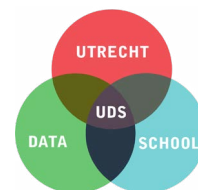




Universiteit Utrecht



‘Wie zoekt, zal vinden’: of toch niet?

Een verkenning van zoekalgoritmen
op vacaturewebsites en het effect
op gelijke kansen

Auteurs

Daniël Everts

Iris Muis

Karin van Es

Marieke van Santen

Met dank aan:

Arthur Vankan

Daphne Cornelisse

In opdracht van het College voor de Rechten van de Mens

Maart 2021

Inhoudsopgave

1. Samenvatting	3
2. Aanleiding	5
3. Discriminatie en algoritmische vooringenomenheid	7
3.1 Discriminatie, vooringenomenheid en algoritmen	7
3.1.1 Vooringenomenheid ingeprogrammeerd in algoritmen	8
3.1.2 Vooringenomenheid uit de samenleving, gereflecteerd door algoritmen	9
3.2 Vooringenomenheid op de arbeidsmarkt	10
3.2.1 Het zoekproces van een werkzoekende op vacaturesites	10
3.2.2 Taalgebruik en gender	12
3.2.3 Het neutraliseren van gendergekleurde taal	13
3.2.4 De werking van zoekalgoritmen	13
3.3 Onderzoek in de context van het gehele sollicitatieproces	14
4. Onderzoeksaanpak	16
4.1 Fase I	16
4.2 Fase II	19
5. Resultaten	21
5.1 Fase I	21
5.1.1 Vacaturesite A	22
5.1.2 Vacaturesite B	22
5.1.3 Vacaturesites C en D	22
5.2 Fase II	22
5.2.1 Mate van neutraliseren: concrete voorbeelden	23
5.2.2 Mate van neutraliseren: totaalbeeld	25
5.2.3 Kwantitatieve check	26
5.2.4 Vooringenomenheid tegengaan, reproduceren of versterken	28
6. Conclusie	30
6.1 Discriminatie op basis van profiel	30
6.2 Genderstereotypen en het zoekalgoritme	30
6.2.1 Het zoekalgoritme gaat genderstereotypen tegen	31
6.2.2 Het zoekalgoritme reproduceert genderstereotypen	31
6.2.3 Het zoekalgoritme versterkt genderstereotypen	32
6.3 Overkoepelende conclusie en aanbevelingen	32
7. Referenties	34
8. Appendix	37

1 Samenvatting

Lang werden algoritmen gezien als neutraal en objectief, een aanname waarvan inmiddels het tegendeel is bewezen. Ze kunnen namelijk beïnvloed zijn door bestaande maatschappelijke stereotypen en vooroordelen. Bijvoorbeeld dat bepaalde beroepen geschikter zouden zijn voor vrouwen dan mannen (of juist andersom), met als risico dat dit van invloed is bij het algoritmisch matchen van banen met mannelijke versus vrouwelijke kandidaten. Een algoritme kan in potentie deze maatschappelijke vooringenomenheid tegengaan, maar ook reproduceren of zelfs versterken. Dit exploratieve onderzoek richt zich op deze effecten van zoekalgoritmen op vacaturewebsites, om te achterhalen of er sprake is van algoritmische vooringenomenheid dat zou kunnen leiden tot mogelijke discriminatie op de arbeidsmarkt.

Werving en selectie vindt tegenwoordig in toenemende mate plaats op de online arbeidsmarkt. Steeds meer werkzoekenden maken gebruik van vacaturewebsites om aan een baan te komen. Het bijeenbrengen van vraag en aanbod op een dergelijke website wordt bewerkstelligd door een zoekalgoritme. Het zoekalgoritme heeft grote invloed op de kansen van de werkzoekende. Ten eerste bepaalt het welke vacatures de werkzoekende te zien krijgt en dus op welke vacatures de werkzoekende in staat is om te reageren. Deze kan immers niet reageren op vacatures die hij of zij niet te zien krijgt. Ten tweede bepaalt het zoekalgoritme het niveau van genderstereotypen dat terugkomt in de getoonde vacatures. Het kan genderneutrale functietitels (zoals 'docent') voorrang geven boven gendergekleurde functietitels (zoals 'lerares'), geen onderscheid maken tussen gendergekleurd en genderneutraal of zelfs voorrang geven aan gendergekleurde functietitels boven genderneutrale. Genderstereotypen in taalgebruik op de arbeidsmarkt kunnen grote invloed hebben op werkzoekenden, in hoe zij beoordeeld worden en op welke vacatures zij zich geroepen voelen om te reageren.

Voor deze verkenning was er geen toegang tot de technische documentatie met betrekking tot de werking van de algoritmen. Daarom is besloten om de onderzoeksaanpak te baseren op *reverse engineering*: het gecontroleerd ingeven van input en het bestuderen van de output, om vervolgens af te leiden wat het algoritme doet. Dit is op een exploratieve manier gedaan: met een beperkte dataset en op handmatige wijze.

Het eerste deel van het onderzoek richt zich op de vraag of werkzoekenden met verschillende profielkenmerken, verschillende vacatures getoond krijgen door het zoekalgoritme. Daarbij zijn in december 2020 vier grote Nederlandse vacaturesites onderzocht door middel van twaalf aangemaakte profielen per website en zoekacties uitgevoerd door de betreffende profielen. Werkzoekenden kunnen op vacaturewebsites een profiel aanmaken en daarbij informatie opgeven over onder andere werkervaring en opleidingsniveau. De zoekalgoritmen zouden deze profielinformatie mee kunnen laten wegen in hun oordeel over wat relevante zoekresultaten zijn voor die specifieke werkzoekende,

en daarbij mogelijk onderscheid kunnen maken in kenmerken die als verboden discriminatiegrond zijn aangemerkt. De profielkenmerken waren gerelateerd aan verschillende discriminatiegronden: geslacht, etniciteit (ras) en leeftijd. De zoekresultaten van de verschillende profielen werden vergeleken: hierbij werd gekeken of eventuele verschillen in resultaten herleid konden worden naar de verschillende profielkenmerken, om eventueel verboden onderscheid te kunnen detecteren. Er zijn hierbij geen indicaties gevonden die wijzen op onderscheid door de vacaturewebsites dat duidelijk toe te schrijven is aan verschil in profielkenmerken.

Het tweede deel van het onderzoek richt zich op de invloed van de werking van het zoekalgoritme op gendergekleurd taalgebruik en genderstereotypen. De algoritmen kunnen vooringenomenheid in de vorm van genderstereotypen tegengaan, reproduceren of versterken, afhankelijk van hun werking. Om dit te onderzoeken is een lijst samengesteld van vijftien gendergekleurde functietitels en hun neutrale tegenhanger(s). Deze functietitels werden als zoekterm ingevoerd op de verschillende vacaturesites. De output werd vervolgens geregistreerd en geanalyseerd. De bevindingen wijzen op het bestaan van alledrie de fenomenen. Ten eerste corrigeren zoekalgoritmen op vacaturesites in sommige gevallen gendergekleurde zoektermen naar genderneutrale zoekresultaten, zoals wanneer met 'lerares' wordt gezocht en daarbij ook zoekresultaten worden getoond voor 'docent'. Hiermee wordt vooringenomenheid tegen gegaan. Ten tweede nemen in sommige gevallen de zoekalgoritmen gendergekleurde zoektermen één-op-één over in de zoekresultaten, zoals wanneer met 'lerares' wordt gezocht en alleen resultaten voor 'lerares' getoond worden en niet ook die voor 'docent'. Hiermee wordt vooringenomenheid gereproduceerd. Ten derde hebben we, in mindere mate, gezien dat een zoekalgoritme vooringenomenheid versterkt door het tonen van lager betaalde functies bij gebruik van een 'typisch vrouwelijke' functietitel als zoekterm, zoals wanneer met 'lerares' wordt gezocht en vervolgens een functie voor 'onderwijsassistent' getoond wordt. De zoektermen 'secretaresse' (ondanks dat dit tegenwoordig niet uitsluitend vrouwen aanduidt), 'schoonmaakster' en 'verpleegster' worden op alle vacaturesites weinig geneutraliseerd. We zijn geen voorbeelden tegen gekomen van gendergekleurde zoekresultaten bij gebruik van een genderneutrale zoekterm (wat vooringenomenheid zou versterken).

Dit kan gerelateerd worden aan het type algoritme dat ingezet wordt: een simpel zoekalgoritme dat slechts letterlijk de ingevoerde zoekterm overneemt, is in ieder geval niet in staat om vooringenomenheid tegen te gaan maar zal dit reproduceren. Een zoekalgoritme dat complexere methoden gebruikt, kan zowel in staat zijn om genderstereotypen tegen te gaan, te reproduceren of te versterken. Welke uitkomst het heeft, hangt af van hoe het algoritme precies is ontwikkeld. De consequentie hiervan kan echter zijn dat door vooringenomenheid van (zowel de werkzoekende als) het zoekalgoritme van een vacaturesite, een vrouw niet altijd dezelfde vacatures te zien krijgt als een man (en vice versa). Het verdient daarom aanbeveling dat vacaturesites hun algoritmen doorlichten om dergelijke effecten zoveel als mogelijk tegen te gaan, bijvoorbeeld door complexere boven simpele zoekalgoritmen te verkiezen.

2 Aanleiding

Het College voor de Rechten van de Mens (het College) heeft signalen ontvangen over mogelijke discriminatie op vacaturewebsites vanuit de algoritmen op basis waarvan deze websites draaien.¹ Algoritmen zijn, simpel geformuleerd, stap-voor-stap instructies waarmee computers acties kunnen uitvoeren. Algoritmen kunnen vooroordelen bevatten, bijvoorbeeld doordat de programmeur het algoritme heeft gemaakt met bepaalde aannames in het achterhoofd.² Het College voor de Rechten van de Mens heeft de Utrecht Data School (UDS), Universiteit Utrecht, verzocht om effecten van dergelijke impliciete aannames in de algoritmen waarop vacaturewebsites draaien te onderzoeken. Middels deze verkenning wenst het College de mogelijke risico's met betrekking tot vooringenomenheid en discriminatie van werving en selectie op vacaturewebsites met behulp van algoritmen nader te onderzoeken en indien nodig te agenderen.

