken te doen over meerdere romans en over leeshoudingen van meerdere lezers. Deze doelstelling roept uiteenlopende theoretische vragen op en doet enkele problematische aannames over wat lezen, literatuur en literatuurwetenschap is of zou moeten zijn. Zo zeg ik rekenschap te willen geven van verschillende leeshoudingen, maar dwing ik participanten tegelijkertijd om hun interpretaties en lezersreacties te reduceren tot antwoorden op een formulier van tien (meerkeuze-)vragen. Vervolgens veronderstel ik mogelijk onterecht dat de antwoorden op die vragen voor elke roman een accurate graadmeter vormen voor de manier waarop personages gerepresenteerd worden. Als die antwoorden immers betekenisloos zijn voor bepaalde romans, zijn generaliserende uitspraken op basis van die formulieren misleidend. Daarnaast staat het formulier andere input dan de gevraagde informatie over de personages niet toe, hoewel personages slechts een van de vele betekenisvolle aspecten zijn in vertellingen. Het is denkbaar dat een roman geen enkel personage opvoert of voornamelijk via andere middelen een narratief overbrengt, waardoor een karakterisering van de personages als enige lezersreactie kunstmatig of geforceerd zou zijn. Tot slot speelt er nog een epistemologisch probleem dat inherent is aan de digitale bemiddeling van deze methode: de aanname dat het zoeken naar patronen meer betekenis of meer legitieme kennis voortbrengt dan het analyseren van de afwijking.

In dit hoofdstuk zal ik deze aannames in meer detail problematiseren. Het theoretisch kader dat zo vorm krijgt, stelt mij in de gelegenheid om enkele hypothesen te formuleren over de functie en waarde van de collectieve en geformaliseerde lezersreacties die ik met crowdsourcing hoop te verkrijgen. Hieronder bespreek ik de verschillende rollen die auteur, tekst, lezer en context kregen in enkele relevante bijdragen aan de hermeneutische discussie van de laatste zestig jaar. Het is daarbij allerminst mijn doel om een uitputtende bespreking van deze traditie te geven. Ik schets slechts enkele posities die ik in mijn ogen in overweging moet nemen om te begrijpen welke specifieke manieren van interpreteren dit project veronderstelt en wat de consequenties daarvan zijn.

4.1 Een hermeneutiek van crowdsourcing

Het is een klassieke vraag binnen de literatuurwetenschap: hoe werken lezer, tekst en context samen in het proces van betekenisproductie dat plaatsvindt wanneer een literaire tekst gelezen wordt? Voor dit onderzoek is deze vraag opnieuw actueel, omdat de applicatie interpretaties van verschillende lezers over verschillende teksten samenvoegt en vervolgens het geheel van die interpretaties opnieuw (digitaal) interpreteert. Die pretentie tart de huidige hermeneutische theorievorming, die doorgaans uitging van één tekst of één (menselijke) lezer (doorgaans de onderzoeker), of vaker nog van abstracties als ‘de tekst’ of ‘de lezer’ of de ‘impliciete lezer’ (vergelijk ook Pieterse 2008: 18; Felski 2008: 16). Om te beginnen bespreek ik de verschillende posities in dit debat over de hermeneutische oorsprong van betekenis. Ik signaleer dat het zwaartepunt historisch gezien geleidelijk verschuift van context naar tekst naar lezer, om uiteindelijk een compromis te bereiken tussen de laatste twee.

Het ligt voor de hand om te wijzen naar de tekst als de bron van literaire betekenis. Die uitleg volgt immers de temporele orde die we ervaren wanneer we lezen: de lezer neemt een boek ter hand, begint te lezen, en permitteert zich – in een gunstig geval – een vorm van reflectie op de betekenis van de gelezen tekst. De reflectie, de interpretatie, gaat niet vooraf aan de tekst, noch aan het lezen. De logische gevolgtrekking luidt dat de tekst de betekenisbron is. Maar strikt genomen is er nog een vroegere oorsprong van literaire communicatie. Teksten ontstaan immers niet spontaan, maar hebben altijd een maker. In de vroege twintigste eeuw wendden critici en literatuurwetenschappers zich daarom tot de

schrijver als de ultieme bron van literaire communicatie. Wimsatt & Beardsley herkenden in hun beroemde essay 'The Intentional Fallacy' (1946) twee centrale vragen in deze vorm van kunstkritiek: '(1) whether the artist achieved his intentions; (2) whether the work of art "ought ever to have been undertaken at all" and so "whether it is worth preserving"' (Wimsatt & Beardsley 1946: 471). Ze werpen tegen dat vraag (1) onmogelijk te beantwoorden is wanneer uitsluitend de tekst object van kritiek is of kan zijn. (2) geldt voor hen wel als een zinnige vraag, maar alleen wanneer het antwoord op vraag 1 daarop niet van invloed is. Critici dienen zich volgens deze formalisten louter op de tekst te baseren voor hun interpretatie of evaluatie, want de auteursintentie is niet toegankelijk en lezersreacties zijn te subjectief en dus instabiel. Het essay van Wimsatt & Beardsley markeert kortom een vroege formulering van een positie die stelt dat de tekst het enige stabiele en toegankelijke referentiepunt is voor de interpretatieve activiteit (vgl. Fish 1980: 2).

Een vergelijkbare en beroemde kritiek op de hermeneutiek van de vooroorlogse literatuurbeschouwing komt van de Amerikaanse schrijver en critica Susan Sontag. In haar essay 'Against Interpretation' (1961) breekt ze retorisch een lans voor de algehele afschaffing van interpretatie, een reactie op kunst die zij reductionistisch en inadequaar vindt. Sontag reageert daarmee voornamelijk op het Freudiaanse en Marxistische interpreteren, waarin iedere tekst vooral als een symptoom gezien wordt van een van deze 'contextuele' modellen van de werkelijkheid. 'Against Interpretation' adresseert kortom de critici die teksten slechts lezen als een weerspiegeling van de psychoanalytische driften van de auteur of van de materialistische ongelijkheden die zich in de tekst nestelden. Sontag ziet de neiging tot interpretatie in de kunstkritiek als een gevolg van de mimetische kunstopvatting die volgens haar de volledige Westerse esthetica na Plato heeft beheerst (Sontag 2009: 4). De esthetica beschouwt mimesis dus als een inherente en onvermijdelijke eigenschap: kunst is altijd een afspiegeling van iets anders en het is aan de kunstcriticus om dat 'iets anders' te zien achter de tekst, het beeld, de muziek: 'Directed to art, interpretation means plucking a set of elements (the X, the Y, the Z, and so forth) from the whole work. The task of interpretation is virtually one of translation. The interpreter says, Look, don't you see that X is really – or, really means – A? That Y is really B? That Z is really C?' (Sontag 2009: 5). De kunstkritiek heeft zo vorm gescheiden van inhoud en het eerste ondergeschikt en instrumenteel gemaakt:

'Interpretation in our own time, however, is even more complex. For the contemporary zeal for the project of interpretation is often prompted not by piety toward the troublesome text (which may conceal an aggression), but by an open aggressiveness, an overt contempt for appearances. The old style of interpretation was insistent, but respectful; it erected another meaning on top of the literal one. The modern style of interpretation excavates, and as it excavates, destroys; it digs "behind" the text, to find a sub-text which is the true one (Sontag 2009: 6).

Volgens Sontag leidt deze lezersreactie tot een reductie en zelfs tot een 'vernietiging' van de tekst: de tekst wordt 'getemd' tot er alleen nog de inhoud over is (Sontag 2009: 8). Sontag pleit dan ook voor een herwaardering van vorm: 'In place of a hermeneutics we need an erotics of art' (Sontag 2009: 14). De obsessie met de inhoud moet plaatsmaken voor een louter vormelijke, 'lichamelijke' ervaring van kunst. Sontag kent dus ook de grootste autoriteit toe aan de tekst en herwaardeert het vormelijke. Met haar oproep tot een andere benadering van kunst dan de gangbare, erkent ze echter ook impliciet dat bepaalde lezersreacties meer legitiem zijn dan andere en dat het mogelijk is om een andere leeshouding aan te nemen. Bij Sontag is de lezer dus assertief maar blijft zij altijd dienstbaar aan de stabiele tekst.

In reactie op dergelijke tekstgerichte of formalistische benaderingen kreeg een nieuwe positie gestalte die juist de veelkantigheid van het lezen en de problematiek van alle vormen van communicatie – dus ook literaire – omarmde. Deze positie wordt wel de poststructuralistische genoemd, met Jacques Derrida als belangrijkste denker. Derrida wees op de hiërarchische tweedeling tussen spreken en schrijven die als een soort dogma een lange traditie in de Westerse filosofie gekend had (Derrida 1988: 5). In die traditie gold het schrift als een minder zuivere vorm van communicatie vanwege de inherente afwezigheid of

distantie van de ontvanger en het gebrek aan een vaste context (zowel linguïstisch – zoals grammatica – als ‘echt’ – in de vorm van een communicatieve situatie). Het spreken werd daarentegen wel gekenmerkt door aanwezigheid (van context, verwachtingen en feedback van de ontvanger etc.) en was daardoor volgens de taalfilosofie in staat tot succesvolle overdracht van ideeën, meningen, intenties. Derrida stemt in met de communicatieve belemmeringen van het schrift, maar beargumenteert vervolgens dat de onzekerheid en onbedoelde betekenisproductie van geschreven taal niet uniek is, maar integendeel ook orale communicatie problematiseert (Derrida 1988: 7). De implicatie is dat *alle* communicatie geteisterd wordt door verlies van context en van de originele, bedoelde betekenis (*vouloir dire*). Daaruit volgt dat het onderscheid tussen schrift en spraak niet langer houdbaar is, en bovenal dat het filosofische vertrouwen in talige communicatie schromelijk overschat is (Derrida 1988: 12). Voor de hermeneutiek betekent die conclusie dat teksten niet de stabiele betekenisbasis bieden die literatuurcritici en vooral de structuralisten veronderstelden. Wie Derrida streng leest, moet daarom vaststellen dat zoets als ‘literatuurwetenschap’ in deze visie nauwelijks nog bestaansrecht heeft. Een taalfilosofie die interpretaties in beginsel wantrouwt, ondermijnt een vakgebied dat kennis en symbolisch kapitaal (of wetenschappelijke autoriteit) wil produceren via de productie van interpretaties (van literaire teksten en de omstandigheden waaronder die ontstaan). We zouden met enig gevoel voor dramatiek kunnen spreken van een crisis in de hermeneutiek.

Het poststructuralisme kent echter vele criticasters die de consequenties van Derrida’s logica zien doorschieten in een literatuurwetenschappelijke ‘methode’. Ze vinden dat dit woord de systematiek die het belooft niet kan waarmaken (zie bijvoorbeeld Bertens 2015). Ze hebben bovendien weinig vertrouwen in de individuele zelfredzaamheid van de deconstruerende of tegendraadse lezer (die toch de interpretatieve autoriteit weer opeist) en verkiezen de autoriteit van de tekst boven de anarchie van het ‘anything goes’-devies. Een van die criticasters was de Italiaanse schrijver en literatuurwetenschapper Umberto Eco. In een lezingenreeks, gepubliceerd onder de titel *Interpretation and overinterpretation* (1992), verwoordt hij zijn standpunt ten aanzien van de poststructuralistische positie:

‘I was studying the dialectics between the rights of texts and the rights of their interpreters. I have the impression that, in the course of the last decades, the rights of the interpreters have been overstressed’ (Eco 1992: 23).

Volgens Eco determineert de tekst wel degelijk het aantal mogelijke of geldige interpretaties. Hij accepteert Derrida’s notie van de potentiële oneindigheid van betekenisproductie door een tekst, maar dat concept betekent volgens hem niet dat interpretatie haar object mag boetseren op een middelpuntvliedende draaischijf die van de originele vorm niets overlaat: ‘to say that a text has potentially no end does not mean that every act of interpretation can have a happy end.’ (Eco 1992: 24). Hij spreekt in dit verband van een ‘intention of the text’, een bepaald communicatief kader waar je als lezer simpelweg niet omheen kunt (Eco 1992: 25). Dat wil zeggen: je kunt ervoor kiezen om die tekstintentie te negeren – en je verliezen in de ‘overinterpretatie’ waar de titel naar verwijst – maar een dergelijke interpretatie zal volgens Eco niet standhouden in de gemeenschap van interpreten. Daarmee maakt Eco een belangrijk punt: de *geldigheid* van interpretaties wordt niet bepaald door de auteur, de tekst of de lezer, maar door de gemeenschap van lezers. En daar, in die sociale interactie, ligt uiteindelijk ook de autoriteit.

Stefan Collini schrijft in de inleiding van *Interpretation and Overinterpretation* dat Eco deze sociale onderhandeling van interpretatie voorstelt als een soort cultureel Darwinisme: ‘certain readings prove themselves over time to the satisfaction of the relevant community’ (Collini 1992: 16). Het is een proces dat doet denken aan Poppers falsificatiecriterium van wetenschappelijke theorie: een interpretatie is ‘waar’, geldig, voor de interpreterende gemeenschap tot het tegenbewijs geleverd is. Zodra meer lezers een tegenbewijs geven, verliest de interpretatie aan zeggingskracht. De individuele lezer heeft weinig autoriteit; die autoriteit ligt bij het collectief. De gemeenschap van lezers bestaat in deze uitleg dus uit een soort *lectores economicae* die overeenstemming kunnen bereiken over wat juiste lezingen zijn dankzij een gedeeld rationeel vermogen tot interpreteren.

Een tweede reactie op het poststructuralisme richt zich juist wel op de individuele lezer en neemt de diversiteit aan mogelijke lezersreacties serieus. Deze beweging wordt aangeduid als *reader response criticism* of receptie-esthetiek, met Wolfgang Iser als een van de grondleggers. Deze ‘fenomenologische’ benadering erkent dat literaire betekenis ontstaat in een ruimte waarvan de grenzen zowel door de tekst als door het leesproces bepaald worden (Iser 1988: 50). Deze virtuele ruimte noemt Iser het ‘Gestalt’ van de tekst, een derde dimensie die voortdurend transformeert tijdens het lezen (Iser 1988: 58). Isers benadering geeft dus rekenschap van het temporele aspect van interpretatie: betekenis ontstaat niet achteraf, maar wordt voortdurend bijgesteld en verrijkt tijdens de interactie tussen tekst en lezer. En daarin is ook voor de auteur nog een rol weggelegd:

‘As we read, we oscillate to a greater or lesser degree between the building and the breaking of illusions. In a process of trial and error, we organize and reorganize the various data offered us by the text. These are the given factors, the fixed points on which we base our “interpretation,” trying to fit them together in the way we think the author meant them to be fitted’ (Iser 1988: 62).

Volgens Iser is de lezers verwachting van wat de auteur bedoeld kan hebben dus leidend voor hoe zij de ‘gefixeerde’ tekstuele ‘data’ ervaart. Iser kiest dus een tussenpositie: zowel de tekst als de lezer is in zekere mate gefixeerd en voorspelbaar. Het is het samenspel tussen tekst en lezer dat de betekenis construeert. Iser spreekt dan ook van een opheffing van de subject-object dichotomie tussen lezer en tekst en doet terloops een uitspraak over het empathische potentieel van literaire teksten:

‘Text and reader no longer confront each other as object and subject, but instead the “division” takes place within the reader himself. In thinking the thoughts of another, his own individuality temporarily recedes into the background, since it is supplanted by these alien thoughts, which now become the theme on which his attention is focused’ (Iser 1988: 67).

Interpretatie wordt in deze uitleg dus sterk persoonsgebonden. De lezer spiegelt haar individuele realiteit aan die in de tekst. Zo mag het lezen weliswaar een samenspel tussen lezer en tekst heten; de invloed van de lezer is niettemin aanzienlijk.

In het verlengde van deze fenomenologische hermeneutiek staat Stanley Fish, die echter in zijn *Is there a text in this class* (1980) het zwaartepunt van de betekenisproductie nadrukkelijk niet bij de individuele lezer legt (zie ook Saïd 1983: 14-15). Net als Iser stelt Fish voor om de dichotomie tussen object en subject los te laten in de manier waarop we naar interpretatie kijken. De tekst gaat niet vooraf aan de lezer, noch bezit die tekst als zodanig enige stabiliteit als universeel referentiepunt in de betekenisproductie. Fish bekritiseert het formalisme omdat die de laatst verkregen interpretatie – na het lezen van de tekst – zou verkiezen boven de betekenissen die de lezer construeert tijdens het lezen: een temporele hermeneutiek zou de werkelijkheid meer recht doen dan een interpretatie die louter uitgaat van de ‘ruimtelijkheid’ van de tekst (Fish 1980: 3). Fish draait de chronologische orde zoals die doorgaans gezien wordt om: interpretatie is de ‘shape of reading’, en daardoor vormt de interpretatie de tekst, in plaats van andersom (Fish 1980: 13). Het potentieel van betekenis ligt dus eerder bij de lezer dan bij de tekst, maar Fish benadrukt dat de lezer daar geen bewuste keuzevrijheid in heeft. De manier waarop hij of zij interpreteert wordt bepaald door zogenaamde *interpretative communities*, gemeenschappen van lezers met dezelfde leeshoudingen. Verschillen tussen interpretaties moeten we volgens hem toeschrijven aan het feit dat die lezers tot verschillende interpretative communities behoren:

‘Indeed, it is interpretive communities, rather than either the text or the reader, that produce meanings and are responsible for the emergence of formal features. Interpretive communities are made up of those who share interpretive strategies not for reading but for writing texts, for constituting their properties. In other words these strategies exist prior to the act of reading and

therefore determine the shape of what is read rather than, as is usually assumed, the other way around' (Fish 1980: 14).

Deze aanname is behalve aannemelijk ook productief voor dit onderzoek, omdat ze zowel de stabiliteit van de tekst als de variatie in lezersreacties op die tekst theoretisch handhaaft. De stabiliteit van de tekst is weliswaar beperkt tot de interpretatieve gemeenschap waarin die gelezen wordt, maar binnen zo'n gemeenschap kennen tekstkenmerken wel degelijk een zekere bepaaldheid. Zo lang de participanten dus deel uitmaken van of toetreden tot dezelfde interpretatieve gemeenschap, is het mogelijk om tot gedeelde en valide kennis over de gelezen teksten te komen. Of nauwkeuriger: kennis over de betekenissen die collectief ontstaan tijdens het lezen van deze teksten.

De vraag rest dan wat die gemeenschappen en hun strategieën precies zijn en hoe die een betekenis vormen. Rita Felski geeft een mooie illustratie van zulke sociale leesmoden in haar essay 'After Suspicion' (2009). Ze beschrijft hoe haar studenten zich bewust worden van nieuwe houdingen ten aanzien van literatuur wanneer zij in de collegebanken in aanraking komen met literaire theorie (Felski 2009: 28). De kritische, 'wantrouwende' benadering van teksten die de soms activistisch georiënteerde theoretische traditie propageert, dwingt hen hun onbevengene of welwillende houding tegenover de kunsten te nuanceren of zelfs op te geven. Felski is echter ontevreden over de scheiding die de academische literatuurkritiek heeft aangebracht tussen hun eigen 'hermeneutics of suspicion' en de 'hermeneutics of admiration' waarmee de literatuur buiten de universiteitsmuren gewaardeerd wordt (Felski 2009: 29). Die scheiding doet volgens haar geen recht aan de veel pluriformere werkelijkheid, waarin academische en niet-academische lezers al dan niet bewust verschillende lezersreacties uit verschillende registers aanspreken. Ze stelt een hermeneutiek voor waarin ruimte is voor zowel de bewonderende, esthetische of onverschillige lezersreacties als voor ethische, kritische en tegendraadse houdingen, want: 'Such an insistence on the radical differences between interpretative communities disallows the possibility of overlapping modes of reading and shared cognitive and affective parameters' (Felski 2009: 30).

Felski accepteert dus dat er verschillende interpretatieve gemeenschappen bestaan, maar relateert tegelijkertijd de verschillen daartussen. In haar boek *Uses of literature* werkt ze deze aanname uit in vier categorieën lezersreacties, 'modes of textual engagement', die ze niet reserveert voor een bepaald type lezer (Felski 2008: 14). In plaats daarvan verwacht ze dat die lezersreacties voortkomen uit leeshoudingen die iedere lezer soms aanwendt tijdens het lezen, ongeacht haar doel of academische opleiding. Felski bekent daarmee kleur aan een fenomenologische opvatting van interpreteren, hoewel ze zich distantieert van de abstracte fenomenologie van Wolfgang Iser waarin de lezer 'bloodless and disembodied' is, 'stripped of all passions as well as of ethical or political commitments' (Felski 2008: 16). De elegantie van de fenomenologische benadering ziet zij in de persoonlijke, anti-universalistische implicatie van het idee dat de wereld slechts kenbaar is in de mate waarin die aan individuen 'verschijnt', gemedieerd door ons bewustzijn, perceptie en waardepatroon: 'We can learn to question our own beliefs; we can come to see that our seemingly spontaneous reactions are shaped by cultural pressures; we can acknowledge, in short, the historicity of our experience' (Felski 2008: 17). Felski benadrukt dus de cultureel-historische bepaaldheid van interpretaties van concrete lezers en theoretiseert hoe die interpretatie gevormd en bijgesteld wordt tijdens het lezen.

Er bestaat kortom een lange en rijke theoretische traditie over de werking van tekst en interpretatie. Verschillende theoretici nemen verschillende standpunten in, en bestrijden elkaar met argumenten van theoretische of poëtische aard. Die theorieën verdiepen ons begrip van wat lezen is en hoe interpretatie kan werken. Maar daar blijft het dan vaak bij. Op de empirische literatuurwetenschap na doen deze theoretici weinig moeite om hun theorieën te testen aan de werkelijkheid van de lezer, noch formuleren ze methoden om hun concepten te operationaliseren. De lezer, en feitelijk ook de tekst, zijn daardoor slechts abstracties. De theorieën klinken mogelijk alleen elegant omdat ze de schimmige en complexe realiteit van

de lezer buiten beschouwing laten. De theoretische relevantie van dit onderzoeksontwerp is dan ook dat het nadrukkelijk wel een beroep doet op ‘echte’ lezers. Door het verzamelen en formaliseren van interpretaties van tientallen en in potentie honderden of duizenden verschillende lezers, kan ik empirisch gefundeerde uitspraken doen over de stabiliteit van intersubjectieve interpretatie en daarmee over de betekenisstabiliteit van de tekst. Net als Fish behoud ik zo de theoretische mogelijkheid van de stabiele tekst zonder te ontkennen dat lezers verschillende leeshoudingen kunnen hebben.

De theoretische traditie dwingt niettemin ook tot een reflectie op mijn opzet. Mijn beroep op verschillende, echte lezers is namelijk ook de achilleshiel van dit onderzoek. Het standaardiseren van verschillende lezersreacties in een bepaald format geeft mogelijk een misleidende suggestie van bepaaldheid en van stabiliteit (zowel van teksten als van lezersreacties). Om beter zicht te krijgen op die mogelijke misleiding analyseer ik in hoofdstuk 5 een drietal romans uit de database, ingevuld door lezers. Ik beperk me in die analyse tot de representatie van personages. Het doel is namelijk niet om aan te tonen welke aspecten van de roman niet worden begrepen door de applicatie (die zijn immers eindeloos), maar om te begrijpen hoe de technologie representaties van personages interpreteert en hoe die interpretatie van dit specifieke verhaalaspect afwijkt van menselijke interpretaties. Als die afwijking duidelijk wordt, krijgt de ‘hermeneutische positie’ van deze vorm van crowdsourcen gestalte ten opzichte van de theoretische traditie.

4.2 Een epistemologie van crowdsourcing

De aanname dat een ‘hermeneutiek’ van het crowdsourcen – een theorie over de manieren van interpreteren die crowdsourcing aanspreekt en veronderstelt – zich kan verankeren in het traditionele denken over menselijk interpreteren vraagt in zichzelf een zekere ruimdenkendheid van de wetenschappelijke of interpretatieve gemeenschap die dit onderzoek aanspreekt. Deze aanname was legitiem geweest als ik als menselijke onderzoeker met alle lezers in contact stond en hun lezersreacties analoog verzamelde, interpreteerde en rapporteerde. Deze taken heb ik echter uitbesteed aan een webapplicatie, een ‘machine’. De kennis die dit onderzoek (hopelijk) oplevert, wordt zo op alle drie de niveaus (verzamelen, interpreteren en representeren van informatie) digitaal gemedieerd. Dat geeft de kennis die de applicatie voortbrengt mogelijk een andere status.⁵ Behalve een hermeneutiek dwingt dit onderzoek dus ook tot een ‘epistemologie’ van het crowdsourcen, hoewel de twee theorieën verwant zijn. De eerste gaat over de status van interpretatie in dit onderzoek, de tweede over de betekenis van digitale kennis in de letterkundige praktijk.

Om te kunnen spreken van een eigen ‘epistemologie’ – ik begrijp dat woord hier simpelweg als een theorie over wat valide kennis is en hoe die tot stand komt – moet ik twee vragen beantwoorden: is er een fundamenteel verschil tussen informatie over literatuur die menselijke lezers genereren en informatie over literatuur die een computer produceert, en zo ja: wat zijn daarvan de consequenties? Als ik verder abstraheer, gaat deze vraag in de kern over het verschil tussen menselijk en computationeel begrip. Wanneer die twee een samenwerking aangaan om een complex probleem op te lossen, zoals in dit onderzoek, veronderstellen we dat de computer op dezelfde manier of althans in essentie niet anders redeneert en communiceert dan de mens. Die veronderstelling lijkt echter moeilijk houdbaar als we in acht nemen dat mensen en computers begrijpen volgens verschillende modi (zoals metaforisch, associatief, verifiërend), media (mondeling, digitaal, papier) en eenheden (zoals afwijkingen, gemiddelden, patronen) van informatie en communicatie. Die verschillen eisen daarom een plaats op in een epistemologie van het crowdsourcen.

⁵ N.B. ‘Crowdsourcing’ gebeurt tegenwoordig vrijwel altijd digitaal, maar volgens mijn begrip van de term (zie pagina 14), is die modus niet per definitie onderdeel van het proces. De informatie die vergaard wordt door bijvoorbeeld referenda, prijsvragen of een uitzending van *Opsporing gezocht* wordt niet per se digitaal gemedieerd. Het zijn evengoed vormen van crowdsourcing.

Toch zijn er ook argumenten voor de stelling dat computers en mensen in essentie weinig verschillen. Volgens de Amerikaanse literatuurwetenschapper Katherine N. Hayles is er niet langer sprake van een fundamenteel onderscheid tussen menselijke computationele cultuur. In haar boek *My Mother Was a Computer* (2005) analyseert ze de verregaande menselijke interactie met 'intelligent machines' op drie domeinen: taal en code, print en 'etext', menselijk bewustzijn en digitale 'cognitie'. Ze demonstreert dat op al die domeinen het een niet langer losgezien kan worden van het ander. *My Mother Was a Computer* is een uitwerking van haar eerdere boek *How We Became Posthuman* (1999), waarin zij de 'feedback loops' tussen kunstmatige mechanismen en biologische organismen analyseert als een oorzaak van een zogenaamde 'crisis of agency', een herziening kortom van wat sinds het cartesiaanse dualisme als een uniek menselijke eigenschap is beschouwd (Hayles 2005: 177). Zulke interacties of 'feedback loops' tussen biologische organismen en geprogrammeerde machines nemen soms concrete vormen aan, bijvoorbeeld bij lichamelijke extensies als een pacemaker of een gehoorapparaat, maar het probleem is ook abstracter. Hayles reageert op theoretici uit de computerwetenschap die signaleren dat er biologische organismen zijn die werken volgens binaire/ digitale procedures, zoals zogenaamde cellular automata (systematische organismen waarin de binaire staat van een biologische eenheid (bijvoorbeeld een cel) bepaald wordt door de binaire staat van de eenheden daaromheen, zoals de binaire vlakken van een schaakbord bepaald worden door de kleur van hun aangrenzende vlakken). Hun verregaande opvatting is dat veel complexe systemen, inclusief biologische organismen zoals de mens, in wezen geregeerd worden door een dergelijk 'Regime of Computation' (Hayles 2005: 17). Zowel digitale machines als organismen zijn in deze visie terug te brengen tot een schaakbord van binaire parameters (*true* of *false*; 1 of 0) die in een economie van interacties of algoritmen complexe systemen kunnen vormen. De consequentie is dat dit regime meer dan slechts een metafoor is die we gebruiken om de werkelijkheid beter te begrijpen. Het is ook een mechanisme dat de werkelijkheid daadwerkelijk reguleert: de term is niet langer constructivistisch maar realistisch (Hayles 2005: 171). Als natuurlijke delen van de werkelijkheid eveneens een binaire logica volgen, de mens inclusief, dan worden we zelf evenzeer bestuurd door computerlogica als de computers waarvan wij de uitvinders dachten te zijn. Hayles gaat niet zo ver mee in deze redenering maar verwerpt wel de relatie van de mens als maker en subject tot de machine als creatie en object: 'this perspective enables me to see that they are neither objects to dominate nor subjects threatening to dominate me' (Hayles 2005: 243). Het menselijk bewustzijn, de menselijke taal en cultuur zijn volgens Hayles in zo'n mate gemedieerd en gefiltreerd door computationele mechanismen en media dat een scheiding tussen die twee domeinen niet meer houdbaar is.

Het is een grote stap om computationeel 'begrip' gelijk te stellen met menselijk begrip op basis van deze filosofische overeenkomst tussen bepaalde biologische organismen en digitale systemen. Maar er zijn ook overeenkomsten op andere niveaus. Zo hebben de recente ontwikkelingen in onze omgang met de computer (in het tijdperk van 'Web 2.0', zie hoofdstuk 4) laten zien dat ook menselijk gedrag zich op grote schaal uiterst goed laat voorspellen en begrijpen door computermodellen, die opnieuw in de kern functioneren volgens een binair systeem. De computer veronderstelt dat menselijke massa's zich gedragen volgens een regime of computation, en heeft daarin vaak gelijk. Zo lanceerde Google een model waarmee zij griepuitbraken konden voorspellen op basis van griep-gerelateerde zoekopdrachten.⁶ De redenering is dus: als er bovengemiddeld veel zoekopdrachten over 'koorts' zijn in een bepaalde geografische regio, wijst dat patroon op een griepuitbraak. Dit simpele model kent één verklaring die theoretisch wankel is, maar in de praktijk vrijwel altijd correct blijkt. Het is dezelfde logica die ten grondslag ligt aan Google's Ei van Columbus, de pagerank-index (voor het eerst voorgesteld in Brin & Page 1998): hoe meer hyperlinks naar een pagina en hoe meer gebruikers die pagina aanklikken na een zoekopdracht, hoe hoger de positie van die pagina in de index. Google is niet geïnteresseerd in de afwijking of de uitzondering maar begrijpt alleen maar de regel. En dat is niet voor niets, want de regel is altijd 'waar' of 'functioneel' voor de menselijke gebruiker: het antwoord op haar vraag staat vrijwel altijd tussen de eerste drie resultaten, omdat

⁶ <https://www.google.org/flutrends/about/>

haar voorgangers dat impliciet als ‘waar’ hebben aangemerkt. Het theoretische probleem dat patronen slechts een deel van de waarheid blootleggen is precies dat: een *theoretisch* probleem. In de praktijk ervaart de internetgebruiker dat de computerlogica van Google haar wereld perfect begrijpt.

Een ander voorbeeld van het modelmatige gedrag van mensen is wat Surowiecki ‘the wisdom of crowds’ noemde: het mechanisme waardoor een massa collectief altijd slimmer is dan de slimste enkeling (Surowiecki 2004: XIII-XIV). Zou je een complexe opdracht uitbesteden aan een grote, diverse groep van autonome en gedecentraliseerde individuen, dan is hun gemiddelde oplossing altijd beter dan die van elk individu in de groep. De groep produceert vanzelf een regressie naar de beste oplossing en elimineert zo ‘automatisch’ de afwijkingen, de ruis. Surowiecki stelt: ‘With most things, the average is mediocrity. With decision making, it’s often excellence. You could say it’s as if we’ve been programmed to be collectively smart’ (2004: 11). Je zou evengoed kunnen zeggen dat menselijk gedrag op grote schaal net zo voorspelbaar is voor de computer als de kleur van vak C4 op een schaakbord als gegeven is dat C3 zwart is. Als de hoeveelheid en variatie van data maar groot genoeg is, dan komt de computer net als het menselijk collectief altijd tot de beste oplossing – of de beste ‘waarheid’.

Het menselijke denken conformeert zich op grote schaal dus aan het patroonmatige paradigma van de computer, of vertoont daarmee althans belangrijke overeenkomsten. Een onderzoeker is niet in staat om zelfstandig dezelfde hoeveelheden data te verwerken als computers kunnen, maar zij zou die kwantiteiten kunnen benaderen als ze met duizenden of honderdduizenden menselijke denkers kon samenwerken. En in zulke groepen kunnen we op basis van de modellen van Google en ‘the wisdom of crowds’ stellen dat zij bij elkaar op vergelijkbare wijze waarnemen en begrijpen als een computer. Een ‘epistemologie van crowdsourcing’ verdient daarom geen eigen status. In de samenwerking tussen grotere groepen lezers en een patroonzoekende machine vullen de twee paradigma’s (menselijk en computationeel) elkaar goed aan. De menselijke lezer verwerkt data die hij begrijpt (literaire teksten) en levert ze interpreteerbaar aan. De applicatie verwerkt die informatie vervolgens op een schaal die voor de individuele mens niet mogelijk is.

Volgens computerwetenschapper en literatuurtheoreticus Stephen Ramsay ligt de toekomst van ‘computer assisted literary criticism’ precies in deze combinatie van computationele meetbaarheid en menselijke subjectiviteit (Ramsay 2011: 32). Zoals vele onderzoekers die actief zijn in de digital humanities werd ook hij geïnspireerd door een gebrek aan methode in de geesteswetenschappen (Ramsay 2011: 7). De literatuurtheorie in het bijzonder was volgens hem nooit geïnteresseerd in daadwerkelijke operationaliseerbaarheid van hun concepten: ‘In point of fact, most of these elaborations have the technical status of hypothesis, since they have not been confirmed empirically in terms of the data which they propose to describe – literary texts’ (Ramsay 2011: 4). In zijn boek *Reading Machines; Towards an Algorithmic Criticism* (2011) doet hij een voorstel voor hoe we tot operationalisering van die theorieën moeten komen in de letterkunde, door middel van de computer maar voorbij het niveau van het ‘fact checken’.

Ramsay beschrijft hoe digitale tekstanalyse hetzelfde doet als wat filologen al eeuwen doen: het herschikken en hervormen van een literaire tekst zodat interpretatie mogelijk wordt, een proces dat hij ‘deformance’ noemt. Hij stelt:

‘Any reading of a text that is not a recapitulation of that text relies on a heuristic of radical transformation. The critic who endeavors to put forth a “reading” puts forth not the text, but a new text in which the data has been paraphrased, elaborated, selected, truncated, and transduced. (...) In every case, what is being read is not the “original” text, but a text transformed and transduced into an alternative vision, in which, as Wittgenstein put it, we “see an aspect” that further enables discussion and debate’ (Ramsay 2011: 16).

De deformaties die de computer toepast op literaire teksten (woordenlijsten, wordclouds, topic models, concordanties) verschillen daarmee in essentie niet van een menselijke close reading: ‘it is the same thing

at a different scale and with expanded powers of observation. It is in such results that the critic seeks not facts, but patterns' (Ramsay 2011: 17). Mijn onderzoek maakt weliswaar geen gebruik van digitale tekstanalyse, maar verzamelt op vergelijkbare wijze aspecten van teksten, in dit geval personagekenmerken. De lemma's in de database zijn deformaties van de teksten, zoals een menselijke romananalyse een hiërarchie van tekstkenmerken aanbrengt teneinde die te kunnen interpreteren. Deze vorm van 'algorithmic criticism' gebruikt de computer dus om patronen op het spoor te komen die met het menselijk oog niet zichtbaar zijn. Deze bewuste, computationele reductie van de tekst geeft vervolgens en idealiter aanwijzingen voor een subjectieve, kritische interpretatie door de menselijke lezer. Ramsay spreekt van een 'desire to use the narrowing forces of constraint to enable the liberating visions of potentiality': de doelbewuste simplificaties leggen een nieuw 'betekenispotentieel' van literaire teksten bloot (Ramsay 2011: 32).

In een 'epistemologie van crowdsourcing' werken menselijke lezers dus samen met de rekenkracht van computers, maar de kennis die die samenwerking voortbrengt is niet fundamenteel anders dan de producten van traditionele analyses. Wel moet ik me bewust tonen van de hogere status van wat we 'patronen' noemen en van de mate waarin de applicatie 'ruis' elimineert in deze epistemologie. Het is niet ondenkbaar dat zulke patronen vanzelfsprekend validiteit opeisen in een wereld die in toenemende mate gestructureerd wordt door het digitale paradigma (vgl. Van der Deijl 2015). Juist in de context van letterkundig onderzoek zijn zulke patronen echter problematisch en misleidend. Literatuur is immers een unieke vorm van 'data'. Ze bezit een zekere zelfbewustheid en kent de mogelijkheid om culturele patronen (scripts, stereotypen, discoursen etc.) te saboteren met parodie, ironie, pastiche en dergelijke (anders dan neutralere data zoals geboortecijfers, financiële boekhoudingen, publicatiecijfers etc.). Om die reden zal ik in hoofdstuk 5 met roman-analyses de grotere patronen proberen tegen te spreken en te nuanceren door middel van een focus op de ruis uit de individuele roman. Alleen door de patronen in dialoog te laten treden met de afwijkingen kan ik beoordelen in welke mate het digitale paradigma de literaire werkelijkheid vervormt of juist verdiept.

5. Analyse

Om de hermeneutiek van de Personagebank te verdiepen analyseer ik de werking van de applicatie hieronder op twee niveaus en via twee modi: verifiërend en falsifiërend. Ik verifieer de kracht en het potentieel van crowdsourcing door enkele datavisualisaties van de Resultatenpagina toe te lichten en van een voorlopige, voorzichtige interpretatie te voorzien. Daarna probeer ik de werking van de applicatie te falsifiëren en te bevragen. Dat doe ik door drie romans uit de database te analyseren die de identiteit van de personages thematiseren en problematiseren: *Het jongensuur* (1969) van Andreas Burnier, *De asielzoeker* (2003) van Arnon Grunberg en *Het gym* (2011) van Karin Amatmoekrim. Behalve dat kenmerk hebben de romans in mijn ogen weinig gemeen, en de keuze is dan ook in zekere mate willekeurig en afhankelijk geweest van de input van lezers. Niettemin hoop ik dat deze drie romans elk een eigen aspect van de hermeneutiek van de applicatie belichten, om zo een nauwkeurig beeld te krijgen van de kansen en beperkingen van deze technologie.

5.1 Grafieken, kaarten, tabellen

Hierboven beschreef ik twee criteria voor het open-data beleid van de Personagebank: de applicatie moest de database zowel exploratief (*exploratory*) als uitleggend (*explanatory*) presenteren. De eerste functie betekent concreet dat gebruikers de data als totaaloverzicht kunnen inzien en data kunnen filteren, om zo een antwoord te vinden op eigen vragen aan de database.⁷ De tweede functie wordt vervuld door een pagina waarop enkele interpretaties van actuele verhoudingen in de database worden getoond door middel van grafieken, kaarten en tabellen.⁸ Die mogelijkheid is een voorbeeld van de uitleggende functie van de applicatie, omdat deze interpretaties geprogrammeerd zijn voor de gebruiker zonder dat zij er invloed op heeft (op de verschillende mogelijkheden voor interactiviteit na). Ik heb dus als ontwikkelaar voor de gebruiker bepaald welke verbanden en verhoudingen belangrijk zijn. Hieronder illustreer ik het potentieel van een dergelijke database door enkele datavisualisaties toe te lichten en te interpreteren. De figuren zijn gebaseerd op screenshots van de webpagina zoals die op 6 juni 2016 gegenereerd werd (met 821 personages en 176 romans).

Gender

Gender lijkt een van de meest kenmerkende eigenschappen voor personages. Slechts van 0.24 % (= 2 personages) van de personages werd het gender gemarkeerd als 'Onbekend'. De Resultatenpagina genereert een cirkeldiagram met de actuele genderverhouding onder alle personages, zie figuur 1. De mannen zijn licht in de meerderheid, maar dat lijkt niet te wijzen op een bepaalde voorkeur. Het verschil is een gevolg van een onevenwichtigheid in de romanselectie: die omvat 120 boeken van mannelijke auteurs en 56 van vrouwelijke. Schrijvers zijn namelijk geneigd om vaker te schrijven over personages van het eigen geslacht (zie tabel 1 in figuur 1). In het licht van dat gegeven, valt de balans nog mee (met een verhouding van 2 mannelijke op 1 vrouwelijke auteur in de database zou je een vergelijkbare verhouding verwachten onder de personages), temeer gezien het feit dat vrouwelijke schrijvers vaker een sprong maken naar het andere geslacht dan mannelijke: mannen schrijven vaker over mannen, en vrouwen schrijven bovendien vaker over mannen dan mannen over vrouwen. We zien die genderverhouding echter wel terug onder de vertellende personages (zie tabel 2 in figuur 1). Hoe dominanter een personage in een vertelling, hoe groter de kans dat het om een man gaat.