Met dit onderzoek bouwt UDS voort op eerder door het College uitgevoerd literatuuronderzoek, waarin uitgebreid naar voren komt hoe algoritmen op vacaturewebsites een discriminerend effect kunnen hebben (College 2020). Hoewel er al enig onderzoek naar discriminatie op vacaturewebsites is verricht³, is er vooralsnog niets bekend over de manier waarop algoritmen een rol spelen in het proces dat werkzoekenden doorlopen wanneer zij zoeken naar vacatures. In dit rapport richten wij ons daarom op de mogelijk uitsluitende werking van zoekfuncties van vacaturesites. Hiermee vullen wij in dit rapport de theoretische kennis over algoritmische discriminatie die in het eerdergenoemde literatuuronderzoek is vergaard verder aan.

Algoritmen hebben de potentie om bestaande vooringenomenheid in de samenleving te verminderen (neutraliseren), één-op-één te reproduceren (naturaliseren) of zelfs te versterken. Dat geldt ook voor de zoekalgoritmen op vacaturesites. In dat kader zullen de volgende onderzoeksvragen worden behandeld: *Zijn er signalen van een discriminerende werking van algoritmen die worden ingezet bij matching van profielinformatie van werkzoekenden en banenaanbod op Nederlandse vacaturewebsites?*

- a. *Lijkt er sprake te zijn van algoritmen op de vier te onderzoeken vacaturewebsites, die onderscheid maken op basis van profielinformatie van werkzoekenden? (Fase I) Zo ja, hoe gaan deze te werk?*
- a. *Hoe gaan de zoekalgoritmen op vacaturewebsites om met gegenderde taal, bijvoorbeeld wanneer gezocht wordt op 'secretaresse'? Neutraliseren zij deze taal actief, of doen zij dat niet en versterken zij daarmee bepaalde gendernormen die leven in de maatschappij? (Fase II)*

¹ Zie onder andere <https://medium.com/@anukritisaurabh/employment-discrimination-blame-it-on-job-recruiting-sites-64edad4eaa8b>, D. Hangartner, D. Kopp, M Siegenthaler 2021, Monitoring hiring discrimination through online recruitment platforms, Nature (589)

² Zie ook de definitie van algoritme uit 'Als computers je CV beoordelen, wie beoordeelt dan de computers?', College voor de Rechten van de Mens, 2 september 2020, pagina 6: "In simpele woorden zijn algoritmen een soort instructieregels waarmee de computer op basis van ingevoerde informatie – zoals de selectiecriteria bij een vacature – besluiten kan nemen."

³ Zie bijvoorbeeld het eerdere onderzoek in opdracht van de Inspectie-SZW: Y. Westerbaan, W. Kleefstra, T. Eefting 2021, Verkennend onderzoek naar discriminatiemogelijkheden op online vacatureplatforms.

Voordat we ingaan op deze vragen en de onderzoeksanpak, zullen we in het volgende hoofdstuk ingaan op de manier waarop algoritmen en specifiek de zoekfunctie bij vacaturesites een discriminerende werking kunnen hebben. Er wordt toegelicht hoe stereotypen en vooroordelen ingebed worden in algoritmen, als ook hoe vooringenomenheid en ongelijkheid door algoritmen kunnen ontstaan, dan wel kunnen worden versterkt.

3 Discriminatie en algoritmische vooringenomenheid

Voringenomenheid komt op alle gebieden van onze samenleving voor en is een menselijke eigenschap die ons helpt om een beeld te vormen over iets of iemand, wanneer er beperkte informatie beschikbaar is. Deze vooringenomenheid komt voort uit stereotiepe maatschappelijke denkbeelden en daarop gebaseerde vooroordelen, en handelen op basis van deze vooringenomenheid kan leiden tot discriminatie. Dit komt ook voor op de arbeidsmarkt (College 2013). Binnen dit hoofdstuk relateren we deze begrippen aan algoritmen om te illustreren dat vooringenomenheid ook binnen een algoritme tot uiting kan komen en dat onderscheid op basis van een vooringenomen algoritme kan leiden tot discriminatie. Ten slotte plaatsen we dit fenomeen in de context van het proces van een werkzoekende die via een vacaturesite op zoek gaat naar werk.

3.1 Discriminatie, vooringenomenheid en algoritmen

Discriminatie kan zowel direct als indirect plaatsvinden. We spreken van directe discriminatie wanneer een persoon op wat voor manier dan ook anders dan anderen wordt behandeld, puur op basis van een verboden grond zoals geslacht, ras en/of leeftijd (zie Algemene Wet Gelijke Behandeling en andere Nederlandse gelijkebehandelingswetgeving, d'Alessandro, O'Neil, en LaGatta 2017, 124).⁴ Dergelijke eigenschappen worden ook wel 'beschermd criteria' genoemd; groepen die worden gedefinieerd op basis van deze criteria worden ook wel 'beschermd groepen' of 'beschermd klassen' genoemd (zie Nederlandse gelijkebehandelingswetgeving, Seaver 2017, 5).

Indirecte discriminatie vindt plaats wanneer de behandeling van een persoon wordt gebaseerd op eigenschappen die weliswaar niet bij wet beschermd zijn, maar wel dusdanig met beschermd criteria correleren dat personen uit een beschermd klasse significant anders worden behandeld (d'Alessandro, O'Neil, en LaGatta 2017, 124). Een duidelijk voorbeeld hiervan is de manier waarop de zwarte Amerikaanse bevolking vaker onderhevig is aan politiecontrole, omdat mensen die tot deze groep behoren doorgaans in armere wijken, waar meer politiecontroles worden uitgevoerd, woonachtig zijn (Alexander 2020, 7, zie ook Kader 1). Een ander voorbeeld is wanneer in een vacature als eis wordt gesteld dat de Nederlandse taal volledig beheerst moet worden, ook al is dat voor de functie an sich niet relevant. Dit benadeelt personen met een andere nationaliteit dan Nederlands, die de taal niet (vloeiend) beheersen, en is als zodanig indirecte discriminatie.

In het begin van de 21^{ste} eeuw is de wereld in rap tempo gedigitaliseerd en tegenwoordig wordt er in

⁴ De tien wettelijk verboden discriminatiegronden bij (toegang tot) arbeid zijn godsdienst, levensovertuiging, politieke gezindheid, ras, geslacht, nationaliteit, hetero- of homoseksuele gerichtheid, burgerlijke staat, handicap of chronische ziekte en leeftijd.

vrijwel alle delen van de samenleving gebruik gemaakt van algoritmen. Lang werden algoritmen gezien als remedie tegen menselijke vooroordelen onder de aanname dat computers enkel doen waarvoor ze geprogrammeerd zijn en zodoende niet bevooroordeeld kunnen zijn op de manier waarop mensen dat wel kunnen zijn – in het Engels noemt men dat ook wel *automation bias* (Zuiderveen Borgesius 2018, 10). Deze presumptie is de afgelopen jaren echter ongegrond gebleken: ondertussen is bekend dat algoritmen wel degelijk bevooroordeeld kunnen zijn.

3.1.1 Vooringenomenheid ingeprogrammeerd in algoritmen

Dergelijke vooringenomenheid kan op drie manieren naar voren komen. Zo kunnen technische limieten binnen computertechnologie leiden tot vooringenomenheid ten aanzien van bepaalde beschermde groepen, bijvoorbeeld wanneer een systeem simpelweg de complexiteit van de werkelijkheid niet aan kan en daardoor wel bepaalde aannames *moet* maken – dergelijke vooringenomenheid wordt ook wel *technical bias* genoemd. Ook komt het voor dat reeds bestaande vooroordelen in individuen en samenlevingen (onbewust) worden 'doorvertaald' naar algoritmische functies (zie Kader II) – in dat geval spreken we over *preexisting bias* (Friedman en Nissenbaum 1996, 334). Als laatste is er *emergent bias*: vooringenomenheid die meestal pas na verloop van tijd naar voren komt wanneer een algoritme in een andere context wordt gebruikt dan de context die de ontwikkelaars voor ogen hadden – het desbetreffende algoritme kan dan niet accuraat inspelen op de nieuwe situatie (Ibid., 335). Denk hierbij bijvoorbeeld aan gezichtsherkenning algoritmen die mensen met een donkere huidskleur niet kan herkennen, omdat de (witte) programmeurs enkel van hun eigen huidskleur zijn uitgegaan.

De verschillende vormen van vooringenomenheid in algoritmen staan meestal niet volledig los van

KADER I: Algoritmische discriminatie in het Amerikaanse rechtssysteem

In 2016 lagen algoritmen onder vuur die medezeggenschap hadden in het Amerikaanse strafrechtelijk proces. Deze algoritmen voorspelden namelijk de waarschijnlijkheid van recidive bij criminelen op basis van eerdere aanrakingen met politie, maar – zo bleek uiteindelijk uit kritisch onderzoek – zonder daar in mee te nemen of criminelen bij deze eerdere aanrakingen daadwerkelijk schuldig werden bevonden aan een misdaad. De aanname van de programmeurs was dat elke aanraking met politie een indicatie was voor mogelijk recidivisme. Deze aanname had echter een indirect discriminerende werking: doordat de zwarte Amerikaanse bevolking disproportioneel in aanraking komt met politie, zorgden deze algoritmen er onbedoeld voor dat zwarte mensen gemiddeld zwaardere straffen kregen opgelegd (O'Neil 2016, 31–32). Zodoende werd een indirecte vorm discriminatie door toedoen van deze algoritmen gesystematiseerd en hervormd tot een veel directere vorm van discriminatie.

elkaar, maar liggen in elkaars verlengde en kunnen in elkaar overlopen. De werkelijkheid classificeren met algoritmen vraagt om een simplificatie van die werkelijkheid, onder andere door het doen van aannames. Aannames van programmeurs over hoe de werkelijkheid in elkaar zit – oftewel reeds bestaande vooroordelen – worden daarom veelal overgenomen in algoritmische functies, die vervolgens een bevooroordeelde en dikwijls discriminerende werking blijken te hebben zodra ze in gebruik worden genomen (Blok, Lammerant, en De Hert 2018; O'Neil 2016, 24; Bowker en Star 1999, 46–47).

Zo kunnen aannames van programmeurs over welke *variabelen* wel en niet relevant zijn voor het maken van bijvoorbeeld voorspellingen over groepen mensen (onbedoeld) tot discriminerende algoritmen leiden (zie voor een sterk discriminerend algoritme Kader I). Dergelijke aannames kunnen zijn geïnformeerd door het persoonlijke, op eigen beleving gebaseerde wereldbeeld van programmeurs. Een veelgenoemd kritiekpunt over de IT-wereld is dan ook dat vooral witte, mannelijke programmeurs vaak verantwoordelijk zijn voor het produceren van algoritmen (Noble 2018b, 161).

3.1.2 Vooringenomenheid uit de samenleving, gereflecteerd door algoritmen

Bevooroordeelde algoritmen zijn echter lang niet altijd het gevolg van vooroordelen van individuele programmeurs. De classificatiesystemen en instructies waar algoritmen uit bestaan komen in de kern tot stand doordat computer- en datawetenschappers patronen ontdekken in relevante ruwe datasets en op basis daarvan een naar computercode vertaald beeld van de werkelijkheid schetsen (O'Neil 2016, 23). Er wordt beargumenteerd dat dergelijke datasets *nooit* neutraal zijn; ze zijn altijd het resultaat van ongelijke sociale, historische en economische omstandigheden in de gehele maatschappij en daarmee altijd bevooroordeeld (D'Ignazio en Klein 2020, 17).

KADER II: Zijn mannen technischer, of hebben ze vaker een technische baan?

Een algoritme voor het voorspellen van technische vaardigheden bij sollicitanten voor technische functies kan (deels) worden gebaseerd op kenmerken van eerdere technische medewerkers. Als deze werknemers allemaal mannen zijn, kan het hebben van het mannelijk geslacht in het algoritme worden geassocieerd met het bezitten van technische vaardigheden. Daardoor kan het algoritme dan onterecht mannen eerder naar voren schuiven dan vrouwen (Grommé et al. 2019, 60–62). Zodoende wordt ongelijkheid in de samenleving (oftewel pre-existing bias) niet alleen gereproduceerd, maar ook versterkt: het gevolg is namelijk dat vrouwen nog minder vaak technische functies krijgen aangeboden, waardoor datasets die in de toekomst mogelijk worden gebruikt om technische vaardigheid te voorspellen nog meer vooringenomen worden.

Naast dat algoritmen op deze manier ongelijkheid in de samenleving kunnen reproduceren, kunnen ze vooroordelen ook versterken (zie Kader II). Als een algoritme een bepaalde groep mensen steeds achterstelt door ze bijvoorbeeld minder snel in aanmerking te laten komen voor een baan, dan zal dat worden gereflecteerd in een volgende set data. Als die set data vervolgens wordt gebruikt voor een volgende versie van het algoritme, zal dat algoritme alleen nog maar sneller concluderen dat die groep mensen ongeschikt is voor werk (Grommé et al. 2019, 60–62). Hoe groter de schaal waarop algoritmen worden gebruikt, hoe groter het gevaar dat ze dergelijke versterkende effecten hebben (O'Neil 2016, 32–33). Dit risico van het gebruik van algoritmen op grote schaal geldt zeker voor vacaturesites: één van de onderzochte vacaturewebsites bevat meer dan 1 miljoen unieke profielen van werkzoekenden en heeft 1,6 miljoen bezoeken per maand.

3.2 Vooringenomenheid op de arbeidsmarkt

Uit onderzoek - dat zich deels tegelijkertijd met het huidige onderzoek heeft voltrokken - is recentelijk gebleken dat het voor werkgevers op meerdere Nederlandse vacatureplatformen mogelijk is om in vacatureteksten actief discriminerende voorkeuren aan te geven, ondanks dat het plaatsen van discriminerende teksten expliciet in de voorwaarden van deze platformen wordt verboden. Ook is het voor werkgevers mogelijk gebleken om in door vacatureplatformen aangeleverde curricula vitae van sollicitanten te filteren op basis van beschermde criteria (Inspectie SZW 2021).

Op het eerder genoemde onderzoek na, is er geen onderzoek naar Nederlandse vacaturewebsites gedaan. Het meeste onderzoek naar digitale arbeidsmarkten is gericht op Amerikaanse of internationaal opererende bedrijven en de Amerikaanse of internationale arbeidsmarkt. Daarbij heeft de focus tot dusver voornamelijk gelegen op discriminatie op basis van geslacht en de manier waarop algoritmen op basis van teksten in curricula vitae voorspellen hoe productief sollicitanten zullen zijn, nog voordat de werkgever sollicitaties onder ogen krijgt (zie Kader III); op basis van dergelijke voorspellingen bepaalt een algoritme dan of de sollicitatie daadwerkelijk bij de werkgever terecht komt (Buolamwini en Gebru 2018, 2). Dankzij onderzoek naar algoritmische discriminatie op de arbeidsmarkt worden computerwetenschappers in staat gesteld om eventuele misstanden te verhelpen; zo heeft Amazon (zie Kader III) ondertussen zijn algoritme dat vrouwen discrimineerde veranderd (Dastin 2018).

3.2.1 Het zoekproces van een werkzoekende op vacaturesites

Discriminatie kan in de digitaal gemedieerde omgeving meer beslaan dan enkel algoritmisch bepaalde classificaties en daarop volgende oneerlijke oordelen en selecties van werkzoekenden. Zo wordt beargumenteerd dat zoekalgoritmen bij zoekmachines evenzeer vooringenomenheid in de vorm van emergent bias kunnen vertonen: ze brengen veelal de meest 'relevante' zoekresultaten naar voren, maar wat doorgaat als relevant wordt bepaald door de programmeurs, wiens aannames over wat zogenaamd relevant is dus worden opgenomen in het algoritme. Doordat de meeste algoritmen nog altijd worden geschreven door heteroseksuele witte mannen die, net als alle andere mensen, bepaalde vooroordelen hebben, kan het dus gebeuren dat een algoritme heel goed is in het weergeven van zoekresultaten die

KADER III: De screening van curricula vitae

Uit recentelijk onderzoek door het bedrijf Amazon in de Verenigde Staten is naar voren gekomen dat een door het bedrijf zelf ontwikkelde algoritme curricula vitae van vrouwen onderop de stapel legde, omdat het op basis van demografische gegevens van huidige werknemers van dat bedrijf had geleerd dat mannen de voorkeur hadden voor technische functies en uit teksten in curricula vitae kon opmaken wanneer deze van vrouwen of van mannen waren (Dastin 2018).¹ Na die ontdekking zette Amazon de ontwikkeling van dit type algoritmen stop. Ook is gebleken dat vooringenomenheid van werkgevers zelf in de praktijk vooroordelen in zelflerende algoritmen teweeg brachten of versterkten: werkgevers klikten op internationaal opererende vacaturewebsites Indeed, Monster en CareerBuilder vaker op curricula vitae van mannen, waardoor de algoritmen van die websites op den duur leerden dat curricula vitae van mannen relevanter waren dan die van vrouwen. Vervolgens werden daardoor de curricula vitae van mannen vaker naar voren geschoven (Chen et al. 2018).

1 Voorbeeld, afkomstig uit Dastin 2018: "Amazon's system taught itself that male candidates were preferable. It penalized resumes that included the word "women's," as in "women's chess club captain." And it downgraded graduates of two all-women's colleges, according to people familiar with the matter."

relevant zullen zijn voor heteroseksuele witte mannen, maar tegelijkertijd niet kan voldoen aan de wens om relevante resultaten voor vrouwen, niet-witte mensen en niet-heteroseksuele mensen weer te geven (Noble 2018a, 20).⁵

Tot dusver is er in relatie tot *zoekfuncties* op vacaturewebsites weinig onderzoek verricht. Het weinige onderzoek dat bestaat richt zich op het sollicitatieproces vanuit het perspectief van de werkgever en is vooralsnog beperkt tot hoe algoritmen teksten in door sollicitanten aangedragen curricula vitae verwerken. Zo is aangetoond dat websites als LinkedIn, waarop werkgevers actief kunnen zoeken naar potentiële werknemers, vrouwennamen automatisch corrigeerden naar mannennamen (Day 2016). Al het overige onderzoek richt zich, zoals eerder genoemd, vooral op het algoritmische selectieproces van curricula vitae. Echter, voordat curricula vitae überhaupt door een algoritme worden verwerkt, moeten werkzoekenden eerst nog de keuze maken om überhaupt voor een functie te solliciteren – een keuze die in essentie bepaalt wie uit de populatie werkzoekenden onderdeel wordt van de kandidatenpool voor een functie (Peng et al. 2019).

5 Hoewel Google ondertussen zijn zoekalgoritmen heeft veranderd, kwam het in het verleden voor dat het met name door witte, heteroseksuele mannen ontwikkelde zoekalgoritme van Google voornamelijk pornografisch materiaal aanleverde wanneer de zoekterm 'black girls' ('zwarte meisjes') werd ingevoerd, terwijl dit bij de zoekterm 'white girls' ('witte meisjes') veel minder was en bij zoektermen die verwezen naar mannen nagenoeg niet (Noble 2018a, 20).

Vanzelfsprekend krijgt een werkzoekende pas de keuze om te reageren op een vacature, wanneer deze die vacature daadwerkelijk te zien krijgt. Hoewel bekend is dat algoritmen grote invloed kunnen uitoefenen op de zichtbaarheid van informatie voor specifieke doelgroepen (Tobin en Merrill 2018), is er met betrekking tot zoekmachines op vacaturewebsites nagenoeg niets bekend. Om inzicht te krijgen in mogelijke vormen van algoritmische discriminatie binnen het proces van zoeken naar vacatures, richt fase I van dit onderzoek zich op het verschil in zoekresultaten tussen verschillende beschermde groepen mensen.