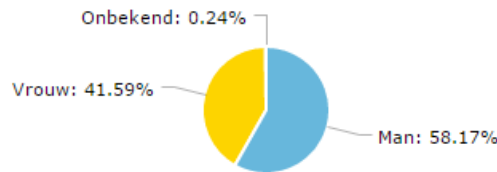
⁷ personagebank.hum.uu.nl/search-3.xq

⁸ personagebank.hum.uu.nl/resultaten.xq

Figuur 1. Genderverhoudingen

Gender (N = 821)

JS chart by amCharts



Tabel 1. Verband tussen gender auteur en personages

	Personage (V)	Personage (M)
Auteur (V)	53.44 %	46.56 %
Auteur (M)	36.6 %	63.4 %

Tabel 2. Man/vrouw-verhouding per rol

	Vrouw	Man	Totaal
Verteller	29.92 %	68.5 %	127
Hoofdpersonage	41.94 %	58.06 %	310
Nevenpersonage	45.05 %	54.69 %	384

Wat kunnen dergelijke patronen ons vertellen over schrijvers (m/v) en hun personages? Ten eerste suggereert het verschil in de mate waarin mannen en vrouwen ‘verbeeldingsprongen’ maken een bepaalde mannelijke norm in de literatuur. Mannen voelen minder de neiging om vrouwelijke personages op te voeren, terwijl de verhouding in romans van vrouwen bijna 50/50 is. Mogelijk verwachten vrouwelijke schrijvers dat zij serieuzer genomen worden als ze over mannen schrijven. Ten tweede wijzen deze patronen op een dominante mannelijke blik in de Nederlandstalige literatuur: vertellers zijn veel vaker mannen dan hoofdpersonages en nevenpersonages. Er lijkt dus een correlatie tussen gender en positie in de roman. Het is goed mogelijk dat mannen wel over vrouwen schrijven, maar dan hoofdzakelijk in ondergeschikte, narratologische posities. Een andere verklaring is echter dat het verschil voortkomt uit een bepaald verwachtingspatroon bij de participerende lezers. De categorieën ‘verteller’, ‘hoofdpersonage’ en ‘nevenpersonage’ zijn niet in steen gebeiteld, dus het is goed mogelijk dat bepaalde verwachtingen van de lezer een rol gespeeld hebben bij de toekenning ervan. Als iemand vooral boeken leest waarin mannen vertellers zijn, dan is het denkbaar dat zij een mannelijk personages sneller interpreteert als een centrale figuur. Beide verklaringen (een representatienorm of een verwachtingspatroon) zijn waardevolle resultaten die vragen om nader onderzoek. Welke (on)bewuste overwegingen maken schrijvers bij de keuze voor mannen of vrouwen? In welke mate spelen culturele verwachtingspatronen, scripts of denkbeelden over gender een rol bij de interpretatie van identiteiten door concrete lezers? In hoeverre hangt de waardering van een roman door critici af van het gender van de personages? Met een omvangrijker corpus in de database, komen antwoorden op dergelijke vragen in zicht.

Leeftijd

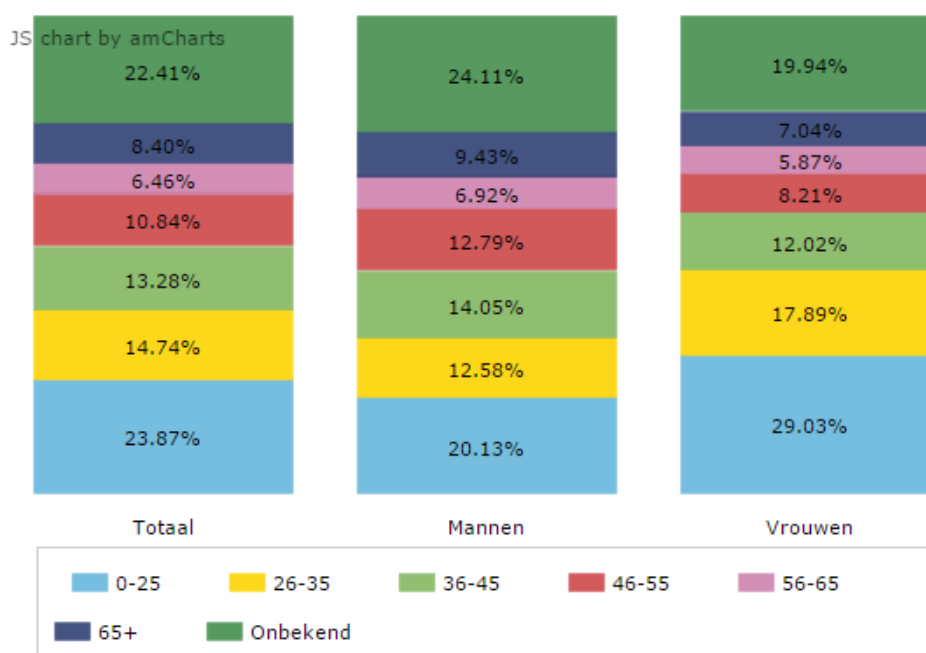
Anders dan gender en etniciteit is leeftijd in veel close readings niet snel een eigenschap waarachter een ideologisch complex wordt gesitueerd (het zogenaamde ‘ageism’ kan doorgaans op minder aandacht rekenen dan seksisme of racisme). Zelden is leeftijd een factor van betekenis voor de identiteit van personages. Dat blijkt bijvoorbeeld ook uit het hoge aandeel personages voor wie de leeftijd onbekend was (22.41 procent), na opleidingsniveau de eigenschap die het vaakst oningevuld bleef. Maar hier toont zich dan ook de waarde van een kwantitatieve benadering. In een database van 821 personages zijn vragen over structurele representaties van leeftijd of correlaties met andere eigenschappen wel mogelijk.

Zo wordt duidelijk dat jonge personages oververtegenwoordigd zijn. Als we de personages voor wie de leeftijd onbekend was niet meerekenen, dan blijkt dat bijna de helft (49.77 procent) onder de 35 jaar oud is (zie ook figuur 2). Verhalen vertellen dus in hoge mate over jonge mensen. Zo’n patroon zegt mogelijk iets over de aard van literatuur en over welke verhalen schrijvers meer of minder het vertellen waard achten. De jeugd geldt mogelijk als de belangrijkste levensfase, de fase waarin schrijvers en personages gevormd worden. Wat zich daarna nog afspeelt is van ondergeschikt belang, en vindt in ieder geval minder vaak zijn weg naar een roman.

Een dergelijk patroon vertelt ons nog niet veel over bepaalde ideologische vooringenomenheid tegenover bepaalde leeftijdscategorieën. Wel zijn zulke resultaten uiterst relevant om te begrijpen welke verhalen verteld worden in een bepaalde periode of cultuur. Als de database groeit en er verschillende romans uit verschillende jaren opgenomen zijn, worden transhistorische vergelijkingen mogelijk: vertelden romans van na de oorlog voornamelijk over de jeugd die een nieuwe toekomst tegemoet ging, of over ouderen die de oorlog(en) hadden meegemaakt? Spelen jonge personages vaker een hoofdrol in romans uit periodes waarin de jeugd door de cultuurgeschiedenis een sleutelrol is toegedicht, zoals in de jaren zestig of tijdens het decennium van de verloren generatie van de jaren tachtig? Zien we de bobbel die de babyboomers trekken door alle statistieken vanaf 1945 ook terug in de literatuurgeschiedenis? Met een uitgebreidere variant van deze database wordt kortom een literatuurgeschiedenis mogelijk van jeugd, generaties en ouderdom.

Figuur 2. Verhouding leeftijdscategorieën

Leeftijd (N = 821)



Opleidingsniveau

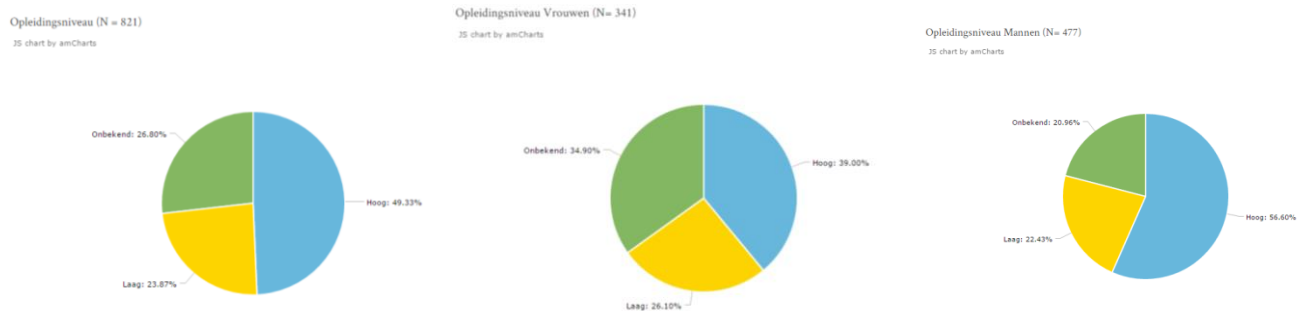
Behalve gender en leeftijd is ook opleidingsniveau een relevante eigenschap voor een demografie van romanpersonages. Ik beschouw het kenmerk hier als een indicator voor sociaaleconomische klasse: hoogopgeleide personages zullen waarschijnlijk tot een hogere klasse behoren en vice versa.

Opleidingsniveau geldt evenwel niet als een belangrijke karaktereigenschap voor personages: voor 26.8 procent van de personages was dit kenmerk onbekend. Het is daarmee van alle variabelen uit het formulier de minst algemene.

Onder de totale groep van personages zijn de hoogopgeleide personages sterk in de meerderheid. Het predicaat was van toepassing op 49.33 procent en slechts 23.87 procent werd getypeerd als 'laagopgeleid' (zie het eerste diagram in figuur 3). De webpagina biedt daarnaast de optie om de groep voor wie het opleidingsniveau niet bekend was te beschouwen als 'missing data'. Als de gebruiker die groep eruit filtert, zoals de pagina suggereert, verschijnt er een nieuw diagram. Daarin is de verhouding 67.39 hoogopgeleiden versus 32.61 procent laagopgeleiden. De Nederlandstalige literatuur houdt zich mogelijk vooral bezig met de hogere regionen van de samenleving.

Ook geeft de webpagina de optie om de verhouding uit te splitsen per gender. Met een klik op de knop transformeert het diagram naar de juiste verdeling voor mannelijke en vrouwelijke personages (diagram 2 en 3 in figuur 3). Daarin valt op dat opleidingsniveau een samenhang lijkt te vertonen met gender. Mannelijke personages zijn vaker hoogopgeleid (56.6 procent) dan vrouwen (39.0 procent) en minder vaak laagopgeleid (22.43 procent) dan hun vrouwelijke soortgenoten (26.1 procent). Daarbij is het opleidingsniveau voor vrouwen vaker onbekend (34.9 procent) dan voor mannen (20.96 procent). Kennelijk worden vrouwen minder in termen van hun maatschappelijke status geportretteerd. Dat resultaat illustreert de registers die schrijvers (en mogelijk de samenleving als geheel) aanwendt bij de representatie van vrouwen. In geautomatiseerde representaties zijn vrouwen vooral jong en mooi; mannen zijn succesvol en hebben een maatschappelijke positie.

Figuur 3. Opleidingsniveau totaal en per gender



Opnieuw kan een uitgebreidere variant van deze database inzicht geven in transhistorische verschuivingen van dit patroon. Zien we de tweede feministische golf terug in het opleidingsniveau van personages uit romans nadien? Vindt er – net als in de Nederlandse en Vlaamse samenleving – een stijging van opleidingsgraad plaats in de literatuur van na 1945 (analoog aan de niet-literaire wereld)? Of in welke mate wordt de zogenaamde democratisering van de literatuur – het fenomeen dat boeken sinds de jaren vijftig steeds goedkoper en dus toegankelijker worden in combinatie met de stijging van de welvaart – weerspiegeld in de maatschappelijke status van haar personages? Dergelijke vragen kunnen inzicht bieden in de manier waarop literatuur cultuurhistorische veranderingen verwerkt en waar de literatuur mogelijk juist conservatiever is dan de cultuur waarin ze ontstond. Het wordt mogelijk om de literatuur als een verzameling bronnen te zien waarmee systematisch een geschiedenis geschreven kan worden zoals de geschiedwetenschap dat doet met historische bronnen.

Beroep

Arbeid is van oudsher een centraal thema in de marxistische kritiek, maar is doorgaans niet het speerpunt van de activistische 'identity politics'. De meeste mensen beschikken immers over een zekere vrijheid om een beroep te kiezen, anders dan etniciteit, sociale achtergrond, gender of seksuele geaardheid. Hoewel arbeid in de niet-literaire werkelijkheid een zeer concrete politieke realiteit is van financiële en economische machtsverschillen, zijn beroepen van personages als zodanig in veel individuele gevallen betrekkelijk neutraal. Ik zeg nadrukkelijk 'als zodanig', want het zijn opnieuw de intersecties die wel betekenisvol zijn. De Marokkaanse taxichauffeur, de Nepalese huishoudster, de oudere werkloze: combinaties van zulke personage-eigenschappen dragen een politiek potentieel. De verklaring voor dat potentieel ligt vermoedelijk in het feit dat de ene eigenschap onlosmakelijk verbonden is met de andere, en dat de een daarvan niet gekozen kan worden.

Een kwantitatieve literatuurbenadering maakt een onderzoek mogelijk naar deze veronderstelde neutraliteit van arbeid en naar de patronen in de intersecties met 'beroep' als personage-kenmerk. De Resultatenpagina genereert bijvoorbeeld een lijst van de meest voorkomende beroepen en publiceert een top 10. De gebruiker kan op twee knoppen voor 'Mannen' en 'Vrouwen' klikken, waarop de lijst van het totale aantal personages uit Tabel 1 wordt vervangen door een van de twee lijsten van het betreffende geslacht. De studenten en scholieren spannen de kroon met een totaal van 69 en 54 gevallen. Die uitkomst komt overeen met het resultaat dat personages relatief jong zijn. Verder zijn de creatieve beroepen goed vertegenwoordigd: de fictieve samenleving kent geen gebrek aan kunstenaars, schrijvers en journalisten. Dit resultaat bevestigt de eerdergenoemde hypothese dat auteurs schrijven over personages die dicht bij hun eigen belevingswereld blijven, maar illustreert nog niet hoe arbeid een politieke status kan hebben in literaire representatie.

Interessanter is de combinatie van beroep en gender. De twee lijsten van meest voorkomende beroepen bij mannen en vrouwen vertonen opvallende verschillen. Staat bij de mannen 'Docent' nog altijd op de derde plek; bij de vrouwen is die ingenomen door 'Huisvrouw', gevolgd door 'Prostituee'. De verschillen tussen deze twee tabellen leggen opnieuw een verschil in register bloot. Vrouwen zijn huisvrouwen, verpleegkundigen en prostituees, terwijl mannen schrijver, ondernemer, journalist en fotograaf zijn. In deze context krijgt 'arbeid' een sterke politieke lading, niet alleen omdat het thema kennelijk niet dezelfde status heeft voor beide gendercategorieën, maar ook omdat het mogelijk een machtig cultureel denkbeeld blootlegt over de ondergeschikte rol van vrouwen op de arbeidsmarkt. Hoe de verhouding tussen fictieve en niet-literaire werkelijkheid ook wordt ingevuld; zulke patronen dwingen tot kritische reflecties op ons denken over man/vrouw-verhoudingen. Kan de emancipatie van vrouwen slagen zo lang de positie van de vrouw in culturele verbeeldingen alles behalve geëmancipeerd is? In hoeverre is een samenleving progressief als conservatieve wereldbeelden dominant blijven in literatuur, films of media? Waar schuilt het kritische potentieel van literatuur als die ene feministische roman overschaduw wordt door de massa?

Tabel 1. Top 10 meest voorkomende beroepen

Top 10 Beroepen (N = 821)			Top 10 Beroepen onder Vrouwen (N = 341)			Top 10 Beroepen onder Mannen (N = 477)		
Rang	Beroep	Frequentie	Rang	Beroep	Frequentie	Rang	Beroep	Frequentie
1	Student	69	1	Student	31	1	Student	37
2	Scholier	54	2	Scholier	29	2	Scholier	25
3	Docent	33	3	Huisvrouw	18	3	Docent	23
4	Kunstenaar	22	4	Prostituee	15	4	Schrijver	17
5	Schrijver	19	5	Docent	10	5	Kunstenaar	14
6	Huisvrouw	18	6	Verpleegkundige	10	6	Arts	13
7	Arts	17	7	Kunstenaar	8	7	Ondernemer	11
8	Prostituee	15	8	Werkloos	5	8	Boer	11
9	Ondernemer	14	9	Manager	5	9	Journalist	10
10	Journalist	13	10	Arts	4	10	Fotograaf	8

De vragen die de database oproept tonen dus de kansen van een kwantitatieve literatuurbenadering voor ideologiekritiek of cultuurkritiek. De combinatie gender + beroep is slechts één van de vele intersecties die de politieke betekenis van beroep aantonen. Ik noemde al voorbeelden van het verband herkomst + beroep: zijn mannen van Noord-Afrikaanse afkomst vaak terroristen (zoals in *De Winters VSV* (2012)) of taxichauffeurs (zoals in Maurits de Bruijns *Broer* (2012)), en wat betekent het als zij een keer tegen de stereotyperingen in ‘Student sociologie’ zijn (zoals in Marjolijn van Heemstra’s *De laatste Aedema* (2012)) of ‘Docent Nederlandse les’ (zoals in Britta Bolts *Heldhaftig* (2012))? Een andere mogelijkheid is beroep + type woonplaats: in hoeverre kennen fictieve steden ook een typisch verstedelijkte economie met bijbehorende beroepen? Zijn personages van het platteland ook vrijwel altijd boer, of is er een grotere diversiteit mogelijk? Verder zijn intersecties interessant die meer dan twee eigenschappen combineren, zoals herkomst + gender + beroep. Tot slot geldt opnieuw het transhistorische perspectief als een waardevolle optie die een database als de Personagebank ontsluit. Door romans uit verschillende jaren te vergelijken worden verschuivingen zichtbaar in de beroepen die mannen en vrouwen vervullen. Daarnaast biedt zo’n database mogelijk inzicht in literaire verwerkingen van sociaal-maatschappelijke ontwikkelingen, zoals de werkloze in tijden van werkloosheid en de komst van de gastarbeider in tijden van arbeidsoverschotten, of het begin van de freelancer en de flexibele economie.

Woonplaats

De Resultatenpagina projecteert de woonplaatsen van alle personages op een wereldkaart, met behulp van de Google Maps API. Door middel van een cluster-functie worden concentraties van personages gerepresenteerd in blauwe, gele en rode stippen met daarin het aantal personages dat binnen dat cluster woont. Bij een kleinere zoom worden de clusters groter, en vice versa (zie kaart 1). Bovendien is de locatie van de cluster nauwkeuriger bij een grotere zoom.

Net als ‘beroep’ is ‘woonplaats’ geen eigenschap met een evidente ideologische status. De woonplaats van personages lijkt in veel romans weinig betekenisvol en is ook vaak onbekend. Toch wordt de ideologische status van de gerepresenteerde domeinen pas zichtbaar in een kwantitatieve benadering. De grootte van de clusters zijn immers een indicatie van de frequentie waarin verhalen zich afspelen in de bewuste domeinen. Ze reflecteren zo de traditionele machtsverhouding tussen centrum en periferie in de literaire verbeelding. Wie bijvoorbeeld inzoomt op Nederland en Vlaanderen, ziet dat Amsterdam een belangrijk centrum is in de fictieve samenleving. Deze verhouding zegt ten eerste iets over de culturele

verdeling van aandacht over verschillende geografische domeinen: lezers krijgen relatief vaak verhalen voorgeschoteld over de hoofdstad. Daarnaast bieden deze clusteringen ook aanleiding om kruisverbanden te onderzoeken: hebben Amsterdamse personages andere eigenschappen dan karakters uit de rest van Nederland? Is de Amsterdamse populatie bijvoorbeeld jonger of hoger opgeleid dan het gemiddelde Nederlandse personage? Het is daarbij denkbaar dat de aandacht voor de hoofdstad samengaat met een dominant beeld van ‘Amsterdamsheid’ of ‘Randstedelijkheid’. Zulke vragen vormen hypothesen die getest kunnen worden met deze database. Zo biedt een kwantitatieve benadering van de eigenschap ‘woonplaats’ een aanknopingspunt voor een analyse van een literaire discours over de relatie tussen identiteit en woonplaats.

Kaart 1. Personages woonachtig in Nederland

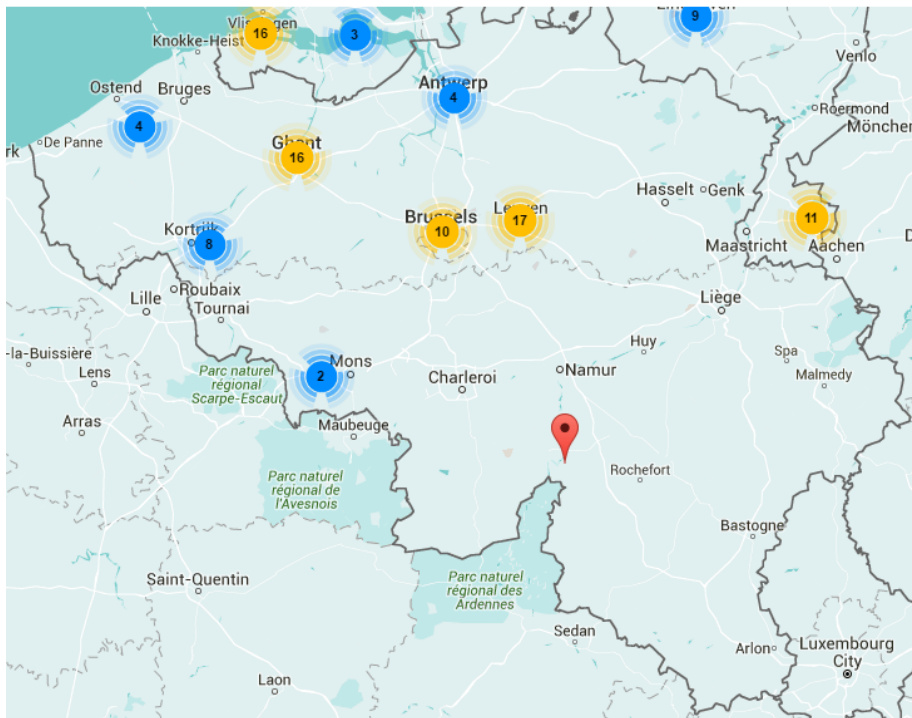


Interessant is bijvoorbeeld ook dat er in het fictieve België minder sprake lijkt van één dominant centrum. Het aantal Belgische personages is nog laag, maar het valt al op dat Antwerpen, Gent, Brussel en Leuven in meer gelijke mate vertegenwoordigd zijn (zie kaart 2). Daarnaast is duidelijk dat Wallonië nauwelijks een rol van betekenis speelt in Nederlandstalige verhalen. Slechts 3 van de 59 Belgische personages van wie de woonplaats bekend is, woont ten zuiden van de taalgrens (4,8 %). Zo lijkt de verspreiding van personages ook een weerspiegeling van politieke verhoudingen en culturele mobiliteit in de wereld buiten de tekst.

Daarnaast lijkt er een verband tussen de gerepresenteerde domeinen en de status van die domeinen in de werkelijkheid van de boekenmarkt. Het leeuwendeel van de uitgevers die in het corpus zijn vertegenwoordigd, houdt in Amsterdam kantoor.⁹ Ook andere literaire instituten zijn in de hoofdstad gevestigd, zoals redacties van kranten en tijdschriften, het Letterenfonds en de CPNB. Amsterdam is kortom het epicentrum van de Nederlandse literaire wereld, en dat lijkt zijn weerslag te hebben op de literatuur die de stad produceert. De kwantitatieve patronen tonen een mogelijke verstrengeling tussen de literaire verbeelding en de politiek-culturele werkelijkheid waarin die gedijt en ontstaat.

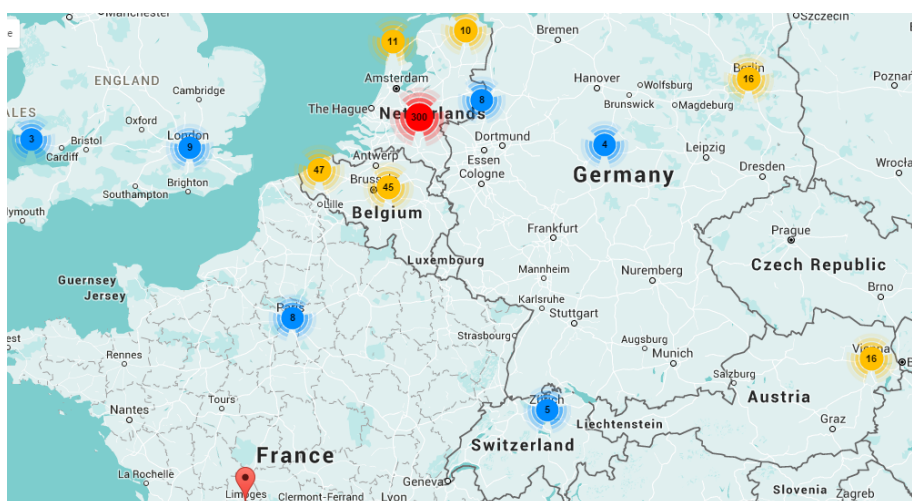
⁹ www.personagebank.hum.uu.nl/blog.xq

Kaart 2. Personages woonachtig in België



Op een hoger niveau blijkt dat de voorkeursstatus van het centrum ten opzichte van de periferie eveneens zichtbaar is buiten het fictieve Nederland en Vlaanderen. Wanneer Nederlandstalige romans zich situeren in andere Europese steden, gebeurt dat vrijwel uitsluitend in de hoofdsteden van die landen. Parijs, Londen, Berlijn en Wenen vertonen clusters in een verder verlaten landschap (zie kaart 3). Het literaire discours lijkt dus in hoge mate verstedelijkt en gecentreerd (vergelijk ook de tweede diagram in figuur 4). Verhalen uit de periferie worden minder vaak verteld, uitgegeven, gelezen en/ of ingevuld door participanten.

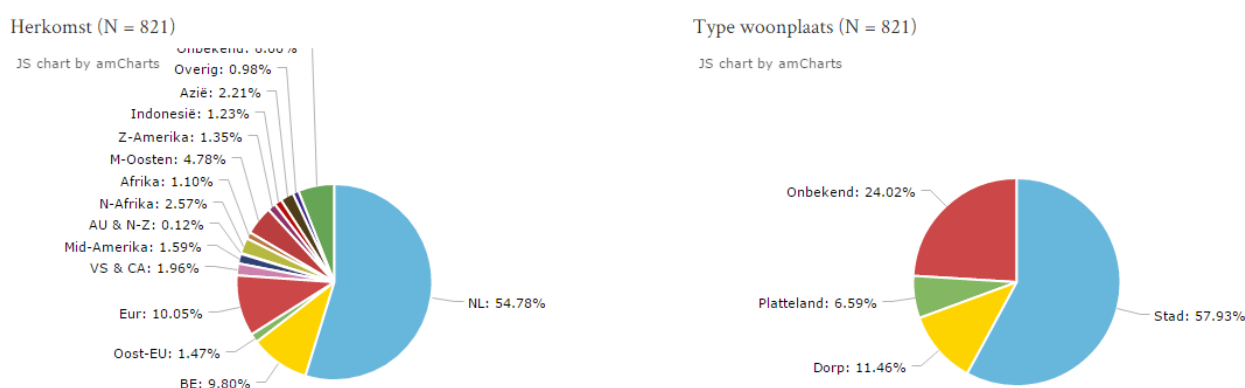
Kaart 3. Personages woonachtig in Europa



Herkomst

De applicatie interpreteert de herkomst van alle personages door hun onderlinge verhoudingen te tonen in een cirkeldiagram (zie figuur 4). Het webformulier stond zestien categorieën toe: Nederland, België, Europa overig, Oost-Europa en Rusland, VS en Canada, Midden-Amerika, Australië en Nieuw-Zeeland, Noord-Afrika, Afrika overig, Midden-Oosten, Zuid-Amerika, Indonesië, Azië overig, Overig, Fictief en Onbekend. Lezers werd gevraagd de herkomst van personages in te delen in een van deze regio's op basis van 'de geboortegrond van het personages, tenzij die afwijkt van zijn of haar ouders. In dat geval is de herkomst van de ouders leidend.'¹⁰ Zo voorkom ik dat migranten van de tweede of derde generatie als 'Nederlands' worden geïdentificeerd, hoewel hun 'Nederlandsheid' mogelijk in belangrijke mate verschilt of gethematiseerd wordt in de roman.

Figuur 4. Herkomst en type woonplaats



In het diagram valt op dat een meerderheid van de personages 'Nederland' als herkomst heeft (54.78 %). Grote minderheden zijn daarnaast personages uit 'België' (9.80 %) en 'Europa overig' (10.05 %). De CBS-categorie 'Niet-Westerse allochtoon' is in mindere mate vertegenwoordigd in de fictieve demografie. 4.78 % van de personages komt oorspronkelijk uit het Midden Oosten, 2.57 % uit Noord-Afrika en 2.21 % uit Azië overig. Het is niet eenvoudig om op basis van deze cijfers uitspraken te doen over de diversiteit van de Nederlandstalige literatuur. Dat meer dan de helft uit Nederland komt lijkt een veelzeggend resultaat, maar voordat ik dit mag interpreteren is een norm nodig, een hypothese. Wijkt dit resultaat af van wat je mag verwachten en welke gegevens rechtvaardigen zo'n verwachting? Het ligt misschien voor de hand om de verhoudingen in de actuele Nederlandse samenleving als ijkpunt te nemen, opgetekend door het CBS of vergelijkbare organisaties. Die keuze roept echter eveneens theoretische vragen op. We hebben immers geen enkele aanwijzing dat de romanliteratuur een accurate graadmeter is of wil zijn voor sociale verhoudingen in de niet-literaire werkelijkheid. Bovendien omvat het corpus de romanproductie van 1945 tot heden, dus een demografische jaarpeiling van het CBS zou altijd arbitrair zijn. Een andere manier zou zijn om de diversiteit van de Nederlandstalige literatuur te vergelijken met die van literatuur uit andere taalgebieden, maar die gegevens zijn (nog) niet voorhanden.

Ik heb dus een werkbare definitie van 'diversiteit' nodig voordat het diagram uit figuur 4 interpreteerbaar is. In de kern verwijst 'diversiteit' voornamelijk naar de mate waarin verscheidene groepen in verscheidene rollen gerepresenteerd zijn in de literatuur. Met deze definitie moet ik diversiteit dus intern-relatieel beschrijven, in termen van de interne verhoudingen tussen de gerepresenteerde groepen, in plaats van extern-relatieel (door te verwijzen naar verhoudingen in de niet-literaire wereld). Zo

¹⁰ www.personagebank.hum.uu.nl/doemec.xq

komen we interne representatiepatronen op het spoor en voorkom ik dat ik de mimetische kwaliteiten of pretenties van de literatuur worden overschat. Met deze opvatting zijn de verhoudingen zoals gerepresenteerd in figuur 4 op zichzelf nog niet direct betekenisvol. Het zijn de intersecties van eigenschappen die aanleiding kunnen geven tot uitspraken over diversiteit en representatiepatronen. Relevante vragen zijn dan: zijn niet-Westerse personages vaker mannelijk of vrouwelijk, hoog- of laagopgeleid, verteller of nevenpersonage etc.? Als er geen sprake zou zijn van stereotyperingen of representatiepatronen, dan zou je verwachten dat gender-, opleidingsniveau- en narratologische verhoudingen gelijk zijn in alle herkomst-categorieën. Als bepaalde groepen significant afwijken van het gemiddelde, kan ik met recht spreken van een representatiepatroon, of nog sterker: van een discours. Naarmate de database groeit zijn dergelijke kruisverbanden eenvoudig aantoonbaar.

Een eerste stap in deze richting is al de zoek- en filterfunctie die de applicatie momenteel biedt (zie figuur 5). In de toekomst hoop ik meerdere zoekfilters te gaan combineren, zodat bijvoorbeeld alle 'Mannen' uit 'Noord-Afrika' opvraagbaar zijn en te vergelijken met alle 'Vrouwen' uit 'Noord-Afrika'.

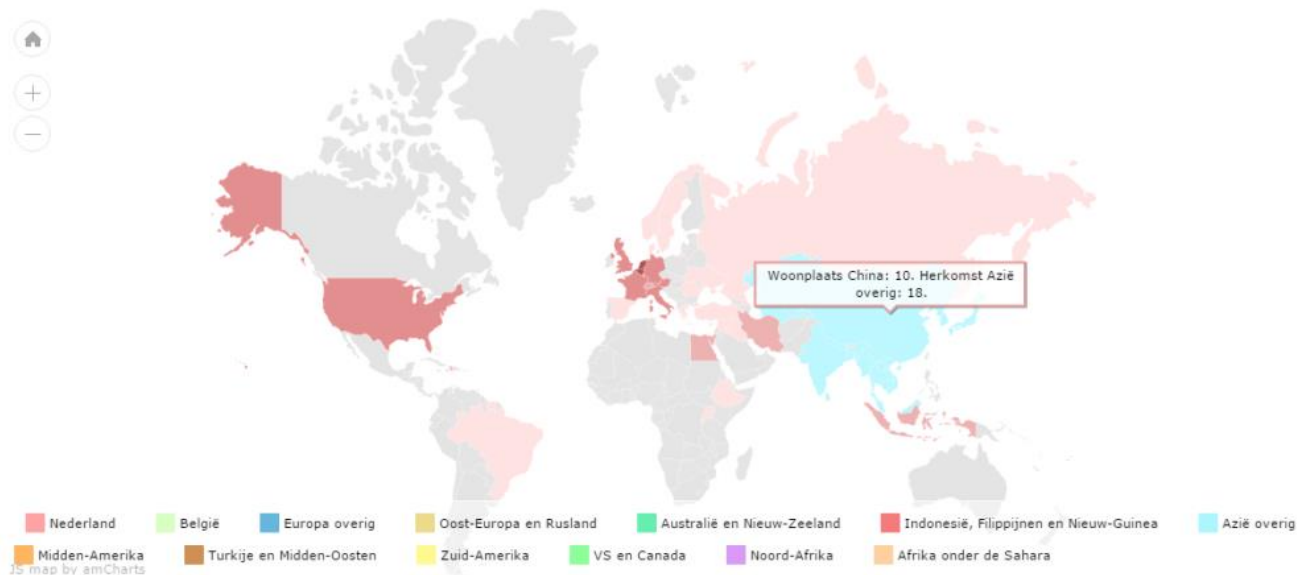
Figuur 5. Zoek- en filterfunctie

Gezocht op: "China" in: *Personage*

Naam	Rol	Gender	Leeftijd	Herkomst	Woonplaats (land)	Opleiding	Beroep	
'Ik'	Verteller	Man	Onbekend	Nederland	Chongqing (China)	Hoogopgeleid	Universitair Docent	Meer info
Ella	Nevenpersonage	Vrouw	0-25	Azië overig	Chongqing (China)	Hoogopgeleid	Student	Meer info
Ling	Hoofdpersonage	Vrouw	26-35	Azië overig	Chongqing (China)	Laagopgeleid	Manager wasmachinefabriek	Meer info
Meester Yao	Hoofdpersonage	Man	65+	Azië overig	Chongqing (China)	Onbekend	Boedhistisch leermeester	Meer info
Mu	Nevenpersonage	Vrouw	0-25	Azië overig	Chongqing (China)	Hoogopgeleid	Student	Meer info
Namyang	Nevenpersonage	Man	Onbekend	Azië overig	Tibet (China)	Hoogopgeleid	Monnik	Meer info
Sheng	Hoofdpersonage	Man	26-35	Azië overig	Chongqing (China)	Hoogopgeleid	Universitair Docent	Meer info

Verder biedt de Resultatenpagina een specifieke zoekfunctie op landen en herkomst met een alternatieve landenkaart (zie Kaart 3). Op die kaart is een land ingekleurd met een bepaalde rood-tint als er één of meer personages uit de database in dat land wonen. De tint wordt donkerder naarmate er meer personages wonen. Door te klikken op een land verschijnt er een tabel met alle personages uit het betreffende land en een label waarin staat hoeveel personages er wonen en hoeveel personages afkomstig zijn uit de regio waartoe het land behoort (zie Kaart 3).

Kaart 3. Herkomst en woonplaats per land



Personages uit China:

Naam	Rol	Gender	Leeftijd	Herkomst	Woonplaats	Opleiding	Beroep	Roman
'Ik'	Verteller	Man	Onbekend	Nederland	Chongqing (China)	Hoogopgeleid	Universitair Docent	Meer info
Ella	Nevenpersonage	Vrouw	0-25	Azië overig	Chongqing (China)	Hoogopgeleid	Student	Meer info
Ling	Hoofdpersonage	Vrouw	26-35	Azië overig	Chongqing (China)	Laagopgeleid	Manager wasmachinefabriek	Meer info
Meester Yao	Hoofdpersonage	Man	65+	Azië overig	Chongqing (China)	Onbekend	Boedhistisch leermeester	Meer info
Mu	Nevenpersonage	Vrouw	0-25	Azië overig	Chongqing (China)	Hoogopgeleid	Student	Meer info
Namyang	Nevenpersonage	Man	Onbekend	Azië overig	Tibet (China)	Hoogopgeleid	Monnik	Meer info
Sheng	Hoofdpersonage	Man	26-35	Azië overig	Chongqing (China)	Hoogopgeleid	Universitair Docent	Meer info
Tang	Nevenpersonage	Man	Onbekend	Azië overig	Chongqing (China)	Hoogopgeleid	Ingenieur	Meer info
Tsayong	Nevenpersonage	Man	Onbekend	Azië overig	Chongqing (China)	Hoogopgeleid	Makelaar	Meer info
Yong	Hoofdpersonage	Man	0-25	Azië overig	Chongqing (China)	Hoogopgeleid	Student	Meer info

5.2 *Het jongensuur*

De lezer die *Het jongensuur* (1969) van Andreas Burnier invulde, markeerde één personage uit dat verhaal, met de volgende kenmerken:

Tabel 1. Lezersinput van *Het jongensuur* van Andreas Burnier

	Personage 1
Naam	ik
Rol	Verteller
Gender	Vrouw
Leeftijd	0-25
Herkomst	Nederland
Woonplaats (land)	Nederland
Woonplaats	meerdere
Woonplaats (type)	Onbekend
Opleiding	Onbekend
Beroep	kind

De inzet van deze analyse is niet om aan te wijzen waar de lezer het in mijn ogen wel of niet bij het rechte eind had. In plaats daarvan wil ik voor enkele van deze kenmerken nagaan welke aspecten van het verhaal ze representeren en welke niet. Zo'n formulier is tenslotte een representatie van een representatie, en de leidende vraag in dit hoofdstuk is hoe de eerste zich verhoudt tot de tweede.

Het hoofdpersonage uit *Het jongensuur* is inderdaad een ik-verteller (haar voornaam wordt een enkele keer genoemd: Simone), een joods meisje dat tijdens haar jeugdige jaren onder de Duitse bezetting moet onderduiken op meerdere plekken in Nederland en dat graag tot jongen wil 'metamorfoser'. Vele lezers zullen deze samenvatting vermoedelijk zien als een nauwkeurige kenschets van het verhaal en van het centrale personage. Simones identiteit en functie in het verhaal (meisje, verteller, niet ouder dan 25, een kind nog zonder opleiding, woont op meerdere plekken in Nederland) zijn dan ook grotendeels gevangen in het formulier. De afwezige kenmerken zijn echter eveneens evident: Simones joodse identiteit, het feit dat het verhaal zich in de Tweede Wereldoorlog afspeelt, het feit dat Simone opgroeit tijdens de vertelling (het 'coming of age'-aspect) en uiteraard: Simones sterke wens om een jongen te zijn (of om behandeld te worden als de jongen die ze is). Het ingevulde formulier representeert dus een interpretatie die tot op zekere hoogte nauwkeurig, herhaalbaar en controleerbaar is. De applicatie produceert echter ook een stereotypering, niet alleen omdat het bepaalde kenmerken vooropstelt en uitvergroot, maar ook omdat het formulier betekenisvolle, complexere dimensies van de vertelling simpelweg niet ziet.