3.2.2 Taalgebruik en gender

Algoritmen zijn dusdanig ingebed in en verweven met andere aspecten van onze cultuur en komen door zoveel verschillende vormen van menselijk handelen tot stand, dat ze eigenlijk niet als geïsoleerde stukken computercode kunnen worden gezien (Seaver 2017, 5). Zo kunnen ze bijvoorbeeld een grote rol spelen in het normaliseren van genderstereotypen in taalgebruik.

In tegenstelling tot bijvoorbeeld het Engels, waarin bijna alle zelfstandige naamwoorden onzijdig zijn, is de Nederlandse taal *gender-geïnfecteerd*: de meeste zelfstandige naamwoorden zijn mannelijk of vrouwelijk. In lijn met de centrale rol die de (witte, heteroseksuele) man traditioneel inneemt in de Westerse samenleving (Dyer 2002, 126), gelden mannelijke voornaamwoorden zoals 'leraar', 'kok' en 'agent' doorgaans als de zogenaamde 'norm' – deze termen kunnen zelfs 'neutraal' worden gebruikt om naar eenieder met een dergelijke functie te verwijzen. Daarentegen kunnen vrouwelijke voornaamwoorden zoals 'lerares', 'kokkin' en 'agente' enkel worden gebruikt om specifiek naar vrouwen te verwijzen, waardoor mannen duidelijk als zogenaamde 'norm' worden geduid (Perez 2019, 19–20). Er is een hedendaagse trend te herkennen in Nederlandse nieuwsmedia waarin vrouwelijke voornaamwoorden geheel worden vermeden, maar dat initiatief wordt hevig bekritiseerd. Volgens critici wordt het feit dat vrouwen, net als mannen, bijvoorbeeld leerkracht kunnen zijn verhuld wanneer het woord 'lerares' wordt afgeschaft, waardoor de norm-positie van de man alleen nog maar sterker in de samenleving wordt gecementeerd (Pous 2020).

In de praktijk blijkt de norm-positie van de man en het mannelijk voornaamwoord aanzienlijk nadelige gevolgen te hebben voor mensen die niet aan de traditionele omschrijving van de (witte, heteroseksuele) man voldoen. Zo is in relatie tot de arbeidsmarkt vastgesteld dat sterk *gegenderde* taal invloed heeft op de beoordeling van sollicitanten door hun potentiële toekomstige werkgevers: alle andere variabelen gelijkblijvend, worden vrouwen die zichzelf middels een van origine mannelijke beroepsnaam zoals 'leraar' beschrijven doorgaans positiever beoordeeld dan vrouwen die zichzelf 'lerares' noemen (Formanowicz et al. 2013). Tegelijkertijd worden vrouwen minder goed beoordeeld wanneer ze solliciteren voor een functie die met een van origine mannelijke beroepsnaam wordt omschreven ten opzichte van wanneer ze solliciteren voor een functie die met een vrouwelijke beroepsnaam wordt omschreven (Perez 2019, 18).

Naast op de beoordeling van sollicitanten, heeft taalgebruik ook invloed op hoe werkzoekenden

vacatures waarnemen. Zo blijkt uit onderzoek dat vrouwen relatief vaker reageren op vacatures waar zowel een mannelijke als vrouwelijke beroepsnaam wordt genoemd – bijvoorbeeld 'leraar/lerares' – dan op vacatures waar de mannelijke beroepsnaam als neutrale term wordt gebruikt – bijvoorbeeld 'leraar (m/v)' (Pous 2020) en dat zowel mannen als vrouwen eerder het gevoel hebben dat ze in aanmerking kunnen komen voor een baan wanneer er zowel mannelijke als vrouwelijke beroepsnamen worden genoemd (Chatard-Pannetier, Guimont, en Martinot 2005; Vervecken 2013). Dat is onder andere te verklaren vanuit het feit dat het gebruik van zowel mannelijke als vrouwelijke beroepsnamen het idee versterkt dat vrouwen succesvol kunnen zijn in traditioneel mannelijke beroepspraktijken (Stahlberg en Sczesny 2001; Vervecken 2013).

3.2.3 Het neutraliseren van gendergekleurde taal

Het maakt dus zowel uit hoe sollicitanten zichzelf kiezen te beschrijven, als hoe vacatures worden omschreven. Deze keuzes staan echter niet geheel los van invloeden van buitenaf: het handelen van mensen wordt beïnvloed door de manier waarover er al binnen een bepaalde cultuur over mensen en meer abstracte zaken wordt gesproken (Foucault 2007 [1976], 161; Hall 1992, 291). Ter concretisering: wanneer het 'als van nature' is om bepaalde functies te koppelen aan vrouwen, dan zullen vacatures voor die functie ook eerder worden omschreven met vrouwelijke beroepsnamen. Tegelijkertijd zullen vrouwen ook eerder geneigd zijn om zichzelf in die functies te zien, naar die functies te zoeken en zichzelf, wanneer ze solliciteren voor die functies, ook als vrouw – dat wil zeggen middels vrouwelijk *gegenderde* taal – te beschrijven.

Dergelijke *naturalisering* van taal is het gevolg van een constante herhaling van taalgebruik, zonder dat er aanzienlijke tegenspraak op geboden wordt (Sturken en Cartwright 2009, 16). Het is met betrekking tot dergelijke naturalisering dat algoritmen in zoekmachines een discriminerende rol kunnen vervullen die verder reikt dan enkel het genereren van zoekresultaten die niet voor iedereen even relevant zijn. Zo bleek eerder bijvoorbeeld dat de Google vertaalmachine vooringenomenheid vertoont bij de vertaling van de Turkse, genderneutrale zin 'hij/zij/het is een ingenieur, hij/zij/het is een kok'. Google vertaalt dit naar 'hij is een ingenieur' en 'zij is een kok' (Morse 2017). Daarmee draagt de vertaalmachine onder andere bij aan het naturaliseren van het in de maatschappij circulerende vooroordeel dat technische beroepen bij mannen horen en vrouwen beter in de keuken kunnen staan (Ibid.).

3.2.4 De werking van zoekalgoritmen

Bovenstaande voorbeelden en inzichten uit onderzoek geven aan dat genderstereotypen in taal een grote rol kunnen spelen in het proces van een werkzoekende. Gendergekleurde taal kan een werkzoekende belemmeren in de zoektocht naar werk. Dit omvat ook de teksten op vacaturewebsites en het proces dat een werkzoekende doorloopt wanneer deze gebruikmaakt van een vacaturewebsite om een geschikte baan te zoeken. Idealiter zou een zoekalgoritme op een vacaturesite geen vooringenomenheid bevatten en zelfs de bestaande vooringenomenheid uit de samenleving (of in dit geval: van de individuele werkzoekende) tegengaan door het te neutraliseren. Wanneer een gebruiker van een vacaturewebsite een input geeft door een woord in de zoekbalk te

tikken, beoordeelt een zoekalgoritme alle vacatures op de betreffende website aan de hand van die zoekterm en geeft ze een relevantiescore. Alle vacatures die een relevantiescore hebben die boven een bepaalde drempelwaarde uitkomt, worden getoond als zoekresultaat. Hoe het zoekalgoritme dat precies doet, kan op verschillende manieren. Ter illustratie worden hieronder enkele verschillende typen zoekalgoritmen toegelicht.

Wanneer een werkzoekende een bepaalde zoekfunctie ingeeft zoals 'kokkin', dan zou een algoritme op verschillende manieren kunnen reageren. Een zoekalgoritme zou bijvoorbeeld alle vacatures kunnen tonen die het woord 'kok' bevatten. Dan zou het zoekalgoritme gebaseerd kunnen zijn op woordstammen en woordfamilies. In dat geval is de kern van het algoritme een taxonomie waarin de stammen van alle mogelijke zoekwoorden de kern vormen. Ingevoerde zoektermen worden dan gerelateerd aan de stam van het ingevoerde woord, dat het algoritme uit haar taxonomie haalt. Bijvoorbeeld, wanneer gezocht wordt met 'kokkin'; 'koken' of 'kok', worden deze zoekwoorden allemaal gerelateerd aan het stamwoord 'kok'. De resultaten die getoond worden zullen hetzelfde zijn voor alle zoektermen die relateren aan dezelfde stam. De taxonomie van een dergelijk algoritme kan geautomatiseerd worden gecreëerd (Sanchez & Moreno 2004).

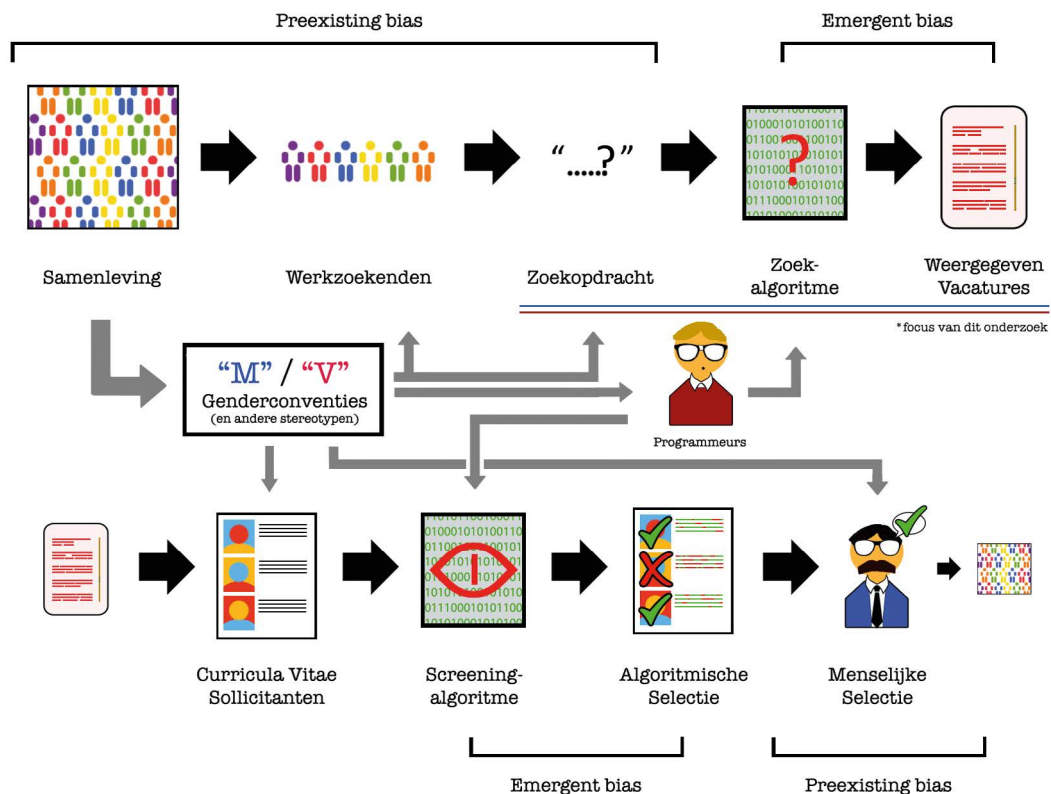
Ten tweede kan een zoekalgoritme gebruik maken van 'word embeddings': een techniek die geavanceerder is dan op basis van woordstammen en woordfamilies. Bij 'word embeddings' krijgen woorden/taal een wiskundige representatie in bijvoorbeeld honderden of duizenden dimensies. Hiermee kan 'echte' semantische overeenkomst gevonden worden, en dat gaat nog een stap verder dan het vergelijken met woordstammen (Ibid.). Ter illustratie: 'een collega' of 'iemand met wie je vaak op kantoor zit' zou semantisch vergelijkbare scores moeten krijgen, ondanks dat je met een simpel synoniem of woordstam hier niet uit zou komen. Een voorbeeld in de context van dit onderzoek: 'docent' en 'baan in het onderwijs' zouden middels dit type algoritme een vergelijkbare score krijgen.

Het is ook mogelijk dat een zoekalgoritme niet met woordstammen of semantische scores werkt, maar puur en alleen zoekt op het woord dat direct in de zoekbalk is ingegeven. Een zoekfunctie ingeven zoals 'schoonmaakster' zal dan alleen vacatures tonen met deze term erin. Een maatschappelijk stereotype dat een schoonmakersfunctie vooral voor vrouwen is, wordt in dit geval door het algoritme gereproduceerd.

Al deze verschillende mogelijke werkingen van zoekalgoritmen, hebben invloed op hoe genderstereotypen gereflecteerd worden op de vacaturesite bij het uitvoeren van een zoekopdracht. Bepaalde typen algoritmen zouden hier beter (dat wil zeggen, minder vooringenomen) mee om kunnen gaan dan andere door stereotypen tegen te gaan.

3.3 Onderzoek in de context van het gehele sollicitatieproces

Op dit moment bestaat er nog geen onderzoek naar dit stereotyperende neveneffect van zoekalgoritmen op vacaturewebsites. Om die reden neemt fase 2 van dit onderzoek juist een stap verder ten opzichte



FIGUUR 1

Maatschappelijke en algoritmische vooroordelen binnen het sollicitatieproces

Bovenstaand schema geeft het proces van zoeken naar vacatures op vacaturewebsites (bovenste rij) en het online sollicitatieproces weer (onderste rij) en maakt inzichtelijk waar vooringenomenheid een rol kan spelen. Vanuit de samenleving komen conventies en stereotypen naar voren, die op verschillende plekken invloed uitoefenen op het sollicitatieproces. Zo kunnen genderconventies maken dat sommige vrouwen niet zoeken op bepaalde functies, niet reageren op bepaalde vacatures of zelfs überhaupt niet gaan werken. Dergelijke conventies kunnen ook invloed uitoefenen op de denkbeelden van programmeurs, wiens vooringenomenheid vervolgens door kan worden gevoerd in zoekalgoritmen en screeningalgoritmen; deze algoritmen beïnvloeden respectievelijk het proces van zoeken naar vacatures en het sollicitatieproces. Conventies en stereotypen hebben uiteraard ook invloed op menselijke besluitvorming. Het gehele proces heeft invloed op de samenleving; zodoende wordt vooringenomenheid vanuit het proces teruggevoerd op de samenleving, met als gevolg dat conventies en stereotypen kunnen worden versterkt.

van eerder onderzoek en richt deze fase zich op de mate waarin het invoeren van gegenderde zoektermen wel of juist niet leidt tot gegenderde vacatures. Daarmee richt dit onderzoek zich dus op twee verschillende, doch gerelateerde processen waarbinnen vooringenomenheid kan opduiken: het zoekproces van werkzoekenden naar vacatures, en het proces van genderstereotypering in functieomschrijvingen die naar voren komen op basis van zoektermen. In Figuur 1 is een schematische weergave gegeven van het gehele proces zoals in dit theoretische kader uiteen is gezet.

4 Onderzoeksaanpak

Utrecht Data School, Universiteit Utrecht, heeft een exploratief onderzoek opgezet en uitgevoerd om het bestaan en de omvang van eventuele discriminatie of vooringenomenheid door algoritmen op vacaturewebsites vast te stellen. Binnen dit onderzoek zijn in december 2020 vier vacaturewebsites onderzocht (hierna aangeduid met A, B, C en D). Drie van de vier maken deel uit van de top vijf grootste vacaturewebsites van Nederland. De laatste is gekozen omdat het zich richt op een specifieke doelgroep, namelijk hoogopgeleiden: zodoende hoopten we om eventuele verschillen in mate van discriminatie tussen arbeidsmarkten die zich richten op hogeropgeleiden versus arbeidsmarkten die daar minder op gericht zijn te doorgronden.

4.1 Fase I

Tijdens fase I probeerden we te achterhalen hoe de zoekalgoritmen op vacaturewebsites omgaan met persoonlijke informatie van werkzoekenden, die zij zelf in hun profiel op de desbetreffende websites in kunnen vullen. Waar we specifiek naar keken is of en hoe het zoekalgoritme resultaten (vacatures) van zoekopdrachten koppelt aan persoonskenmerken van de werkzoekende (uit het profiel).

We onderzochten de websites door middel van 'reverse engineering': door verschillende inputs te geven in de vorm van profielinformatie en zoektermen, en de output – in de vorm van vacatures die als zoekresultaat terugkomen – te meten. Hierdoor probeerden we de werking te achterhalen van de zoekalgoritmen van de vacaturewebsite.

Onze werkwijze bestond uit het aanmaken van een aantal profielen op de vacaturewebsites. Die profielen bevatten informatie die direct of indirect iets zeggen over een kenmerk wat beschermd wordt door een verboden discriminatiegrond. Dat waren:

- geslacht
- etniciteit
- leeftijd

Met etniciteit duiden wij op het concept 'ras', dat in de wet wordt genoemd als beschermde discriminatiegrond. Om deze drie discriminatiegronden te kunnen onderzoeken, zijn er twaalf verschillende nep-profielen aangemaakt. Omdat de onderzochte vacaturewebsites gebruikers niet in staat stellen om expliciet geslacht, etniciteit en leeftijd aan te geven, is ervoor gekozen om de profielen van onze nepsollicitanten te voorzien van voor- en achternamen die doorgaans indicatief zijn voor

iemands geslacht of etniciteit.⁶ Leeftijd werd op twee manieren aan de profielen toegevoegd. Enerzijds werd aan alle accounts een Gmail account gekoppeld waar wel direct leeftijd kon worden ingevoerd. Anderzijds werd op de vacaturewebsites aangegeven hoeveel werkervaring de nep-personen hadden en wanneer zij hadden gestudeerd. Zodoende zagen de variaties in de nep-profielen er als volgt uit:

- Profiel met een mannelijke, Nederlandse naam, 30 jaar, 10 jaar werkervaring
- Profiel met een mannelijke, niet-Nederlandse naam, 30 jaar, 10 jaar werkervaring
- Profiel met een vrouwelijke, Nederlandse naam, 30 jaar, 10 jaar werkervaring
- Profiel met een vrouwelijke, niet-Nederlandse naam, 30 jaar, 10 jaar werkervaring
- Profiel met een vrouwelijke, Nederlandse naam, 45 jaar, 20 jaar werkervaring
- Profiel met een vrouwelijke, Nederlandse naam, 23 jaar, 1 jaar werkervaring

Op de vacaturewebsites moest tevens de opleiding en eerdere functies worden aangegeven. Bij alle profielen is daar hetzelfde opgegeven: psychologie aan de Hogeschool van Utrecht, met werkervaring als psycholoog. Volgens de opgegeven informatie waren geen van de nep-personen nog werkachtig bij hun vorige werkgever. Voor elk van de bovenstaande variaties zijn er twee profielen aangemaakt, waardoor het totaal uitkwam op twaalf nep-profielen.⁷ De reden voor het aanmaken van twee profielen per variatie is dat op die manier een extra check wordt verkregen voor de invloed van willekeur op de zoekresultaten.

Nadat we twaalf profielen op alle vier de websites hadden aangemaakt, stelden we de zoektermen vast waarmee we de profielen zoekopdrachten lieten uitvoeren. Zaak was om zoektermen te identificeren die spreiding hadden over verschillende sectoren, een aantal termen dat specifiek duidde op mannelijke functies, een aantal specifiek gericht op vrouwelijke functies en een aantal 'genderneutrale' termen. De zoektermen hebben we vastgesteld in overleg met het College. De zoektermen zijn:

- Timmerman
- Voorman
- Eindredactrice
- Secretaresse
- Programmeur
- Communicatieadviseur
- Accountmanager
- Servicemedewerker
- Leiding productie
- Directie

⁶ Blijkens het rapport uit 2021 van de Inspectie SZW geven sommige vacaturesites wel de mogelijkheid om kandidaten te filteren op basis van beschermde discriminatiegronden. Dat lijkt dus niet het geval bij de vier websites die in dit onderzoek zijn meegenomen.

⁷ Zie Appendix voor een weergave van de twaalf gebruikte profielen.