Een zo'n dimensie is de genderidentiteit van Simone. *Het jongensuur* wordt omgekeerd chronologisch verteld: de vertelling begint in 1945 en eindigt in 1940. Het verhaal laat zich dan ook lezen als de omgekeerde ontwikkeling van Simones identiteit als meisje. Deze identiteit is in de gehele vertelling complex, maar Simone lijkt die in het eerste hoofdstuk te ervaren als een voldongen feit wanneer zij uit het zwembad geweerd wordt tijdens het jongensuur: "“Ze moet eruit, ze is een meisje,” zei de dikste jongen' (Burnier 1969: 13). Tegelijkertijd identificeert Simones moeder haar dochter als jongen wanneer zij – met kortgeschoren kapsel – thuiskomt na de bevrijding: "“Wie is die jongen?” vroeg mijn moeder tenslotte' (14). Voor Simone zelf ligt haar gender evenmin vast. Soms ziet ze zichzelf als jongen, zoals in het volgende fragment:

‘Om verschillende redenen was een spoedige metamorfose noodzakelijk. In de eerste plaats moest mijn vrouwelijke anatomie en fysiologie niet al te lang meer duren, want dan zou mijn lichaam voorgoed zijn verpest en zou de verandering niet meer lukken. In de tweede plaats was het vernederend voor mij, als jongen, nog veel langer door anderen als meisje te worden behandeld.’ (Burnier 1969: 18)

Op andere momenten beschouwt de hoofdpersoon haar jongen-zijn als een verbeelding:

‘Ik was elf, en ik kon mij nu wel verbeelden ook een jongen te zijn, maar ik had iets wat oom Christfield niet had en nooit zou krijgen: een moederlijk vermogen om te kalmeren, gerust te stellen en een warm gevoel van veiligheid te geven.’ (Burnier 1969: 87)

De tekst biedt kortom voldoende argumenten om Simone ofwel als ‘Vrouw’, ‘Man’, of ‘Transgender’ (de opties die het formulier biedt) te classificeren, maar bovenal suggereert deze roman dat Simones gender onbepaald is en dat al die termen misleidend of ontoereikend zijn. De tekst saboteert deze neiging om te bepalen, te ‘kennen’ of te determineren met zijn omgekeerde chronologie. De lezer wordt meteen geconfronteerd met het onvermijdelijke lot: het is Simone niet gelukt om helemaal een jongen te worden, maar in zijn chronologische omkering schort het verhaal dat lot op, het houdt een verlangen in stand om tegen beter weten in – als een omgekeerde tragedie – aan dat lot te ontkomen. Datzelfde vergeefse verlangen om aan haar (lichamelijke) predestinatie te ontsnappen herkent Simone in haar ergernis over de protestantse lotsbestemming:

‘Je kon zelf aan je predestinatie niets toe- of afdoen. Het was mogelijk dat je je psalmen ongelooflijk langzaam zong, dat je je een ongeluk bad, dat je de bijbel van links naar rechts uit je hoofd kende, en toch voor de eeuwige Hel was bestemd in Zijn Voorzienigheid en Glorie.’ (Burnier 1969: 43)

Bij wijze van verzet tegen deze predestinatie benadrukt de vertelling met de vorm de tijdelijkheid en fluiditeit van Simones gender, ondanks de ogenschijnlijke onontkoombaarheid van haar identiteit: soms voelt ze zich een meisje, soms een jongen, soms homoseksueel en dan weer heteroseksueel.

Rijmend met de chronologische omkering thematiseert het verhaal een algemener verlangen om allerlei logische ordes en ogenschijnlijk stabiele gronden – naast de heteroseksuele genderdichotomie – omver te werpen. Zo vragen Simone en haar vriendinnetje Tessa in het eerste hoofdstuk een passant om een tijdmachine. De man antwoordt: “Wat bedoel je?” vroeg de man. “Willen jullie weten hoe laat het is?” “Nee meneer, we willen een tijdmachine” (Burnier 1969: 17). De man biedt aan om de tijd vast te stellen, maar de kinderen willen juist ontsnappen aan ieder tijdstip en daarmee aan de temporele orde. Een ander voorbeeld is Simones wiskundelessen van ‘Oom Christfield’, een student die op hetzelfde onderduikadres de verplichte tewerkstelling ontduikt. Hij leert haar dat een rekensom als 2a- 3a mogelijk is zolang je afspreekt dat negatieve en daardoor ogenschijnlijk irrationele cijfers, wel logisch kunnen zijn (Burnier 1969: 85). En behalve de cijfers verliezen tenslotte ook de woorden hun veronderstelde stabiliteit wanneer Simone Engelse lessen krijgt: ‘Een gevoel van ontreddering overvalt mij als Tom mij zulke dingen vertelt. Niets kun je weten, alles wankelt als de woorden hun vertrouwde betekenis verliezen.’ (Burnier 1969: 94)

De tweede dimensie die de applicatie niet begrijpt, hangt sterk samen met de eerste: Simones joodse identiteit. Net als haar mannelijke zelfbeeld wordt ze gedwongen om haar joodse eigenheid te onderdrukken en te verstoppen tijdens het lange ‘jongensuur’ dat de Tweede Wereldoorlog als historische periode was. Ook haar joods-zijn zou ze daarom van zich willen afwerpen: ‘Ik zou graag blond en protestant willen zijn, zes jaar op één lagere school zitten, en altijd weten wat er de volgende week gaat gebeuren’ (Burnier 1969: 99). Bovendien is deze etnische identiteit eveneens fluide: ‘In mijn familie nemen

de joodse kenmerken toe met de ouderdom', zoals Simones 'vrouwelijke' lichaamskenmerken toenemen met de jaren (Burnier 1969: 27). Opnieuw ontsnapt Simone zo aan de bepaaldheid van haar identiteit.

Die unieke ontsnapping krijgt in *Het jongensuur* ook een historische lading. Want voor Simone mag haar identiteit als joods meisje/ transgender en homoseksueel complex zijn, onder het nazisme waren zulke categorieën behalve gefixeerd ook levensgevaarlijk. De Shoah speelt allerminst een bijrol in de vertelling: in het eerste hoofdstuk beschrijft Simone hoe zij na de bevrijding bij het bureau van het Rode Kruis informeert naar de vele vermiste familieleden. Zo verschijnt uiteindelijk achter de naam van haar vriendje Werner – die de lezer pas in het laatste hoofdstuk (in 1940) weer tegenkomt – 'vermoedelijk omgekomen' (Burnier 1969: 23). Het eerste en het laatste hoofdstuk eindigen met een verwijzing naar Werner, achtereenvolgens het bericht van zijn dood en de ontmoeting tussen Simone en Werner. Simone ontsnapt aan het lot dat zovele joden ten deel viel tijdens de oorlog en haar a-chronologische verhaal contrasteert met het chronologische narratief waarmee de Westerse cultuur de slachtoffers van de Holocaust doorgaans herinnert (van maatschappelijke isolatie tot de deportatie en uiteindelijk de dood). De circulaire verwijzing naar Werner functioneert zo als een monument, om de lezer ervan te vergewissen dat Simones niet-circulaire en tegendraadse verhaal een uitzondering en buitencategorie was.

Door de parallellen tussen Simones gender en haar joods-zijn wordt haar kinderlijke verzet tegen deze bepaaldheid ook al snel politiek-feministisch. Wanneer de burgemeester na de bevrijding aankondigt dat de vrouwen die 'met de bezetters hebben geheuld' publiekelijk kaalgeschoren zullen worden, vraagt Simone haar vader verontwaardigd wat er met de mannen zal gebeuren die fout waren (Burnier 1969: 36-37). Ze stelt daarop de onderdrukking van de jood gelijk aan die van de vrouw: 'Vrouwen en joden, dat is bijna hetzelfde dacht ik. Ze kunnen niets terug doen, ze zijn altijd schuldig' (Burnier 1969: 38). Hier speelt ook nog mee dat Simone eerder zichzelf kaalgeknipte uit verzet en frustratie na haar eerste menstruatie: ze is daarmee niet meer te onderscheiden van de vrouwen die de gewelddadige consequenties van een dominante moraal ook aan den lijve ondervonden. Simones lichamelijke predestinatie lijkt hier dus uit te groeien tot een metafoor voor de universele lotsbestemming van alle vrouwen.

Ik beschreef hierboven dat het doel van de applicatie moest zijn om representatie en diversiteit van romanpersonages met kwantitatieve data inzichtelijk te maken. Wanneer ik die data stel tegenover de kwalitatieve analyse van een roman als *Het jongensuur*, dan zijn enkele uitspraken over de aard en functie van deze data mogelijk.

Ten eerste laat deze roman van Burnier zien dat de schema's (zoals man/ vrouw/ transgender) die de applicatie hanteert bij haar verwerking van de romans, politiek kunnen zijn. Met 'politiek' bedoel ik hier dat de onderzoeksopzet een bepaalde machtsverhouding in stand houdt en misschien zelfs versterkt. Burniers verhaal verzet zich zowel in vorm als in inhoud tegen de maatschappelijke en politieke denkschema's voor identiteiten die Simone als onderdrukkend en levensbedreigend ervoer. De Personagebank giet dit verhaal dus onverbiddelijk in de mal waaraan het nu juist met literaire en narratologische middelen probeerde te ontsnappen. In het hermeneutische spectrum ontnemt de applicatie dus de autoriteit van zowel de tekst als de lezer. De computer is niet alleen een nieuwe factor geworden in het proces van betekenisgeving; hij geniet in dit onderzoek ook de grootste autoriteit.

Ten tweede laat *Het jongensuur* zien dat de applicatie mogelijk een wat al te eenzijdig beeld van diversiteit onderzoekt. Simone verdwijnt in de dataset op de grote hoop van vrouwen en Nederlanders in de leeftijdscategorie van 0 tot 25 jaar. Teruggebracht tot die kenmerken is ze interpreteerbaar of 'kenbaar' voor de geautomatiseerde berekeningen van de verhoudingen in de database (tussen gender- en leeftijdscategorieën, woonplaatsen etc.). De aannames van de applicatie gaan daarmee echter voorbij aan de specifieke betekenis van Simone als 'vrouw' (voor zover dat label überhaupt nauwkeurig is): die identiteit is tenslotte onlosmakelijk verbonden met haar eigenheid als joodse. De dataset biedt dus weliswaar de potentie om met statistiek de zogenaamde 'intersecties' – verbanden tussen identiteitskenmerken – bij alle personages op te sporen (zoals het verband tussen opleidingsniveau en gender), maar voor de mogelijke individuele intersecties van personage-identiteiten in een verhaal is het

volledig blind. De data worden dus gedestilleerd uit de verhaalcontext op twee niveaus: wanneer de lezer een roman interpreteert en het formulier invult, en wanneer de technologie alle formulieren bij elkaar neemt en de totale dataset interpreteert. *Het jongensuur* illustreert dat er in dat destilleerproces verbindingen tussen de atomen en moleculen van het verhaal (joods – transgender / man – vrouw / goed – fout) afgebroken worden die juist voor een onderzoek naar diversiteit zeer wezenlijk en relevant zijn. Het residu (de gedecontextualiseerde data) is bruikbaar en biedt inzicht in bepaalde patronen van diversiteit, maar de validiteit van die patronen is nog ongewis zolang het gissen blijft naar de elementen en mogelijke patronen die in de destillatietoren verloren gingen.

5.3 *De asielzoeker*

De asielzoeker is het verhaal van de Nederlandse Christian Beck, vertaler van gebruiksaanwijzingen en woonachtig in Göttingen, wiens vriendin ‘de vogel’ een dodelijke ziekte krijgt (vermoedelijk kanker) en in haar laatste maanden besluit te trouwen met een asielzoeker. De asielzoeker trekt bij de twee in, waarna Beck voortaan onder de kapstok slaapt. Het verhaal van dit huishouden wordt afgewisseld met flashbacks naar een tijd waarin Beck en de vogel in de Israëlische badplaats Eilat woonden, waar de vogel onderzoek deed en Beck elke dag een bordeel bezocht. In dat bordeel wekt een Russische prostituee, Sosha Minkiewicz, op een dag zo veel agressie bij hem op dat hij haar een oog uitsteekt met een schroevendraaier. De politie verhoort Beck over deze mishandeling, maar vervolgt hem niet.

Van *De asielzoeker* (2003) van Arnon Grunberg zijn 3 personages ingevuld, weergegeven in tabel 1.

Tabel 2. Lezersinput van *De asielzoeker* van Arnon Grunberg

	Personage 1	Personage 2	Personage 3
Naam	Christian Beck	‘de vogel’	‘de asielzoeker’
Rol	Verteller	Hoofdpersonage	Nevenpersonage
Gender	Man	Vrouw	Man
Leeftijd	36-45	36-45	26-35
Herkomst	Nederland	Nederland	Noord-Afrika
Woonplaats (land)	Duitsland	Duitsland	Duitsland
Woonplaats	Göttingen	Göttingen	Göttingen
Woonplaats (type)	Stad	Stad	Stad
Opleiding	Onbekend	Hoogopgeleid	Onbekend
Beroep	Vertaler; voormalig schrijver	Wetenschappelijk onderzoeker	bouwvakker

Opnieuw valt op dat deze lezersinput een nauwkeurige kenschets is van de drie personages. Het klopt bijvoorbeeld dat uitsluitend naar ‘de vogel’ (Christian Becks vriendin) wordt verwezen met die bijnaam, en hetzelfde geldt voor het personage ‘de asielzoeker’. Verder wordt inderdaad gezegd dat Beck een vertaler is maar vroeger verhalen schreef, dat ‘de vogel’ wetenschappelijk onderzoek doet en dat ‘de asielzoeker’ bijkluist als bouwvakker. Ze wonen gedrieën in het Duitse Göttingen, en hebben de middelbare leeftijd, behalve de asielzoeker, van wie wordt gezegd dat hij een stuk jonger is. De applicatie lijkt dus goed te werken in haar functie als ‘objectievere’ of inter-subjectieve maat om personage-identiteiten te classificeren.

Die vaststelling neemt niet weg dat ook *De asielzoeker* problematisch is vanuit een meer nauwgezette, menselijke lezing. Een eerste voorbeeld is de functie van namen en naamgeving in de roman. De applicatie vraagt de lezer van elk personage de naam, maar doet daar in haar 'data-publication' functie vervolgens niets mee. Voor de applicatie hebben namen dus geen betekenis, hoewel naamgeving (van plaatsen of personages) in veel romans wel degelijk een zekere lading draagt. De invloed van namen op de manier waarop personages worden gerepresenteerd wordt bijvoorbeeld ook duidelijk wanneer personages naamloos blijven. 'Naamloos' betekent dan al snel 'stemloos', een personage kortom dat bewust of onbewust in de periferie van het verhaal blijft. De specifieke namen in de roman zijn een uiting van een vertellende of focaliserende naamgever. In *De asielzoeker* is die rol weggelegd voor Christian Beck, bij wie vrijwel de gehele vertelling de focalisatie ligt. De bijnamen 'de mismaakte' (voor een minnaar van zijn vriendin, o.a. p. 79), 'de asielzoeker' en 'de vogel' komen dus van hem, en het feit dat zij verder zelf nauwelijks 'gekend' of benoemd worden is een direct gevolg van Becks dominantie in het verhaal. Ook irriteert het Beck als hij die naamgevende macht verliest, bijvoorbeeld bij de eveneens naamloze boordeelhoudster in Eilat, die hem voortdurend 'duiveltje' noemt: 'Ik ben geen duiveltje', bijt hij haar toe (98). Maar de asielzoeker verzet zich niet. Het betekent dat hij altijd in de hoedanigheid van 'asielzoeker' optreedt. Hij krijgt niet de gelegenheid om meerdere rollen te spelen of een dubbelzinnige identiteit te vormen; de lezer ziet hem altijd door Becks ogen.

Die beperking van rollen geldt in zekere mate ook voor 'de vogel', maar voor haar ligt het ingewikkelder. Want zij neemt haar eigen bijnaam over:

'Beck noemt zijn vrouw al jaren, eigenlijk al sinds ze elkaar kennen, 'vogel' en ook wel 'vogeltje'. Op een gegeven moment heeft ze de gewoonte ontwikkeld over zichzelf in de derde persoon te spreken. Vooral op de meer gelukkige en intieme momenten. 'De vogel gaat nog even wat water kopen,' kan ze bijvoorbeeld zeggen. Een gewoonte die hem (...) ontroert' (Grunberg 2003: 11).

Momenten waarop de vogel even Becks perspectief overneemt, internaliseert, ervaart hij als intimiteit en wekken bij hem ontroering. Dat zijn dus momenten waarop ze de individuele eerste persoon vervangt voor de afstandelijke derde persoon. Voor Beck ontstaat intimiteit wanneer de vogel zich van zichzelf distantieert en zich in het licht plaatst waarin hij haar graag ziet. Beck ontzegt de asielzoeker en de vogel zo de eigen identiteit en verlangt van hen dat zij zich naar zijn wereldbeeld bewegen voordat – paradoxaal genoeg – intimiteit kan ontstaan en de ander gekend kan worden. Andersom geredeneerd bestaat er dus distantie zolang de vogel (en wellicht ook de asielzoeker) haar individualiteit behoudt. Beck houdt deze afstand bewust in stand, bijvoorbeeld door de asielzoeker te blijven vovoyeren ondanks aandringen van de vogel om 'je' te zeggen (47). Pas na de dood van de vogel zoekt hij enige toenadering: 'Beck is hem [de asielzoeker, LvdD] gaan tutoyeren, maar veel verder dan dat wil hij de intimiteit niet noemen' (Grunberg 2003: 307).

Deze afstand tot de ander wordt niet alleen gecreëerd door de naamgeving. De roman lijkt de blik van Beck te thematiseren. De asielzoeker blijft niet alleen op afstand, hij is vaak niet meer dan een (racistische) karikatuur. Zo wordt de asielzoeker regelmatig als aap en crimineel voorgesteld: 'De man ontbloot zijn gebit voor de tweede keer die dag. Zijn grijns, valt Beck nu op, is een dreigement. Hij heeft niet voor niets in een *gang* gezeten. 'Ik neem een grapefruit,' zegt de man, die Beck telkens weer, op de raarste momenten, aan een aapje doet denken.' (Grunberg 2003: 48). Een grijns als dreigement: het associeert de asielzoeker met een gorilla. Bij zulke typering wordt de focalisatie benadrukt en daarmee de persoonsgebondenheid van deze verbeelding. Vergelijk het zinnetje: 'Het aapje [de asielzoeker, LvdD] kijkt hem verbaasd aan.' (Grunberg 2003: 50). Het focaliserend subject ligt hier besloten in de 'hem': het is niet de asielzoeker die hier waarneemt, maar nog altijd Beck die slechts ziet dat dat 'de nieuwe man van zijn vrouw' naar hem kijkt. In een ander voorbeeld wordt de asielzoeker door Becks blik een crimineel: 'Beck kijkt de asielzoeker aan, een man die ooit voor geld heeft gevochten en god mag weten waarvoor nog meer, een crimineel waarschijnlijk.' (279) Beck construeert een beeld van de asielzoeker door hem te

bekijken, en de lezer kan niet om die blik heen. De hoofdpersoon schrijft datzelfde kijken bovendien toe aan zijn vriendin: wanneer een kennis op de dag van vogels crematie aan Beck vraagt wat ze eigenlijk te zoeken had in een azc, antwoordt hij: ‘Nou, wat ik zei, ze bestudeerde ze. Je hebt er geen verrekijker bij nodig, maar voor de rest zijn het net beesten. En die zijn ook het bestuderen waard zoals je zult weten.’ (300) Beck weet het antwoord eigenlijk niet, maar veronderstelt dat zijn vriendin naar het azc ging om de asielzoekers te bekijken zoals hij hen bekijkt en zoals onderzoekers beesten bestuderen. *De asielzoeker* illustreert dat ‘representatie’ altijd een handelende of representerende instantie vereist. De applicatie formaliseert representaties dus geïsoleerd van de representerende autoriteit. De participant die de formulieren invult kan dus niet anders dan de retoriek van de roman volgen. In het ontwerp van de applicatie blijkt de autoriteit of de mondigheid van de lezer dus betrekkelijk.

Een concreet gevolg daarvan is terug te zien in de personages die ontbreken in de input. De drie personages die ingevuld zijn (Beck, de asielzoeker en de vogel) figureren in een groot deel van de vertelling en de keuze voor deze personages is begrijpelijk. Met de nadruk op deze personages blijft echter ook een aanzienlijk deel van de vertelling buiten zicht. Zoals gezegd wordt de hoofdverhaallijn afgewisseld met flashbacks naar een periode waarin Beck met zijn vriendin in Eilat woonde. De personages uit die flashbacks zien we dus niet terug in de database, maar vanuit een genderkritische of postkoloniale lezing – die zich richt op representaties – is juist deze verhaallijn bijzonder relevant. Deze verhaallijn vertelt hoe Beck regelmatig een Russische prostituee bezoekt, Sosha Minkiewicz, die naar zijn inschatting ziek en lelijk is (Grunberg 2003: 110). Na een van Becks bezoeken aan het bordeel verwijt Sosha hem dat hij haar pijn gedaan heeft. Beck wordt dan zo boos dat hij haar een oog uitsteekt met een schroevendraaier. Beck wordt verhoord door de politie, maar een gerechtelijke veroordeling van zijn mishandeling blijft hem bespaard. Ondanks de centrale rol die Sosha speelt in het verhaal en voor de karaktervorming van Beck, moffelt de tekst haar weg als een onbelangrijke figuur. Beck heeft de vrijheid om prostituees te mishandelen; hij gaat vrijuit. Daarnaast reduceert hij haar – zoals hij deed bij de asielzoeker – tot een dier dat slechts bestaat om met geweld verkracht te worden:

‘Beck zag haar ogen, de ogen van een dier. Niets was beledigender voor haar dan haar nog als mens te herkennen, niets was smadelijker voor haar dan haar vast te houden zoals je een kind vasthoudt. Niets was gemener dan haar ongeneukt achter te laten’ (Grunberg 2003: 178).

De misogynie lading van dergelijke scènes behoeft geen uitleg. Relevant is hier dat een dergelijk personage vanuit de missie van de Personagebank zeer betekenisvol is, maar door de lezer niet opgepikt of herinnerd wordt als een relevante figuur. In een onderzoek naar diversiteit en daarmee ook naar stereotyperingen en rollenpatronen wordt een prostituee niet opgemerkt als die door de tekst gemarginaliseerd wordt. Die conclusie relativeert het vermogen van de applicatie om een kritische of op zijn minst onderzoekende leeshouding over te brengen: het blijft mogelijk ingewikkeld om tegen de retoriek van de tekst in te lezen. Ze nuanceert ook de betrouwbaarheid van de methode: achter elk personage uit de database schuilt mogelijk een personage dat *niet* ingevuld werd.

Tot slot is er nog een derde betekenisdimensie waaraan de applicatie voorbijgaat in haar benadering van een roman als *De asielzoeker*. In een natuurlijk leesproces zullen meer factoren een interpretatie vormen dan aanwijzingen uit de tekst alleen. Een hedendaagse lezer zal een roman als *De asielzoeker* nooit los van de politieke actualiteit kunnen lezen, en ook bij verschijnen in 2003 waren de aantallen asielzoekers zodanig hoog dat hun lot een politiek thema was.¹¹ In het licht van de beeldvorming en angst over recente vluchtelingen – dat Westerse vrouwen niet langer veilig zouden zijn in de nabijheid van deze seksueel gefrustreerde ‘mannen’ – groeit dit verhaal uit tot allegorische proporties: over een man wiens vrouw wordt afgepakt door een asielzoeker. Lezers die Grunberg kennen in de geëngageerde maar vaak cynische rol die hij regelmatig inneemt, zijn daarom misschien geneigd dit verhaal als een politiek

¹¹ CBS cijfers.

commentaar te lezen op de manier waarop West-Europa met haar asielzoekers omgaat. In zo'n lezing wordt de roman een satire, van een samenleving die migranten weliswaar opneemt, maar hen nauwelijks laat integreren, die hen lange tijd zowel sociaal als juridisch als 'asielzoekers' blijft zien, die deze groep op afstand houdt zoals Beck de nieuwe man van zijn vrouw op afstand houdt. Het huwelijk tussen de vogel en de asielzoeker is dan een metafoor voor de institutionele en juridische acceptatie van statushouders tegenover de gelijktijdige sociale uitsluiting. Ondanks zijn huwelijk wordt de asielzoeker nooit een echte echtgenoot, zoals asielzoekers met een verblijfsvergunning misschien moeilijk tot volwaardige medeburgers mogen uitgroeien ondanks de juridische erkenning van hun staatsburgerschap. Als de lezer zo'n interpretatie verkiest, dan is de dominante focalisatie van Beck functioneel, namelijk als een verbeelding van de macht om te representeren in de alledaagse samenleving. In die functie draagt de relatie tussen Beck en de asielzoeker een kritisch potentieel: met narratologische middelen wordt er een karikatuur gemaakt van een buiten-literaire, politieke realiteit. Het moge inmiddels duidelijk zijn dat dergelijke, complexe mechanismen van interpreteren niet begrepen worden in dit onderzoek. Een tekstuele database zoals die door de Personagebank wordt ingezet, stelt onvermijdelijk het woord-niveau voorop. Betekenissen die voorbij of naast het woordniveau bestaan, blijven ongrijpbaar voor de digitale techniek.

Over *Het jongensuur* merkte ik op dat een menselijke lezing van die roman laat zien dat de applicatie geen rekening houdt met intersecties tussen verschillende identiteiten en politieke denkschema's eerder in stand houdt dan bekritiseert. Een nauwkeurige lezing van *De asielzoeker* brengt andere problemen aan het licht. In aanvulling op de al genoemde reserves, evalueer ik de applicatie opnieuw en met nieuwe argumenten.

Ten eerste toonde dit verhaal de functie van namen in romans. Ik liet zien dat namen betekenis kunnen dragen, zowel over de machtsrelaties tussen de naamgever en de genoemde als over de manier waarop een personage gerepresenteerd wordt. De applicatie verzamelt namen van elk personage, maar die namen worden niet geïnterpreteerd. Daarmee gaan mogelijk betekenissen verloren die voor een menselijke lezer evident kunnen zijn. Namen zijn dus een voorbeeld van een betekenisvol aspect van de representatie van een personage dat zich op het lexicale niveau laat interpreteren (anders dan bijvoorbeeld intertekstuele relaties of verwijzingen naar de actualiteit), maar desondanks niet begrepen wordt.

Ten tweede illustreert *De asielzoeker* dat focalisatie een cruciale rol speelt voor de manier waarop personages gerepresenteerd worden. Het formulier registreert niet door wiens ogen de lezer een personage waarneemt, hoewel dat een wezenlijk onderdeel is van de betekenis van zo'n representatie. De applicatie heeft weliswaar relevante informatie over de asielzoeker – afkomstig uit Noord-Afrika, nevenpersonage, tussen 26 en 35 etc. – maar zijn identiteit wordt in hoge mate geconstrueerd door het personage Christian Beck. De kwantitatieve en geformaliseerde benadering van dit onderzoek suggereert een zekere meetbaarheid en objectiviteit, maar is wel gebonden aan de sterke subjectiviteit van de tekst. De lezer mag mondig heten, maar de applicatie moedigt haar aan om simpelweg de retoriek van de tekst te volgen. Als 'de asielzoeker' immers uitsluitend door de ogen van Beck wordt waargenomen, rest de lezer niets anders dan te leunen op de aanwijzingen die Beck haar geeft over de identiteit van de asielzoeker.

Ten derde heeft deze volgzzaamheid van de lezer aan de tekst consequenties voor de data die wel en niet in de database terechtkomen. Vanuit het streven van het project om diversiteit en representatie in kaart te brengen, was de Russische prostituee Sosha Minkiewicz een zeer betekenisvol personage, juist omdat zij gemarginaliseerd wordt, als vrouw en slachtoffer van (seksueel) geweld. De lezer die de input leverde, vulde haar echter niet in. Daarvoor zijn ongetwijfeld talloze verklaringen denkbaar die niet meteen met de narratologische positie van dit personage te maken hebben. Punt is hier dat de applicatie op dit vlak niet meet wat ik ermee hoop te meten, en dat is mogelijk een gevolg van de relatieve onmogelijkheid van de lezer om zich kritisch te verhouden ten opzichte van de representaties zoals de tekst die aan haar presenteert. Met haar vooropstelling van tekstuele aanwijzingen biedt de applicatie bovendien geen alternatief. De Personagebank manoeuvreert zich zo in de merkwaardige positie waarin zij participanten indirect aanspoort om significante gegevens achterwege te laten.

Een laatste inzicht dat ik opdeed met mijn lezing van *De asielszoeker* betreft de mogelijke satirische of ironische betekenis van personagekenmerken. Wanneer de applicatie de verhoudingen in de database berekent en interpreteert, gaat ze ervanuit dat die data eenduidig en interpreteerbaar zijn. Satirische teksten maken echter gebruik van de meerduidigheid en dubbelzinnigheid van taal, haar vermogen om A te zeggen en B te communiceren. Deze metaforische potentie van taal is een klassiek probleem voor digitaal begrip van taal en dwingt tot enige reserves bij de geprogrammeerde interpretaties van verhoudingen in de database zoals die op de Resultatenpagina getoond worden.

5.4 *Het gym*

Het gym van Karin Amatmoekrim (2011) vertelt over het Surinaamse meisje Sandra uit een anonieme achterstandswijk, ‘de wijk’. Haar moeder zit in de bijstand en haar vader ziet haar niet staan wanneer hij tijdelijk zijn nieuwe vrouw verlaat om weer thuis te komen wonen. Sandra mag als enige uit haar omgeving en vriendenkring na de basisschool naar het gymnasium, ‘het gym’, gelegen in een nabijgelegen rijkeldorp. De confrontatie met haar klasgenoten uit een spierwitte middenklasse brengt haar in een identiteitscrisis en een loyaliteitsconflict tegenover haar mede-‘wijkenaren’.

De lezer die *Het gym* las, vulde één personage in. Haar kenmerken staan in tabel 2.

Tabel 2. Lezersinput van *Het gym* van Karin Amatmoekrim

Kenmerk	Personage 1
Naam	Sandra
Rol	Verteller
Gender	Vrouw
Leeftijd	0-25
Herkomst	Zuid-Amerika overig
Woonplaats (land)	Nederland
Woonplaats	Onbekend
Woonplaats (type)	Stad
Opleiding	Hoogopgeleid
Beroep	Scholier

Opnieuw vormen deze kenmerken een accurate representatie van Sandra’s identiteit zoals de roman die verbeeldt. Suriname, Sandra’s land van herkomst, valt onder ‘Zuid-Amerika overig’, ze is een vrouw van 13-14 jaar en woont in een onbekende stad. Haar rol als verteller is weliswaar discutabel en ook de typering ‘hoogopgeleid’ is nog niet zo van toepassing op een brugklasser, maar de keuze voor deze kenmerken is evenwel zeer begrijpelijk. Sandra is namelijk geen ik-verteller, maar de focalisatie ligt wel uitsluitend bij haar. Daarnaast worden Sandra’s aanleg om te leren en intelligentie gethematiseerd in deze roman, dus in dat licht is de typering ‘hoogopgeleid’ zo gek niet. Hoe dan ook blijkt opnieuw dat deze lezer goed nagedacht lijkt te hebben en de vragen serieus genomen heeft.

In de twee romananalyses hierboven zijn al verschillende problemen aangekaart met de aannames en wetenschappelijke consequenties van deze methode. Een analyse van *Het gym* kan daar in mijn ogen twee overwegingen aan toevoegen. Een eerste aanmerking hangt samen met het probleem dat ik eerder signaleerde over de input van *De asielszoeker* en dat zich voordoet als een lezer personages achterwege laat. In dit geval is slechts één personage gekarakteriseerd. Voor deze specifieke roman is dat een ingewikkelde oogst, omdat Sandra’s identiteit voornamelijk in relatie tot andere personages vorm krijgt. Sandra is hier niet zomaar een ‘vrouw’ in de leeftijd van ‘0-25’ uit ‘Zuid-Amerika overig’, maar ze is vooral *niet* wit, *niet* middenklasse en van buiten het dorp van haar gymnasium. De roman thematiseert die tegenstelling en

staat bol van scenes waarin Sandra's achtergrond en identiteit botsen met die van haar nieuwe omgeving. Zo'n moment van (letterlijke) zelfreflectie vindt plaats wanneer Sandra met haar nieuwe vriendinnen voor de spiegel staat op het toilet van de school. Als Sandra geïrriteerd tegen hen uitvalt dat ze moeten ophouden met het vragenvuur waaraan ze onderworpen wordt, antwoordt haar vriendin Mirte: "Je bent gewoon heel anders dan wij!" (59). Voor de constructie van Sandra's identiteit en andersheid zijn haar tegenspelers dus cruciaal. Het is daarom misleidend als die tegenspelers volledig ontbreken in het formulier. Een menselijke lezer zou Sandra immers altijd beschrijven in relatie tot de wereld van het gymnasium. Voor de applicatie komt identiteit ook relationeel tot stand, maar dan via de verhouding tot andere personages uit de database. Dat duidt dus op een hermeneutisch verschil tussen menselijke en computationele interpretatie van deze specifieke roman.

Een ander probleem van een digitale verwerking van *Het gym* sluit aan bij wat ik schreef over *Het jogensuur*, dat deze applicatie uitgaat van een beperkte invulling van het concept 'diversiteit'. De Personagebank laat personages classificeren op basis van hun 'herkomst', dat in dit onderzoek een synoniem is voor 'etniciteit'. Die benadering van personages is adequaat voor deze roman, want het onderscheid op basis van herkomst is hierin zeer relevant. Dat blijkt bijvoorbeeld uit Sandra's overdenkingen – in de vrije indirecte rede – over racisme en uitsluiting door witte Nederlanders. Een voorbeeld:

'Sandra en haar moeder waren wel een beetje zwart, maar niet helemaal. Dat verschil was heel subtiel. Je kon te zwart zijn, zoals Tanya, en dan wilden witte mensen niet met je omgaan. Als je een beetje meer bruin was, zoals Sandra's familie, dan was je oké. Zolang je tenminste altijd even aardig deed en het niet te veel had over dingen die niet Nederlands waren' (Amatmoekrim 2011: 12).

Sandra ervaart dan ook regelmatig hoe zij gereduceerd wordt tot een groepsidentiteit of een stereotype dat men baseert op haar herkomst. Hoe die scheidslijnen tussen verschillende groepen getrokken worden, blijkt bijvoorbeeld ook uit het volgende fragment:

"Ik wil gewoon werken. Maar hun pikken al onze banen in, en zo is dat." Hij wierp een blik op Sandra. "Sorry, meis, maar zo is het wel."

Chantals moeder ging zuchtend aan tafel zitten. Ze keek Sandra met een verontschuldigend lachje aan. "Hij heb het niet over jou, hoor. Dat weet je, toch? Jij bent anders als hun. Jij en je ma zijn geen echte buitenlanders'" (Amatmoekrim 2011: 203).

Chantals vader identificeert Sandra niet als het vriendinnetje van zijn dochter, maar als lid van een groep Nederlanders die zijn baan afpakt. Vanuit deze denkschema's redeneert ook de Personagebank, met zijn indelingen van personage op grond van hun herkomst. In die zin is de methode in lijn met haar object en kan ik erop vertrouwen dat haar metingen betekenis hebben. Maar tegelijkertijd komt Sandra herhaaldelijk tot de conclusie dat haar andersheid niet geconstrueerd wordt door haar afwijkende etnische herkomst, maar vooral ook door de sociaaleconomische klasse waartoe ze behoort. Meer nog dan een verhaal over botsende etnische groepen is *Het gym* zo een roman over klassenverschillen. Een centraal citaat is het volgende:

'Hij had haar zwakke plek feilloos gevonden. Niet het feit dat ze een buitenlander, allochtoon of nieuwe Nederlander was. Haar kleur was op het gymnasium niet onoverkomelijk. Nee, haar zwakke plek was de wijk, die armoede met zich meezeulde in de naam, in de krantenberichten over zwarte scholen en kansarme gezinnen, in de aanblik van die hoge grijze flats en de uitgestrektheid van de parkeerplaatsen eromheen. Bart Willink had haar zwakke plek nog eerder gevonden dan zijzelf zich ervan bewust was geweest' (Amatmoekrim 2011: 69).

De diagnose van dit verhaal lijkt dus te zijn dat de meritocratie nog altijd sterker is dan discriminatie op basis van huidskleur ('haar kleur was op het gymnasium niet onoverkomelijk'), maar dat er vooral wantrouwen blijft bestaan jegens mensen uit andere sociale klassen. Die vaststelling wordt ook gevoeld door Sandra's ontmoeting met een andere zwarte jongen op het gymnasium:

'Een geadopteerde neger. Nu snapte ze het. Een geadopteerde neger was niet hetzelfde als een gewone neger. Een geadopteerde neger uit een rijke kakfamilie was gewoon een rijke kakker. Zijn kleur maakte hem niet hetzelfde als zij. Zijn geld maakte hem hetzelfde als kakkers'
(Amatmoekrim 2011: 81).

Geld bepaalt identiteit, niet huidskleur. Vanuit dit standpunt is sociale klasse een grote afwezige factor in dit onderzoek naar sociale diversiteit in de Nederlandstalige romanliteratuur. Twee onderdelen van het formulier kunnen weliswaar indicatoren zijn van sociale klasse (opleidingsniveau en beroep), maar *Het gym* laat zien dat een meer expliciete variabele als 'klasse' niet misstaan had.

Deze twee kanttekeningen bij *Het gym* geven samen aanleiding tot een meer fundamentele kritiek op de werking van de Personagebank. De paradox die de input over deze roman blootlegt is namelijk dat een roman die de diversiteitsproblematiek en de macht van representatie thematiseert in de relaties tussen personages niet als zodanig waargenomen wordt in een onderzoek naar diversiteit en representatie in literatuur. In dit onderzoek is *Het gym* slechts een roman over een meisje uit een anonieme stad met een Zuid-Amerikaanse herkomst. De focus op kwantificeerbare en digitaal verwerkbare gegevens is niet alleen een focus; voor de applicatie hebben uitsluitend kwantificeerbare patronen betekenis. Individuele romans die interessante inzichten kunnen opleveren voor het onderzoek, worden niet als relevante casus herkend. Dat betekent allerminst dat dit onderzoek zinloos is: ik was immers nadrukkelijk op zoek naar kwantificeerbare en controleerbare tekstkenmerken die de diversiteit in vele verhalen bij elkaar inzichtelijk konden maken. Ik moet mij niettemin bewust tonen van de theoretische en methodologische implicaties van het feit dat personages en romans slechts nog instrumenteel zijn. Wanneer de romans verwerkt worden door de lezer en daarna door de applicatie, worden de personages ontdaan van hun naam, hun politieke betekenis, hun satirische of actuele lading en hun context. Daarbij komt dat de applicatie blind is voor allerlei facetten van 'identiteit' zoals menselijke lezers dat concept begrijpen. Ook daarvan heb ik verschillende voorbeelden genoemd: de intersecties tussen verschillende groepsidentiteiten, sociaaleconomische klasse, de relaties tot andere identiteiten. Een laatste probleem is nog dat personages niet in de samenhang van hun kenmerken geïnterpreteerd worden. In plaats daarvan berekent de applicatie verhoudingen en frequenties van kenmerken afzonderlijk. De Resultatenpagina toont dus verhoudingen van mannen en vrouwen, van hoogopgeleide en laagopgeleide personages en van jonge en oudere karakters. Sommige intersecties worden wel getoond, zoals het verschil tussen mannelijke en vrouwelijke beroepen, of het verschil in opleidingsgraad tussen mannen en vrouwen. Maar ook daarvan geldt dat de database als geheel geïnterpreteerd wordt – en niet als een verzameling intern samenhangende personages. De applicatie reduceert alle personages tot een reeks verwerkbare gegevens. Het gevaar van zo'n hermeneutiek van crowdsourcing is dus dat het digitale denken en het computationeel kenbare een grote stempel drukken op de betekenis en kennis over de literatuur die het project moet voortbrengen. Informatie die digitaal begrepen wordt krijgt mogelijk een grotere legitimiteit dan kennis die een individuele, menselijke lezer voortbrengt, en dat laat zich epistemologisch gezien moeilijk verantwoorden.

5. Receptie en procesevaluatie

Op 18 april 2016 ging de Personagebank officieel live met een begincorpus van 100 romans en 620 personages. Zeven weken later, op 6 juni 2016, stond de teller op 176 romans en 821 personages. Daarbovenop zijn 5 romans en 10 personages verwijderd vanwege dubbele inputs (onder andere de romans *De avonden* van Gerard Reve en *De ontdekking van de hemel* van Mulisch werden twee keer en met verschillende ISBN-nummers ingevuld). Als ik die bijdragen wel meereken, kom ik op een totale lezersinput via crowdsourcing van 81 romans en 211 personages in zeven weken tijd. Dat is een gemiddelde van 11.6 romans en 30.1 personages per week, hoewel ik moet opmerken dat bijna de helft van de input in de eerste week werd toegevoegd (40 romans). Na het momentum van de lancering zakte de respons dus, maar nog altijd is die niet stilgevallen. Een nieuwe promotiecampagne kan de participatie vermoedelijk weer een impuls geven. Ik schat het aantal unieke respondenten op 50 (sommige lezers vulden meerdere romans in, te zien aan bijdragen die kort na elkaar gedaan worden).