Deze tien zoektermen zijn door alle twaalf profielen ingevoerd op de vacaturewebsites, in december 2020. Het uitvoeren van zoekacties is handmatig gedaan. Dit komt neer op 120 zoekacties per website en 480 zoekacties in totaal. Het uitvoeren van zoekacties is niet geautomatiseerd, omdat in de Algemene Voorwaarden van drie van de vier websites het gebruik van scraping tools expliciet werd verboden.⁸

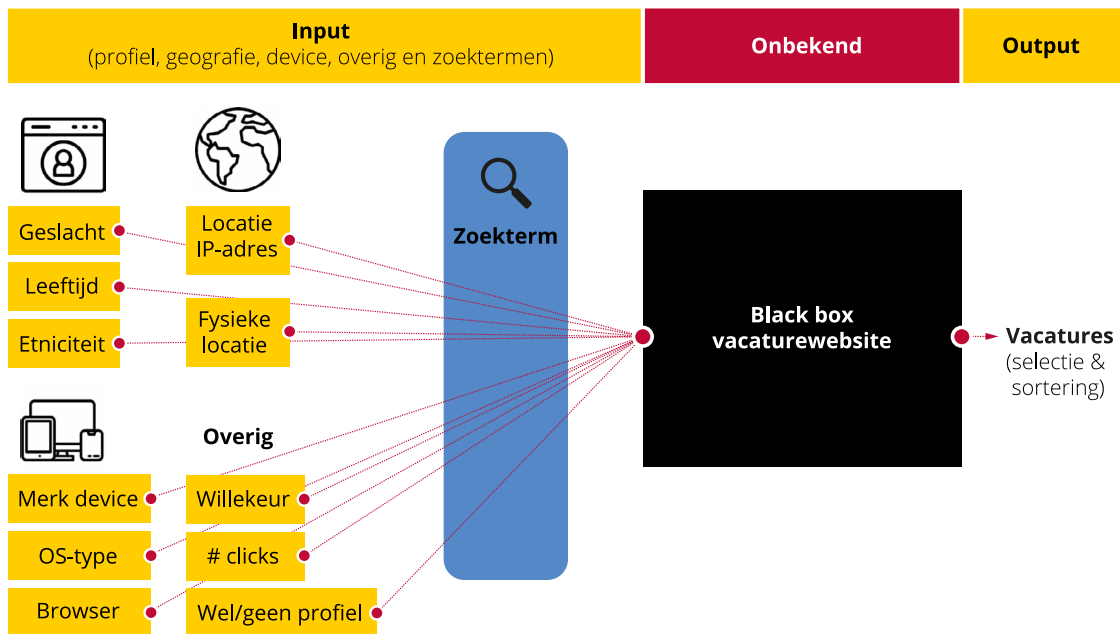
De werkzoekende heeft zelf veel mogelijkheden om resultaten te filteren en te rangschikken, bijvoorbeeld op basis van geografische positie. Bij elke website is voor de filteroptie 'relevantie' gekozen. Dit betekent dat het zoekalgoritme van de vacaturewebsite zelf bepaalt wat het meest relevant is voor de gebruiker en op basis daarvan een lijst aan zoekresultaten toont. De output van elke zoekactie is vastgelegd: hierbij zijn de bovenste drie resultaten opgeslagen. Dit leverde 360 gevonden vacatures op per website. Per zoekresultaat zijn drie kenmerken van de vacature opgeslagen: de titel, of de vacature een geadverteerde vacature is en of de vacature als 'nieuw' wordt aangemerkt.

Aandachtspunten bij het verzamelen van de data waren verschillen in inputfactoren. Drie verschillen in input hebben we bewust aangebracht: namelijk verschillen in de profielen gebaseerd op de drie eerder genoemde discriminatiegronden. Daarnaast kunnen er andere verschillen bestaan in inputfactoren. Zelfs met drie identieke profielen zouden de requests, de zoekopdrachten, op de website verschillen in tijd. De tien verschillende zoekopdrachten kunnen immers niet allemaal op exact hetzelfde tijdstip gedaan worden. Ook de volgorde waarin de zoekopdrachten gedaan werden kan invloed hebben op de output. Daarnaast kan informatie in cookies invloed hebben, waardoor er een relatie tussen request één, twee, ..., N zit. Tenslotte zijn andere uitgangspunten zoals locatie, IP-adres en het type besturingssysteem (OS) van belang, omdat die ook invloed kunnen hebben op de output.

Om deze mogelijke beïnvloedende factoren te ondervangen hebben we een aantal strategieën toegepast bij de dataverzameling. We noteerden bij elke zoekopdracht het precieze tijdstip en de datum van de zoekopdracht, om de factor 'tijd' in kaart te brengen en hier eventuele verschillen in output aan te kunnen relateren. Alle cookiegegevens zijn gewist voor elke zoekopdracht en de incognitomodus werd gebruikt bij het uitvoeren van de zoekopdrachten. Het IP-adres werd gelijk gehouden. Per account van werkzoekende en per website werden alle zoekopdrachten zo kort mogelijk achter elkaar uitgevoerd. Tegelijkertijd ingelogd zijn op meerdere accounts in verschillende tabbladen bleek niet mogelijk te zijn. Daarnaast randomiseerden we de volgorde van de zoekopdrachten en het gebruik van de verschillende profielen.

De werkwijze wordt schematisch weergegeven in figuur 2.

⁸ Scraping en het gebruik van crawlers wordt door een groot aantal vacaturewebsites expliciet verboden in hun Algemene Voorwaarden, waaronder drie van de vier onderzochte websites in dit onderzoek. De vierde website noemt crawlen of scrapen niet expliciet maar verbiedt wel data mining en robots voor het onttrekken van gegevens.



FIGUUR 2

Schematische werkwijze onderzoeksanpak fase I

Er bestaan verschillende factoren binnen de input voor het zoekalgoritme op vacaturewebsites. Deze kunnen allemaal invloed hebben op de waargenomen output. Het in kaart brengen en zoveel als mogelijk ondervangen van mogelijke inputfactoren zorgt ervoor dat beter geanalyseerd kan worden waardoor verschillen in output veroorzaakt worden. Binnen dit onderzoek zijn de volgende inputfactoren beheerd: Profielinformatie (geslacht, leeftijd, etniciteit), locatie (IP-adres en fysieke locatie), Device informatie (merk device, OS-type en browser), wel/geen profiel en aantal clicks. Alleen de inputfactor willekeur kon niet worden beheerd (dat is pas enigszins te corrigeren bij een zeer grootschalig, geautomatiseerd onderzoek).

4.2 Fase II

Voor de tweede fase is in overleg met het College besloten om nader onderzoek te doen naar de omgang van zoekalgoritmen met gendergekleurde zoektermen en functietitels. Hierbij hebben we besloten vacaturesite D niet mee te nemen gezien de sterke overeenkomsten met vacaturesite C (en hun gedeelde moederbedrijf).

Met 'gendergekleurde' functietitels bedoelen wij functietitels die duidelijk doelen op mannelijke of vrouwelijke personen, zoals 'secretaresse'. Dit woord duidt van oudsher een vrouw aan, ook al is inmiddels deze perceptie enigszins aan het verschuiven.⁹ Wanneer er wordt gezocht op 'secretaresse', zullen sommige vacaturewebsites vacatures laten zien die ook het woord 'secretaresse' in de functietitel hebben. In deze gevallen worden vacatures met 'secretariaatsmedewerker', 'office manager' en 'management assistent' – de meer genderneutrale varianten van secretaresse – niet getoond. Andere websites laten juist wél ook deze vacatures zien bij de zoekresultaten van 'secretaresse'. Dit geeft

⁹ Zie <https://onzetaal.nl/taaladvies/secretaaris-secretaresse/#:~:text=Een%20secretaresse%20kan%20zowel%20een,maar%20deze%20betekenis%20is%20verouderd>

aanleiding om te onderzoeken hoe vacaturewebsites de gegenderde taal die bezoekers gebruiken actief neutraliseren, of in andere gevallen juist zouden kunnen reproduceren doordat bestaande stigma's 'genaturaliseerd' worden. Vooringenomenheid treedt niet alleen op door algoritmen, maar doordat mensen ook bias hebben over zichzelf (zoals al eerder genoemd in het theoretisch kader). Of en in welke mate vacaturewebsites hieraan bijdragen, hangt af van de werking van hun zoekalgoritme.

Om dit te onderzoeken hebben we een lijst samengesteld van vijftien gendergekleurde functietitels en hun neutrale tegenhanger(s).¹⁰ Deze functietitels werden als zoekterm ingevoerd op de verschillende websites. Bij het uitvoeren van de zoekopdrachten was de onderzoeker niet ingelogd met een bepaald profiel.

Drie van de vijftien gendergekleurde functietitels hebben we uitgebreider onderzocht. Dat waren 'secretaresse', 'leraar' en 'lerares', en 'gastvrouw', samen met hun genderneutrale tegenhangers. Secretaresse hebben we gekozen omdat het een veel voorkomende, gangbare functietitel is die ogenschijnlijk gendergekleurd is. Ondanks dat mannen in een dergelijke functie zelden worden aangeduid met 'secretaresse', kan de term tegenwoordig op zowel vrouw als man slaan. Aannemelijk is echter dat de associatie vooral met het vrouwelijk geslacht valt. Bij de keuze voor de andere twee termen hebben we wederom gelet op de gangbaarheid als functietitel en op de spreiding over verschillende sectoren. Bij zoekopdrachten met die functietitels hebben we de titel en functieomschrijving van de eerste tien zoekresultaten opgeslagen. Omwille van het exploratieve karakter van dit onderzoek, zijn van de overige twaalf gendergekleurde functietitels slechts de bovenste drie vacatures die als zoekresultaat verschijnen opgeslagen, eveneens met titel en functieomschrijving. Aanvullend is een kwantitatieve check gedaan waarbij het totale aantal zoekresultaten per zoekterm is opgeslagen.

Daarnaast zijn ook zoekopdrachten uitgevoerd waarin we de tegenhangers gebruiken van de gendergekleurde functietitels, de 'genderneutrale' varianten zoals 'secretariaatsmedewerker', 'docent' en 'gastvrouw/-heer'. Bij de eerder genoemde drie functietitels hebben we de tien bovenste zoekresultaten opgeslagen. Bij de andere twaalf neutrale tegenhangers beperkten we ons tot de bovenste drie vacatures.