Op de 18^e april begon de promotie op facebook (120 likes) en twitter (50 followers). Een week later verstuurde ik een persbericht naar redacties van kranten, literaire tijdschriften en radioprogramma's. De vrijdag ervoor had Arjen Fortuin al over de Personagebank geschreven in zijn column voor *NRC Handelsblad*. Na de persberichten volgden interviews voor BNR Nieuwsradio en RTV Utrecht, er verschenen stukken in *Opzij* en *Knack* en op verschillende blogs en fora werd het onderzoek besproken. De reacties in de verschillende media zal ik hieronder kort analyseren.¹²

6.1 Reacties in de media

In dit onderzoek staan twee vragen centraal:

- (1) *In hoeverre kan crowdsourcing via geformaliseerde romaninterpretaties valide en generaliseerbare kennis voortbrengen over een specifieke aspect (representaties van personages) van deze romans en (2) in welke mate bemoeilijkt een verschil in leeshouding onder de participanten deze generaliseerbaarheid?*

Ik heb geprobeerd een voorlopig antwoord te geven op vraag 1 door hierboven enkele romans uit de database, door lezers ingevuld, opnieuw te lezen en te analyseren. Voor een antwoord op vraag 2 is echter een andere aanpak nodig. Dit onderzoek is geen empirische literatuurwetenschap in de traditionele vorm, waarin lezers geënquêteerd worden over hun ervaring van of attitudes tegenover bepaalde literaire teksten. Om principiële redenen en conform mijn zesde ethische criterium voor het ontwerp, slaat de applicatie geen enkele data op over de identiteit of het gedrag van gebruikers. Ik weet dus niet wie welke romans heeft ingevoerd, laat staan waarom, wat hun leeshouding was tijdens het lezen of hun ideeën over een project als de Personagebank. Het is daarom mogelijk ambitieus om toch iets te willen zeggen over de leeshoudingen van de participanten. Ik doe hieronder toch een poging door verschillende reacties op het project, op social media, blogs en in de krant, te beschrijven en te duiden. Die reacties zeggen iets over de mate waarin deze mondige lezers bereid zijn de analytische leeshouding aan te nemen waartoe de Personagebank aanzet. Ik beschouw hun reacties als een indicatie voor de manier waarop grotere groepen lezers met een dergelijk project om willen gaan.

Om te beginnen valt op dat sommige critici wel degelijk welwillend lijken te staan tegenover het project. *NRC*-recensent Thomas de Veen (18 april) en zijn voormalige 'chef Boeken' Pieter Steinz (18 april) deelden een link via Twitter (De Veen: 'Hé, interessant. De @Personagebank gaat data crowdsourcen op zoek naar de vraag: hoe divers is de NLse literatuur?'). Diezelfde week wijdde Arjen

¹² Zie bijlage 3 voor een overzicht of: <http://www.uu.nl/nieuws/veel-media-aandacht-voor-personagebank>

Fortuin zijn vaste column in de wekelijkse boekenbijlage van het *NRC Handelsblad* aan de Personagebank. Hij verbindt daarin verhoudingen in de database aan een poëticaal verlangen: ‘Zo overviel mij bij de Librisnominaties (...) vooral een groot verlangen naar meerstemmigheid of, ouderwetsser nog, een alwetende verteller. Romans die ons niet een mens willen tonen, maar een wereld. Diep doordringend in véél personages’ (Fortuin 2016). Verder schrijft hij: ‘Intussen maak ik me zorgen om het personage in de hedendaagse literatuur. Niet om de hoofdpersoon, maar om de rest. Nederlandse schrijvers kruipen steeds liever “in de huid” of “in het hoofd” van één hoofdpersoon. (...) Dat kan mooie boeken opleveren en niet elke verteller is hetzelfde (...), maar ik kon een gevoel van monotonie niet onderdrukken’ (Fortuin, 22 april 2016). Fortuin rekent zich met deze subtiele onvrede eerder tot het kamp van Anil Ramdas (schrijvers leveren een wanprestatie door zich niet in te leven in meerdere identiteiten) en de voorstanders van heteronome literatuur dan dat hij een zuivere poëtica uitdraagt zoals velen van zijn collega-critici vaak deden. Deze critici van *NRC Handelsblad* – toch geen onbelangrijke medium in de literatuurkritiek – lijken nauwelijks een grens te ervaren met de interpretatieve gemeenschap van de academische kritische traditie.

Enkele dagen later verschijnt er op de website van *Opzij* een stuk waarin Aniek Hofkens de Chef cultuur van het tijdschrift, Femke van Wiggen, interviewt over de eerste resultaten van de Personagebank. Ook Van Wiggen is positief:

‘Het is heel goed dat dit gedaan wordt. Ik denk dat we ons helemaal niet bewust zijn dat in de meeste boeken de wereld vanuit een mannelijk perspectief wordt benaderd. Pas wanneer we dit beseffen, kun je onder ogen zien dat er ook verhalen zijn die niet verteld worden. Dat zijn verhalen die gelijkwaardig zijn aan de ‘mannelijke’ verhalen en die samen een genuanceerder beeld van de werkelijkheid geven’ (Hofkens 2016).

Opvallend is Van Wiggens overtuiging dat literaire beeldvorming de werkelijkheid vormt en dat Nederlandstalige romans kennelijk nauwelijks mimetisch (willen) zijn:

‘Het is bizar dat uit deze pilot blijkt dat vrouwelijke personages vooral student, scholier, huisvrouw en prostituee zijn. Dit zijn natuurlijk pas de eerste resultaten, maar het geeft al wel een zorgelijk beeld. Je kan mij niet wijsmaken dat volwassen vrouwen in werkelijkheid vooral huisvrouw en prostituee zijn, maar deze boeken uit 2013 scheppen wel dat beeld’ (Hofkens 2016).

Met dit besef keert Van Wiggen zich tegen het ‘zuivere’ literatuurideaal. Volgens dat geheel aan normen hebben schrijvers immers de artistieke vrijheid om personages vorm te geven zoals zij hen voor zich zien. Het ideaal stelt dat men niet van de literatuur mag verwachten dat zij een accuraat sociologisch beeld van de werkelijkheid schept, want dat is niet haar doel. In de traditie van de zuivere poëtica zou Van Wiggens vaststelling vermoedelijk gelden als een denkfout, een inadequate reactie op literatuur (vgl. Zwagerman 1997; Hermse 2013; Peeters 2014). Voor de interpretatieve gemeenschap van *Opzij* geldt het tellen en classificeren van personages echter als vanzelfsprekend. En die gemeenschap strekt zich uit over zowel de journalisten als de lezers, getuige de respons op het artikel op de Facebookpagina (29 april 2016) van het blad.¹³ Die was relatief hoog in vergelijking met andere berichten op de pagina: het werd 13 keer gedeeld, 37 keer geliked en er werd veelvuldig gereageerd. De *Opzij*-lezer vindt het niet vreemd om te letten op de rollenpatronen in de literatuur en lijkt eraan gewend om representaties kritisch te beschouwen.

Een meer kritische houding werd uitgedragen door Maarten Dessing, die over de Personagebank schreef op de website van *Knaak*. Dessing doet verslag van de doelstelling en resultaten van het project en besteedt veel aandacht aan de methodische kwesties en de wetenschappelijke potentie van de database. Zo is hij kritisch over de mogelijke generaliseerbaarheid van de gegevens:

¹³ <https://www.facebook.com/opzij>

‘Maar wat als tussen de geselecteerde boeken toevallig een historische roman zit waarin de schrijvers het vrouwbeeld aan het einde van de negentiende eeuw - toen vrouwen alleen hoer of huisvrouw konden zijn - bekritiseert? En een satire over snuivende en hoererende politici?’ (Dessing 2016).

Los van de vraag of vrouwen in de negentiende eeuw daadwerkelijk alleen maar hoeren of huisvrouw konden zijn, benoemt Dessing dus terechte methodologische tekortkomingen van het onderzoek. De valkuilen die hij aanwijst zijn echter niet fundamenteel van aard: in beginsel onderschrijft Dessing deze literatuurbenadering. Hij betwijfelt hooguit in hoeverre deze specifieke resultaten een ideologiekritiek (in dit geval van een bepaald vrouwbeeld) legitimeren. Er is volgens hem geen sprake van schadelijke representaties als ze historisch of satirisch zijn, met andere woorden wanneer schrijvers bewuste keuzes gemaakt hebben bij de precieze identiteiten die ze hun personages gaven.

6.2 Reacties op blogs en fora

Een soortgelijke kritiek verwoordde schrijver Mark Boog in een blogpost op 11 mei 2016. Hij reageert op een van de interpretaties bij een tabel van meest voorkomende beroepen onder vrouwelijke en mannelijke personages op de Resultatenpagina. Die zin luidt: ‘De werkvloer in de fictieve ziekenhuizen is exemplarisch: artsen zijn vooral mannen, verpleegkundigen vrouwen.’¹⁴ Deze resultaten raken Boog, omdat hij zelf een roman schreef over de ziekte van zijn zoontje en waarin veel mannelijke artsen en vrouwelijke verpleegkundigen voorkomen. ‘Het boek is, helaas, gebaseerd op werkelijke gebeurtenissen’, stelt hij, ‘Het aantal mannen en vrouwen erin strookt ook met de werkelijkheid zoals ik die heb ervaren.’ Dat leidt Boog tot de vaststelling: ‘De hierboven aangehaalde observatie is dus wel van toepassing op *Het lot valt altijd op Jona*, maar daaruit zijn eerder maatschappelijke dan literaire conclusies te trekken.’ Literaire representaties zijn legitiem en onproblematisch zolang ze waarheidsgetrouw zijn, zo lijkt Boog te zeggen.

Naast deze critici en schrijvers vond ik meerdere reacties van ‘reguliere’ lezers in een forumdiscussie op Boekmeter.nl naar aanleiding van het artikel in *Opzij*. Als ‘Donkerwoud’ (man, 31 jaar) het artikel deelt op 29 april, reageert ‘liv2’ (vrouw, 56 jaar) diezelfde dag nog:

‘ik lees al lang, en hier ga ik absoluut niet mee akkoord. De schrijver/ster van dit artikel moet ofwel meer boeken lezen/ ofwel eens anders gaan lezen. De mooiste literatuur gaat meestal over de aantrekking tussen man en vrouw. Het moeilijke pad vooral. Meestal wordt er, door goede schrijvers, heel liefdevol omgegaan met dit gegeven. (...) Meestal gaat het over onmacht, de hunkering en diepe liefde voor de vrouw. Misschien ook kwaadheid, maar dat mag.’¹⁵

Deze lezer trekt de resultaten dus in twijfel op basis van zijn of haar eigen leeservaring. Hij geeft hier bovendien een literair oordeel: goede literatuur zijn verhalen die de aantrekking tussen man en vrouw thematiseren. Interessant is wel dat ze lijkt te erkennen dat het mannelijke, heteroseksuele perspectief dominant is. Het gaat om de ‘onmacht, de hunkering en diepe liefde voor *de vrouw*’ [mijn cursivering, LvdD].

Een andere lezer, ‘eRCee’ (man, 32 jaar), stelt:

‘Vind het nogal een eigenaardig type onderzoek. Met literatuurkritiek heeft het alvast niks te maken. Ook de kunst moet blijkaar worden uitgedrukt in maat en getal. Het kan natuurlijk gezien worden als sociaal onderzoek, maar ik vraag me af of je daar uit gaat komen. Als de manier

¹⁴ www.personagebank.hum.uu.nl/resultaten.xq

¹⁵ <http://www.boekmeter.nl/forum/3/759/0#141995>

waarop vrouwen in romans figureren inderdaad op deze manier gekleurd is dan zal dat meest waarschijnlijk een representatie zijn van bepaalde sentimenten in de samenleving, en niet andersom. Blijft over of vrouwelijke auteurs duidelijk anders benaderd worden door de literatuurkritiek. Daar zou je in kwantitatieve zin wellicht iets zinnigs over kunnen zeggen, en daarmee houdt het denk ik wel op.¹⁶

eRCee verwoordt een algemeen en belangrijk ongemak met veel benaderingen uit de digital humanities in het algemeen, dat zelfs de kunst nu wordt uitgedrukt ‘in maat en getal’. Ook de rest van haar/ zijn reflectie vind ik waardevol: de lezer acht het onderzoek weinig relevant, maar verwoordt wel dat mogelijke representatiepatronen (indien valide aangetoond) ‘een representatie zijn van bepaalde sentimenten in de samenleving’. Die overtuiging gaat dus in tegen het zuivere literatuurideaal waarin de schrijver de volledige artistieke vrijheid geniet, en erkent dat culturele discoursen, verwachtingspatronen of ‘sentimenten’ uit de niet-literaire wereld de literatuur kleuren. Deze lezer benoemt vervolgens dat het haar of hem belangwekkend(er) lijkt om te kijken naar ‘genderbias’ in de literatuurkritiek. De lezer staat dus sceptisch tegenover het onderzoek, maar toont tegelijkertijd de bereidwilligheid om een kritische leeshouding aan te nemen ten aanzien van literatuur of literatuurkritiek.

Dat ligt anders bij ‘The Cheshire Cat’ (man, 42 jaar):

‘Puberaal gezeur vind ik het. In de echte (westerse) wereld is zowat alles al bereikt, dus gaan ze het nu maar in de literatuur zoeken. Sowieso is de uitkomst van dit onderzoek bij voorbaat al negatief, want het is duidelijk dat feministen alleen maar dat zien wat ze willen zien.’¹⁷

Volgens deze lezer is het onderzoek dus bevooroordeeld door de activistische idealen van de onderzoeker(s). Voor hem kent deze benadering dus geen enkele relevantie of legitimiteit. Zijn reactie lijkt voornamelijk bepaald door zijn houding ten aanzien van het feminisme (irrelevant want emancipatie is voltooid in de westerse wereld), maar dat is dan ook een belangrijk resultaat: attitudes tegenover ideologiekritische of activistische tradities sturen bij sommige lezers en althans in dit geval in sterke mate de houding tegenover dit onderzoek en bepalen de bereidwilligheid om te participeren. Hij zal niet snel een kritische leeshouding overnemen.

Eenzelfde mechanisme zie ik in de reactie van Pantha Rei (vrouw, leeftijd onbekend):

‘Dat ik op mijn eigen manier in het leven sta, heb ik aan het feminisme te danken. Het echte feminisme bedoel ik dan, niet die radicalinski's die van allerlei triviale zaken een enorme maatschappelijke betekenis toekenden.

Zie bijv. het bericht in de OP [*Opzjj*, LvdD]. Dat soort geneuzel dus.

Opzjj zou zich eens moeten verdiepen in de zgn. flutromannetjes. Harlekijn en boeketreeks dus. Deze boekjes hebben een ware evolutie beleefd.

In mijn jonge jaren waren deze boekjes heel hoofs: er kwam geen spatje sex in voor hooguit even elkaars hand vasthouden en een zeldzame kuise zoen.

Later veranderde dat in een grote sexpartie waarbij de heldin het ene orgasme na het andere had. En nu is het de vrouw die de dingen bepaald; ze is een afgestudeerde bolleboos, alleenstaande moeder met een drukke baan en een soort vrijetijdsdetective.

De tijden zijn echt veranderd wat dat betreft.¹⁸

¹⁶ <http://www.boekmeter.nl/forum/3/759/0#142000>

¹⁷ <http://www.boekmeter.nl/forum/3/759/0#142051>

¹⁸ <http://www.boekmeter.nl/forum/3/759/0#142058>

Interessant aan deze reactie is het feit dat deze lezer mijn onderzoek en de discussie erover in *Opzij* ‘geneuzel’ noemt, maar als argument een literaire analyse aandraagt van de historische ontwikkeling van het vrouwelijk personage in de boeketreeks. Ze beschrijft de recente emancipatie van de vrouw aan de hand van de boeketreeks door de jaren heen, een vorm van *cultural analysis* die niet had misstaan binnen de universiteitsmuren. Haar kritiek richt zich dus niet op de methode of leeshouding zelf, maar op de interpretatie of corpusselectie van mijn onderzoek.

Deze samenvatting van enkele reacties geeft slechts een indruk van mogelijke reacties op dit project en geeft allerm minst een alomvattend beeld van de werkelijke receptie onder verschillende typen lezers. Niettemin geeft deze receptie aanleiding tot twee conclusies. Ten eerste blijkt dat het te schematisch is om te stellen dat ‘de literaire kritiek’ nauwelijks bereid is tot ideologiekritische lezingen en dat de zuivere literatuuropvatting regeert op krantenredacties. Geen van de critici van het *NRC Handelsblad* die publiekelijk reageerden op de Personagebank toonden zich sceptisch en leken, sterker nog, ontvankelijk voor deze vorm van literatuuronderzoek. Ook de geluiden in *Opzij* en *Knack* – beide tijdschriften besteden kolomruimte aan literatuur – geven geen enkele blijk van het ‘zuivere’ literatuurideaal.

Ten tweede kan ik stellen dat niet alle lezers vanzelfsprekend bereid zijn om de kritische leeshouding aan te nemen die de applicatie stimuleert. De discussies op blogs en fora tonen dat lezers sceptisch staan tegenover dit onderzoek als hun attitude over ideologiekritische tradities zoals het feminisme negatief is. Uit de procesevaluatie blijkt echter ook dat er genoeg lezers zijn die de gevraagde leeshouding probleemloos overnemen: 81 romans en 211 personages werden ingevuld door lezers, en dat aantal groeit nog elke dag. Vooralsnog bleek geen enkele participant een ‘trol’¹⁹ en leek alle input serieus ingevoerd. Lezers die sceptisch staan tegenover het onderzoek doen vermoedelijk simpelweg niet mee. Gebruikers die wel meedoen, doen dat serieus en met aandacht. Er is dus nauwelijks sprake van een vervuiling van data. Vanwege de veronderstelde positieve, bereidwillige houding van de gebruikers die meedoen, is optimisme over de generaliseerbaarheid van de ingevoerde data gerechtvaardigd. De participanten hadden zonder twijfel verschillende leeshoudingen bij het lezen van de romans, maar voor hun bijdrage aan het onderzoek waren zij bereid om tijdelijk de lees- of interpreteerhouding aan te nemen die de applicatie van hen vroeg. Dat resultaat bevestigt Felski’s eerdergenoemde stelling dat de leeshoudingen van onderzoekers, critici en andere lezers uiteindelijk veel overeenkomsten vertonen. In dit onderzoek bleken deze verschillende groepen bereid om tijdelijk toe te treden tot dezelfde ‘interpretatieve gemeenschap’, en daardoor lijkt de generaliseerbaarheid en validiteit van de data in hoge mate en in veel contexten (wetenschappelijk, journalistiek, recreatief, activistisch) gegarandeerd.

¹⁹ Internetslang voor een gebruiker die een applicatie of forumdiscussie doelbewust ontregelt of provokeert. Zie ook: <http://www.urbandictionary.com/define.php?term=troll>

7. Conclusie

In dit onderzoek heb ik geprobeerd de rol aan te nemen van wat Stephen Ramsay een ‘algorithmic critic’ noemde. Ik ging op zoek naar ‘some type of machinery’ – in de woorden van digital humanities *godfather* Roberto Busa – om een complex letterkundig probleem op te lossen (geciteerd in Ramsay 2011: 85). Aanleiding was een poëtische discussie over diversiteit in de Nederlandstalige literatuur. Het probleem dat ik wilde oplossen betrof de vraag hoe we neutrale of intersubjectieve gegevens over literaire representaties kunnen verkrijgen binnen een debat dat gepolariseerd lijkt langs poëtische scheidslijnen. Door die poëtische kampen leek een neutrale literatuurwetenschap – een wetenschap die voor beide kampen valide kennis zou produceren – nauwelijks meer mogelijk. Mijn voorstel was een kwantitatieve literatuurbenadering door middel van crowdsourcing, waarin ik grote groepen lezers zou oproepen om personages te classificeren in tien eigenschappen met als doel om betekenisvolle patronen te vinden. Die aanpak had twee voordelen: ten eerste gaf de oproep aan ‘de crowd’ de gegevens een zekere intersubjectieve geldigheid en een bredere generaliseerbaarheid. Ten tweede biedt deze methode in potentie de mogelijkheid om een groot corpus van romans efficiënt te onderzoeken op personagekenmerken.

Mijn operationalisering kreeg concreet de vorm van een crowdsourced database en webapplicatie, ‘Personagebank’, waarmee in 7 weken tijd 81 romans en 211 personages werden onderzocht door reguliere lezers. Samen met het begincorpus ontstond er een database van 179 romans en 821 personages, die nog altijd groeit. Deze database biedt in mijn ogen een cruciaal onderzoeksinstrument voor wat mijn begeleider dr. Saskia Pieterse een ‘sociologie van de verbeelding’ heeft genoemd: een letterkundige of cultuuranalytische discipline die zich bezighoudt met sociale verhoudingen en rollen in fictieve werkelijkheden, om zo op grote schaal representatiepatronen te begrijpen in transhistorisch en transgeografisch perspectief. Daarnaast maakt de database kruisverbanden, intersecties of correlaties tussen personagekenmerken zichtbaar die op hun beurt vragen om kwalitatieve bestudering van de bewuste teksten. De data worden dus zowel exploratief als descriptief gepresenteerd. Ik beschouw de Personagebank dan ook als een voorbeeld van wat Stephen Ramsay ‘algorithmic criticism’ noemde, waarin de subjectieve fijngevoeligheid van een menselijke lezer gecombineerd wordt met de relatieve objectiviteit of neutraliteit die de rekenkracht van de computer mogelijk maakt. In een overzicht en korte duiding van de verschillende datavisualisaties die de applicatie publiceert, toonde ik waar de wetenschappelijke kansen liggen voor zo’n methode en dataset.

Maar zoals alle methode kende ook de mijne haar beperkingen. In het theoretisch kader heb ik de aannames geïmpliciteerd die aan de applicatie ten grondslag lagen. Centraal stonden de hermeneutiek en epistemologie die crowdsourcing in de letterkundige context veronderstelt. In reactie op de hermeneutische traditie en enkele invloedrijke denkers uit de digital humanities kwam ik tot de hypothese dat er fundamentele verschillen zijn tussen menselijk denken en computerlogica. De problemen die uit die verschillen voortkomen voor de inzet van digitale instrumenten in de literatuurwetenschap zijn echter niet onoverkomelijk. Het is ten eerste de kunst om te erkennen dat de computer patronen niet alleen vooropstelt, maar *uitsluitend* patronen begrijpt. Dat geeft een serieuze verstoring van de informatie die bovendien niet altijd herkend wordt in een wereld die geregeerd wordt door computationele mechanismen en retoriek. Om die verstoring te minimaliseren is het noodzakelijk dat de hiaten in het computationele waarnemen opgevuld worden met de hermeneutische kwaliteiten van de menselijke lezer. In een analyse van drie romans uit de database analyseerde ik de kloof tussen enerzijds het computationele interpreteren van de applicatie en anderzijds het menselijke interpreteren van mij als onderzoeker en in mindere mate van de participerende lezer. In dit tweesporenmodel – analyse van zowel mogelijkheden als de beperkingen van de applicatie – heb ik crowdsourcing als hermeneutisch experiment in letterkundig onderzoek zowel verifiërend als falsifiërend geëvalueerd.

Uit een procesevaluatie kon ik uiteindelijk concluderen dat crowdsourcing inderdaad ook effectief blijkt om neutrale of valide gegevens over literatuur te genereren. Hoewel ik in mijn analyse van het debat over diversiteit twee kampen of ‘interpretatieve gemeenschappen’ kon onderscheiden, was er geen sprake van een botsing van zulke gemeenschappen in reactie op dit project. Integendeel, critici van kranten en opiniebladen reageerden welwillend, en aan de respons en de afwezigheid van datavervuiling te zien zijn er genoeg reguliere lezers die een dergelijke literatuurbenadering volkomen legitiem vinden. Tegelijkertijd bleek uit enkele reacties op social media en blogs dat sommige lezers wel degelijk sceptisch staan tegenover literatuuronderzoek dat geïnspireerd is door ideologiekritische of zelfs activistische tradities. Mijn verwachting is dat zulke lezers simpelweg niet meedoen aan het onderzoek. Het is daarom voorbarig om te stellen dat deze data en deze methode *alle* grenzen van interpretatieve gemeenschappen overstijgen. Wel is gebleken dat de grenzen tussen academische, kritische en recreatieve literatuurbeschouwingen mogelijk en althans in dit geval niet zo strikt zijn als we soms willen geloven. Dit onderzoek sterkt daarmee Rita Felski’s hypothese: behalve de geconstrueerde en reële verschillen tussen academische en niet-academische lezers vertonen onze manieren van interpreteren *ook* belangrijke overeenkomsten. Dat resultaat legitimeert een meer fundamentele conclusie: met crowdsourcing als letterkundige onderzoeksmethode kunnen we uiteindelijk niet alleen de diversiteit van de Nederlandstalige romanliteratuur transhistorisch en transgeografisch aantonen, maar produceert de letterkunde bovendien kennis die voortkomt en recht doet aan een grotere diversiteit aan voorkeuren, leeshoudingen en poëtische opvattingen dan een solistische letterkundige ooit kan bereiken.

9. Literatuur

Primair

- Amatmoekrim, K., *Het gym*, Amsterdam: Prometheus, 2011.
Burnier, A., *Het jongensuur*, Amsterdam: Querido, 1969.
Grunberg, A., *De asielzoeker*, Amsterdam: Nijgh & Van Ditmar, 2003.

Secundair

- Amatmoekrim, K., 'Een monoculturele uitwas. De ondraaglijke witheid van de Nederlandse letteren'. *De Groene Amsterdammer*, 20-08-2015.
Barthes, R., 'The Death of the Author'. In: R. Barthes, *Image Music Text*, Londen: Fontana Press, 1977: 142-148.
Benkler, Y., 'Coase's Penguin, or, Linux and the Nature of the Firm'. *The Yale Law Journal* 112 (2002) 3: 429.
Bertens, H., 'Frontale aanval op de humanistische traditie'. *Voors* 33 (2016) 2: 52- 61.
Best, S. & S. Marcus, 'Surface Reading: An Introduction'. *Representations* 108 (2009) 1: 1-21.
Brabham, D. C., 'Crowdsourcing: A Model for Leveraging Online Communities'. In: A. A. Delwiche & J. Henderson (Eds.), *The participatory cultures handbook*. New York: Routledge, 2013: 120-129.
Brin, S. & L. Page, 'The Anatomy of a Large-Scale Hypertextual Web Search Engine'. *Computer Networks and ISDN Systems* 30 (1998): 107-117.
Broeke, A. ten, 'Literaire bloedarmoede'. *de Volkskrant*, 5-6-2015.
Boer, T. de, 'Ontmaskeraars van de goede smaak'. *De Groene Amsterdammer*, 6-5-2009.
Boog, M., 'Stukkie 28: "Artsen zijn vooral mannen, verpleegkundigen vrouwen"'. www.markboog.nl, 11 mei 2016.
Collini, S., 'Introduction : Interpretation terminable and interminable'. In : U. Eco, *Interpretation and Overinterpretation*, Cambridge: Cambridge University Press, 1992: 1-22.
Coster, S. de, 'Mannen, wees gewaarschuwd. Vrouwelijke schrijvers rukken op'. *de Morgen*, 15-01-2015.
Deijl, L.A. van der, 'De semiotiek van het algoritme. Het structuralistische denken in de eenentwintigste eeuw'. *Voors* 33 (2015) 2: 40-51.
Deijl, L.A. van der, S.A. Pieterse, M. Prinse & R.J.H. Smeets, 'Mapping the Demographic Landscape of Characters in Recent Dutch Prose: A Quantitative Approach to Literary Representation'. *Journal of Dutch Literature* 7 (2016) 1: 20-42.
Derrida, J., 'Signature Event Context'. In: J. Derrida, *Limited Inc*, vert. S. Weber & J. Mehlman, Evanston, IL: Northwestern University Press, 1988.
Dessing, M., 'Zijn vrouwen in Nederlandstalige romans werkelijk vooral prostituee?'. *Knack*, 07-05-2016.
Eco, U., *Interpretation and Overinterpretation*. Cambridge: Cambridge University Press, 1992.
Felski, R., 'After Suspicion'. *Profession* (2009): 28-34.
Felski, R., *Uses of Literature*. Oxford: Blackwell Publishing, 2008.
Fish, S., *Is there a text in this class? The authority of interpretative communities*. Cambridge MA: Harvard University Press, 1980.
Fortuin, A., 'Schrijvers, kom uit die hoofden! En geef ons een keisnijder'. *NRC Handelsblad*, 22-04-2016.
Franssen, T en D. Stoffelsen, 'De witte motor. De verblindende blankheid van het boekenvak'. *De Groene Amsterdammer*, 20-8-2015.
Hayles, K. N., *How We Became Posthuman. Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago: University of Chicago Press, 1999.
Hayles, K. N., *My Mother Was a Computer. Digital Subjects and Literary Texts*. Chicago: University of Chicago Press, 2005.
Hermse, J., 'Schrijven vrouwen wel genoeg?' Recensieweb, 13-05-2013.
Hofkens, A., 'In de Nederlandse literatuur zijn vrouwen huisvrouw of hoer'. *Opzij*, 28-04-2016.
Huff, P., 'Literatuurcriticus in Nederland is een man van middelbare leeftijd'. *de Volkskrant*, 06-11-2014.
Joosten, J., 'Scheve verhoudingen in literatuur zitten dieper'. *Trouw*, 02-05-2013.
Kuitenbrouwer, J., 'Discriminatie in literaire wereld? Show, don't tell'. *NRC Handelsblad*, 30-05-2015.

- Lezeres des Vaderlands, '#LekkerTellen', 28-03-2016.
<https://lezeresdesvaderlands.wordpress.com/2016/03/>
- Lezeres des Vaderlands, 'Door mijn leesbril'. *De Groene Amsterdammer*, 02-03-2016
- Lezeres des Vaderlands, 'Toverstaf en tijdmachine. Hoe ons literatuuronderwijs vrouwen buitensluit'. *De Groene Amsterdammer*, 20-04-2016
- Marsman, L., 'Het vingertje van Ouariachi. Hoe Jamal Ouariachi toch geen feminist werd'. *Tirade*, 05-03-2014.
- Mersbergen, J. van, 'Een mooi tapijt', 22-04-2016. www.janvanmersbergen.nl
- O'Reilly, T., 'What is Web 2.0? Design Patterns and Business Models for the Next Generation of Software', 30-09-2005, www.oreilly.com.
- Ouariachi, J., 'Vrouwen, leg de lat eens hoger!' *Vrij Nederland*, 07-03-2014.
- Ouariachi, J., 'Slachtofferschap'. *Trouw*, 21-08-2015.
- Paassen, D. van, 'Lekker ding kan schrijven'. *Opzij*, 23-04-2013.
- Peeters, C., 'Waarom jij, lezer, echt geen lege huls bent'. *Vrij Nederland*, 28-07-2014.
- Pieterse, S., 'De emancipatie van de lezer'. NY, 09-09-2014.
- Pieterse, S., *De buik van de lezer. Over spreken en schrijven in Multatuli's Ideën*. Nijmegen: Vantilt, 2008.
- Ramdas, A., 'Moedwil en kwade trouw bij blanke schrijvers. Niemand heeft oog voor het vreemde'. *NRC Handelsblad*, 14-03-1997.
- Ramsay, S., *Reading Machines. Towards an Algorithmic Criticism*. Urbana: University of Illinois Press, 2011.
- Riesch, H., & Potter, C., 'Citizen science as seen by scientists: Methodological, epistemological and ethical dimensions'. *Public Understanding of Science* 23 (2014) 1: 107-120.
- Rouw, E., 'Literatuur blijft te wit'. *NRC Handelsblad*, 16-05-2015.
- Slot, P., 'De eregalerij blijft gesloten voor boeken van vrouwen'. *Trouw*, 27-03-2013.
- Speet, F., 'Schrijfsters en literaire waardering'. *De Morgen*, 15-01-2014.
- Stoffelsen, D., 'Slachtofferschap, een positiebepaling'. *De Revisor*, 10-09-2015.
- Tapscott, D. & Williams & A.D., *Wikinomics. How Mass Collaboration Changes Everything*. London: Penguin, 2008.
- Howe, J. 'Crowdsourcing: A definition', 2006.
www.crowdsourcing.com/cs/2006/06/crowdsourcing_a.html
- Iser, W. 'The Reading Process: A Fenomenological Approach'. In: J.P Tompkins, (ed.) 1988 (1986). *Reader-Response Criticism. From Formalism to Post-Structuralism*. Baltimore/London: The Johns Hopkins University Press: 50-69.
- Ruiter, F. & Smulders, W., *Literatuur en moderniteit in Nederland 1840-1990*. Amsterdam: De Arbeiderspers, 1996.
- Said, E., 'Opponents, Audiences, Constituencies and Community'. In: W. J. T. Mitchell (ed.), *The Politics of Interpretation*, Chicago: University of Chicago Press, 1988: 7-32.
- Siegel, E. & A. Retter, *eXist. A NoSQL Document Database and Application Platform*. Sebastopol (CA): O'Reilly Publishers, 2014.
- Sontag, S., *Against Interpretation*, New York: Dell Publishing, 2009.
- Surowiecki, J., *The Wisdom of Crowds*. New York: Anchor Books, 2004.
- Vaessens, T., *De revanche van de roman. Literatuur, autoriteit en engagement*. Nijmegen: Vantilt (2009).
- Veer, J. van der, 'Grands écrivains?' *Vooy's* 31 (2013) 1: 8-18.
- Walmsley, *XQuery. Search Across a Variety of XML Data*. Sebastopol (CA): O'Reilly Publishers (2007).
- Weijers, N., 'Vrouwen, mannen, literatuur'. *NRC Next*, 13-11-2014.
- Wimsatt, W.K. & M.C. Beardsley, 'The Intentional Fallacy'. *The Sewanee Review* 54 (1946): 468-488.
- Zwagerman, J., 'Literatuur als inlooppcentrum'. *NRC Handelsblad*, 28-03-1997.

Bijlage 1. Voorbeeldlemma database

```
<database>
  <roman num="107">
    <metadata>
      <titel>De avonden</titel>
      <jaar>2007</jaar>
      <eerste_druk>1947</eerste_druk>
      <uitgever>De Bezige Bij</uitgever>
      <ISBN>9789023427063</ISBN>
      <KBentry>http://opc4.kb.nl/DB=1/PPN?PPN=305106724</KBentry>
      <auteur>
        <naam>Reve, Gerard</naam>
        <gender>Man</gender>
        <geboortejaar>1923</geboortejaar>
      </auteur>
    </metadata>
    <personages>
      <personage num="638">
        <naam>Frits van Egters</naam>
        <rol>Hoofdpersonage</rol>
        <gender>Man</gender>
        <leeftijd>26-35</leeftijd>
        <etniciteit>Nederland</etniciteit>
        <woonplaats>Amsterdam</woonplaats>
        <land>Nederland</land>
        <woonplaats_type>Stad</woonplaats_type>
        <beroep>Kantoormedewerker</beroep>
        <opleiding>Hoogopgeleid</opleiding>
        <map_data>
          <lat>52.3702157</lat>
          <lon>4.8951679</lon>
        </map_data>
      </personage>
    </personages>
  </roman>
</database>
```

Bijlage 2. Formulierinstructies

Wanneer de gebruiker op een van de tien invulvelden klikt, verschijnt er links in beeld een korte richtlijn over hoe zij of hij deze variabelen kan interpreteren. Die instructies luiden als volgt:

Naam

Vul de naam in waarmee het personage wordt aangeduid in de roman. Wees zo volledig mogelijk. Als de achternaam of de voornaam niet bekend is, laat die dan weg. Als het personage een bijnaam heeft, vul die dan in. Als de naam onbekend is, schrijf dan 'Onbekend'.

Rol

Als het vertelperspectief op enig moment in het verhaal bij dit personage ligt, kies dan voor 'Verteller'. Als het personage een belangrijke ontwikkeling doormaakt of als het een belangrijke rol speelt in het verhaal, kies dan voor 'Hoofdpersonage'. In alle andere gevallen kies je voor 'Nevenpersonage'.

Gender

Als de roman suggereert dat het personage niet te beschrijven valt in de termen 'Man' of 'Vrouw', kies dan 'Transgender'. Als het gender niet bekend is, kies dan voor 'Onbekend'.

Leeftijd

Kies de leeftijdscategorie die voor het grootste deel van het verhaal van toepassing is. Ga bij twijfel uit van de hoogste leeftijd die het personage bereikt in de roman. Als de leeftijd niet expliciet genoemd wordt, schat de leeftijd dan: een vrouw met een studerende zoon is 56-65, een student jonger dan 25, een grootmoeder ouder dan 65. Als de roman geen enkele aanwijzing geeft over de leeftijd, kies dan 'Onbekend'.

Herkomst

De 'herkomst' verwijst naar de geboortegrond van het personage, tenzij die afwijkt van de geboortegrond van zijn of haar ouders. In dat geval is de herkomst van de ouders leidend. Bij zogenaamde 'migranten van de tweede of derde generatie' kies je dus de herkomst van de eerste generatie. Bij een fictief land van herkomst, kies je 'Fictief'.

Woonplaats (land)

Kies het land waar het personage woont. Als het personage tijdens het verhaal in meerdere landen woont, ga dan uit van de belangrijkste plaats of van de plaats waar het personage het langst gewoond heeft.

Woonplaats

Vul zo nauwkeurig mogelijk de belangrijkste woonplaats van het personage in. Heeft het personage zijn hele leven in Parijs gewoond, maar woont hij tijdens de vertelling in Amsterdam, vul dan 'Amsterdam' in. Als het personage tijdens de vertelling op meerdere plekken woont, vul dan meerdere plaatsnamen in. Als de roman geen enkele aanwijzing geeft over de woonplaats, vul dan 'Onbekend' in.

Woonplaats (type)

Kies voor een type woonplaats. Als het personage tijdens het verhaal op meerdere plaatsen woont, ga dan uit van de belangrijkste plaats of van de plaats waar het personage het langst gewoond heeft.

Beroep

Het beroep kan breed worden opgevat. Als de term onbruikbaar is, ga dan uit van 'belangrijkste dagelijkse bezigheid'. Als het personage nog naar school gaat of studeert, vul dan 'Scholier' of 'Student(e)' in. Als de roman suggereert dat een personage niet werkt, vul dan 'Werkloos' in. Als het niet duidelijk is of het personage wel of niet werkt, vul dan 'Onbekend' in.

Opleidingsniveau

De term 'hoogopgeleid' verwijst naar wat in de volksmond 'WO/HBO'-denkniveau heet. 'Laagopgeleid' is van toepassing op iedereen die een beroepsopleiding afgerond heeft, een middelbare schooldiploma

behaald heeft, of ongeschoold is. Als het opleidingsniveau van het personage niet bekend is, ga dan uit van het opleidingsniveau dat zijn of haar beroep vereist. Als zowel het opleidingsniveau als het beroep niet bekend zijn, kies dan 'Onbekend'.

Bijlage 3 Media-overzicht Personagebank

Artikelen over de Personagebank:

- Arjen Fortuin, '[Schrijvers, kom uit die hoofden! En geef ons een keisnijder!](#)', *NRC Handelsblad*, 22 april 2016.
- Anoeek Hofkens, '[In de Nederlandse literatuur zijn vrouwen huisvrouw of hoer](#)', *Opzij*, 28 april 2016.
- Maarten Dassing, '[Zijn vrouwen in Nederlandstalige romans werkelijk vooral prostituee?](#)', *Knack* 7 mei 2016.

Blogs en website-artikelen over de Personagebank:

- Neder-L, '[Personagebank online](#)', 18 april 2016.
- Lezeres des Vaderlands, '[Huwelijks crisis in de boekenbijlagen: #lekkertellen week 13](#)', 18 april 2016.
- Internationale Vereniging voor Neerlandistiek, '[Neerlandici publiceren crowdsourced database van romanpersonages](#)', 28 april 2016.
- Schrijven Online, '[Masterstudenten tellen personages in Nederlandstalige romans](#)', 1 mei 2016.
- utrecht.nieuws.nl, '[Personages in Nederlandstalige romans in kaart brengen](#)', 1 mei 2016.
- ViFa Benelux, '[Neu im Web: Personagebank](#)', 6 mei 2016.
- Mark Boog, '[Stukkie 28: "Artsen zijn vooral mannen, verpleegkundigen vrouwen"](#)', www.markboog.nl, 11 mei 2016.
- Iedereen Wetenschapper, '[Hoe divers is de Nederlandstalige literatuur? Een literaire volkstelling](#)', 12 mei 2016.

Interviews met initiatiefnemer Lucas van der Deijl:

- Mark Beekhuis voor de rubriek '[Wetenschap Vandaag](#)' van BNR Nieuwsradio, 26 april 2016, 05:10 pm.
- Conny Kraaijeveld voor het programma 'Aan Tafel' van RTV Utrecht, 3 mei, 13:35-13:45.

Promotie vanuit de universiteit:

- De Personagebank was een van de uitgelichte onderzoeken op de homepage van de universiteit van 22 tot 25 april 2016.
- Blog Nederlandse taal en cultuur UU, '[Doe mee met het vullen van de Personagebank](#)', 15 april 2016.
- uu.nl, '[Neerlandici publiceren crowdsourced database van romanpersonages](#)', Nieuws 21 april 2016.
- uu.nl, '[Veel media-aandacht voor Personagebank](#)', Nieuws 29 april 2016.

Bijlage 4.
 Informatieschema
 Serverresponse na invoer van
 ISBN-nummer