Ten slotte hebben we het totale aantal zoekresultaten van de gebruikte zoekopdrachten bekeken. Dat wil zeggen, het cijfer dat aangeeft hoeveel vacatures er in totaal op de onderzochte website staan die gerelateerd zijn aan de gebruikte zoekopdracht.

10 Zie Appendix

5 Resultaten

De resultaten van het onderzoek naar de vacaturesites, aangeduid als A t/m D, worden hieronder weergegeven. De werking van de zoekalgoritmen op betreffende websites zijn onderzocht aan de hand van zoekacties waarbij functietitels zijn ingevoerd in de zoekbalk van de websites. Tijdens fase I is zijn de zoekacties uitgevoerd terwijl de onderzoeker ingelogd was op de website. Daar zijn nepprofielen van werkzoekenden voor aangemaakt. Tijdens fase II was de onderzoeker niet ingelogd op een specifiek profiel tijdens het uitvoeren van de zoekacties. De resultaten van de zoekacties, in de vorm van weergegeven vacatures in een lijst, zijn geanalyseerd.

5.1 Fase I

Tijdens deze fase hebben we gekeken naar het bestaan van een verband tussen verschillen in profielkenmerken van werkzoekenden en verschillen in zoekresultaten die deze profielen te zien kregen. Wanneer profielen met verschillende kenmerken op het gebied van geslacht, ras of leeftijd andere zoekresultaten te zien krijgen, zou dat kunnen duiden op ongelijke behandeling.

De eerste stap van de analyse was het vergelijken van de vacatures die de verschillende profielen te zien krijgen bij dezelfde zoekactie. Zitten daar verschillen in? Wat zijn die verschillen? Hierbij waren twee uitkomsten denkbaar:

1. Er zijn geen verschillen zichtbaar in de vacatures die de verschillende profielen met dezelfde zoekactie te zien krijgen. Of de verschillen zijn minimaal en kunnen herleid worden naar zaken als het tijdstip van de zoekopdracht of andere inputfactoren waar rekening mee is gehouden.
2. Er zijn verschillen zichtbaar in de vacatures die verschillende profielen te zien krijgen, gebruikmakend van dezelfde zoekterm. In dat geval is nader onderzoek naar desbetreffende vacaturewebsite gewenst.

Het is belangrijk om de invloed van willekeur te benadrukken. Veel algoritmen, bijvoorbeeld bij recommender systems, laten (bewust) een vorm van willekeur toe. Door populaties met verschillende input in contact te laten komen, kan hiermee ook geleerd worden over hoe mensen daarop reageren (net als met A-B-testing). Het feit dat je een ander resultaat observeert, hoeft dus niets te zeggen over vooringenomenheid en vooraf ingeprogrammeerde zaken.

Geen van de onderzochte websites geeft de werkzoekende de gelegenheid om directe informatie in te vullen over geslacht, leeftijd of afkomst. Daardoor kon alleen op een indirecte manier informatie over deze drie kenmerken ingevuld worden in de profielen. Andere bevindingen rangschikken we hieronder per onderzochte vacaturewebsite.

5.1.1 Vacaturesite A

Bij vacaturesite A viel op dat er bij de bovenste drie zoekresultaten geen geadverteerde vacatures werden getoond. Daarnaast was het bovenste zoekresultaat altijd gelijk, onafhankelijk van welk profiel de zoekopdracht uitvoerde. Bij het tweede en derde zoekresultaat was een klein verschil in output te zien: 6% van de zoekresultaten op de tweede en derde positie waren 'unieke' vacatures, de overige zoekresultaten kwamen overeen met de zoekopdrachten van andere profielen. Deze verschillen konden niet worden gelinkt aan profielkenmerken: er was geen verband te vinden tussen de ingegeven profielinformatie over leeftijd, achtergrond of geslacht en het zoekresultaat. We vermoeden dat het tijdstip van de zoekopdracht of een bepaalde factor van willekeur in het algoritme hierbij een rol heeft gespeeld.

5.1.2 Vacaturesite B

Vacaturesite B vertoonde exact dezelfde zoekresultaten bij alle profielen. Kortom, het maakte niet uit met welk profiel de zoekopdracht werd uitgevoerd, er kwamen altijd dezelfde zoekresultaten tevoorschijn. Dit was alleen zo voor niet-geadverteerde vacatures. Vacatures die waren geadverteerd, bij vacaturesite B altijd op positie twee in de lijst met zoekresultaten, verschilden bij sommige zoekopdrachten wel. De oorzaak van de verschillen in output bij geadverteerde vacatures hebben wij niet kunnen achterhalen, in ieder geval hebben we geen link gevonden met de profielkenmerken.

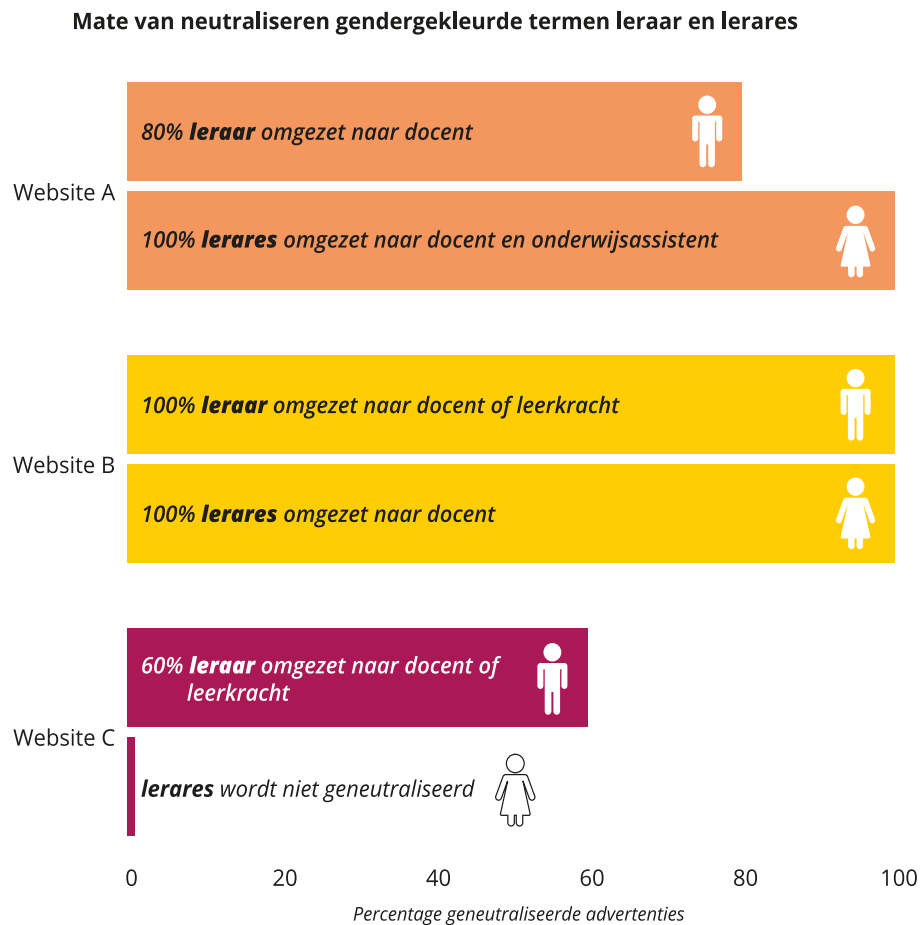
5.1.3 Vacaturesites C en D

Vacaturesites C en D vertoonden gelijksoortige resultaten: ze lieten geen noemenswaardige verschillen in zoekresultaten zien wanneer die uitgevoerd worden door verschillende profielen. Alleen bij een klein aantal geadverteerde vacatures was dit het geval, deze waren echter niet te herleiden naar verschillen in profielkenmerken of andere inputfactoren zoals bijvoorbeeld tijd, maar vloeien waarschijnlijk voort uit advertentiemogelijkheden. Waarschijnlijk komt dit doordat deze twee websites hetzelfde moederbedrijf hebben en algoritmen gebruiken die sterk op elkaar lijken.

Op basis van deze fase van het exploratieve onderzoek hebben wij geen indicaties gevonden die wijzen op het discrimineren op basis van profielkenmerken door vacaturewebsites. Wat opvallend was bij de analyse van de resultaten, was het gebruik van 'gendergekleurde' of 'gegenderde' functietitels. Daarom is ervoor gekozen om ons tijdens fase II op dit fenomeen te richten.

5.2 Fase II

Tijdens fase II van het onderzoek zijn we dieper in de precieze werking van de zoekalgoritmen gedoken. De bevindingen zullen wij hieronder illustreren aan de hand van een aantal concrete voorbeelden, te beginnen met de zoektermen 'leraar' en 'lerares'. Ter herhaling: gekeken wordt naar de mate waarin de zoekalgoritmen vooringenomenheid tegengaan (oftewel neutraliseren), vooringenomenheid reproduceren (oftewel naturaliseren) of zelfs vooringenomenheid versterken. Neutraliseren betekent

**FIGUUR 3**

Percentage neutraal geformuleerde zoekresultaten bij gebruik van gendergekleurde zoekterm

in de praktijk dat wanneer een gendergekleurde zoekterm ingevuld wordt, de vacaturewebsite zoekresultaten laat zien waarin een genderneutrale functietitel getoond wordt. Het versterken van vooringenomenheid is uiteraard ongewenst en het tegengaan (neutraliseren) van vooringenomenheid is gewenst. Ook het reproduceren, of naturaliseren, van vooringenomenheid levert een punt van aandacht op: het zoekalgoritme normaliseert dan als het ware de bestaande vooringenomenheid die heerst in de samenleving.

5.2.1 Mate van neutraliseren: concrete voorbeelden

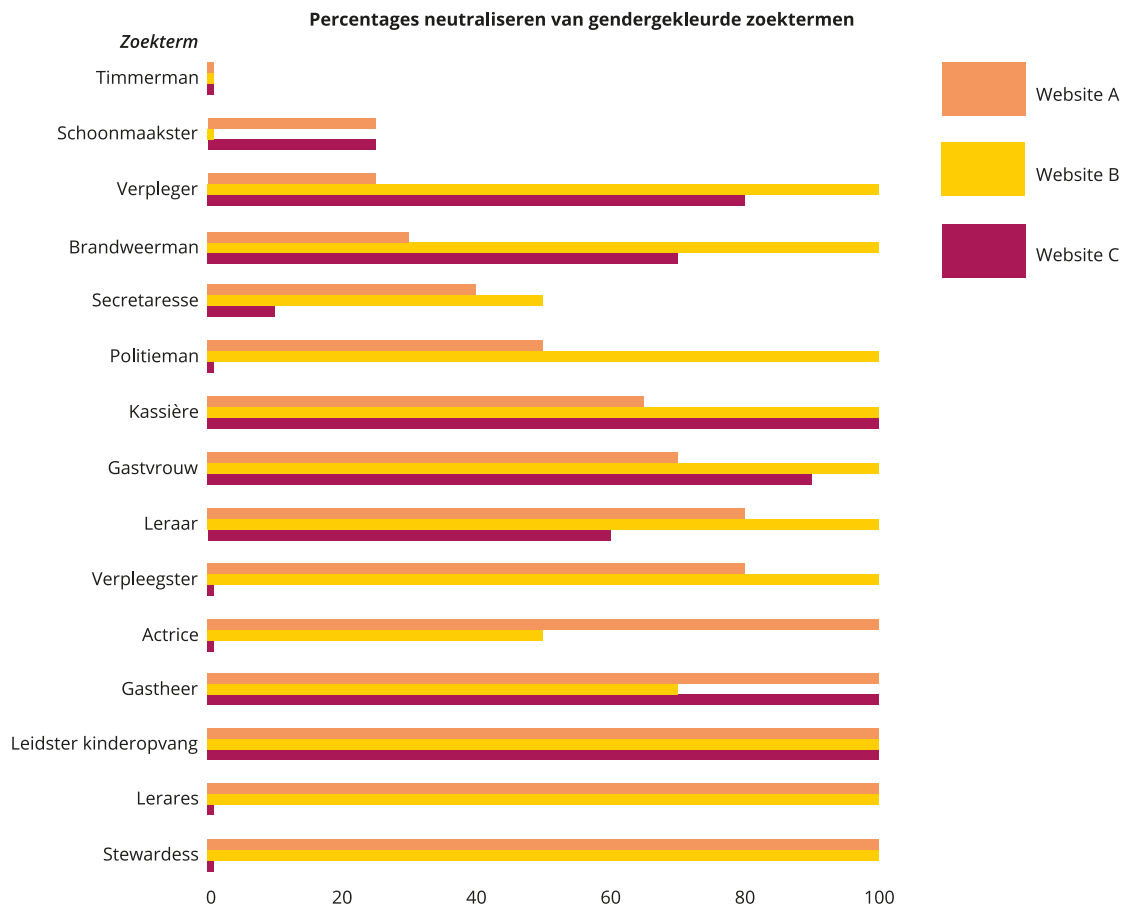
Figuur 3 laat zien of de vacaturesites binnen de onderzochte advertenties de gendergekleurde term 'leraar' en 'lerares' neutraliseren (oftewel vacatures met een genderneutrale functietitel tonen). Hiermee gaan ze vooringenomenheid tegen. Zowel vacaturesite A als B lijken binnen de, door ons onderzochte, advertenties een hoog percentage van de gendergekleurde zoektermen te neutraliseren (80 - 100%). Bij de vacaturesite C ligt dit percentage veel lager. Daarnaast lijkt vacaturesite C bij te dragen aan het reproduceren van vooringenomenheid door 'leraar' wel redelijk vaak te neutraliseren

tot vacatureresultaten voor een 'docent', terwijl 'lerares' geen enkele keer wordt geneutraliseerd. Dit betekent dat wanneer gezocht wordt met de term 'lerares', de vacaturesite enkel vacatures toont die letterlijk de term 'lerares' herhalen en de vacatures voor een docent of leraar niet toont. Zo wordt de persoon die zoekt uitgesloten van het zien op deze vacatures, terwijl iemand die zoekt met 'leraar', de vacatures te zien krijgt voor leraar én docent. Hetzelfde zien we bij vacaturesite A waarbij de zoekterm 'lerares' resulteerde in een getoonde vacature voor een 'onderwijsassistent', een term die in feite niet genderneutraal is en qua functie niet hetzelfde is als een docent maar een functie die onder de docent of leerkracht valt. Dit was het geval bij slechts één zoekresultaat van de tien, we kunnen daarom de uitkomst niet generaliseren maar slechts aangeven dat het een aandachtspunt zou kunnen zijn.

De uitkomst van het onderzoek met de termen 'secretaresse' levert een vergelijkbaar beeld op. Vacaturesite B neutraliseert het meeste, in 50% van de gevallen wordt 'secretaresse' omgezet naar een meer genderneutrale functietitel. Vacaturesite A volgt daarop met 40% en vacaturesite C neutraliseert nauwelijks met maar één van de tien zoekresultaten die een andere term dan 'secretaresse' bevat (namelijk een vacature met de functietitel 'administratief medewerker/secretaresse').

De zoekacties op de drie websites met de term 'gastvrouw' geven in de grote meerderheid van de gevallen vacatures terug met een genderneutrale functietitel. De getoonde vacatures zijn voor 'gastvrouw/gastheer', 'host(ess)' of 'host m/v'. Bij deze zoekterm zijn geen noemenswaardige verschillen gevonden tussen de resultaten van de drie websites.

Opmerking verdient dat Figuur 3 op een beperkte dataset is gebaseerd: het gaat hierbij slechts om de analyse van de tien bovenste zoekresultaten bij gebruik van de zoektermen. Hierbij zijn zowel de functietitel als de vacaturetekst onderzocht op het bevatten van gendergekleurde, danwel genderneutrale taal. In paragraaf 5.2.3 checken we deze bevindingen door een analyse te doen van het totale aantal zoekresultaten dat getoond wordt bij gebruik van eerdergenoemde zoektermen (zie Figuur 5).

**FIGUUR 4**

Percentage neutraal geformuleerde zoekresultaten bij gebruik van gendergekleurde zoekterm

5.2.2 Mate van neutraliseren: totaalbeeld

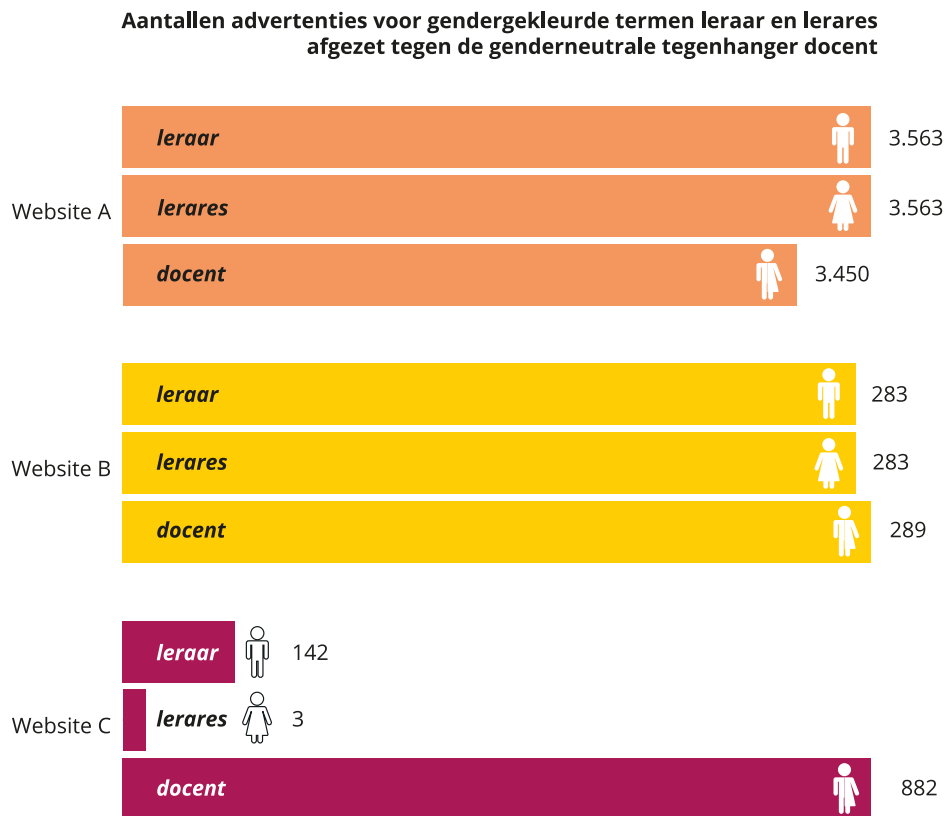
Figuur 4 laat voor alle, door ons onderzochte, gendergekleurde termen zien in welke mate de vacaturesites deze termen in de bekeken advertenties neutraliseren. Op de y-as zien we de gendergekleurde zoektermen die zijn ingevuld in de zoekbalken van de vacaturewebsites. Op de x-as zien we het percentage zoekresultaten dat geneutraliseerd wordt. Dit betekent in de praktijk dat wanneer een gendergekleurde zoekterm ingevuld wordt, de vacaturewebsite zoekresultaten laat zien waarin een genderneutrale functietitel getoond wordt. Dus wanneer je op vacaturesite B zoekt naar 'verpleegster', krijg je alleen zoekresultaten te zien waarin de functietitel 'verpleegkundige' te zien is. Hetzelfde gebeurt wanneer je zoekt met de zoekterm 'verpleger'. Dit is te zien in de linkerkolom van Figuur 4, in de onderste twee balkjes. Conform Figuur 3 is te zien dat vacaturesite C de zoekterm 'lerares' nooit omzet in een genderneutrale variant. Dat is ook in Figuur 4 te zien in de rechterkolom, waar het het balkje van 'lerares' leeg is.

Als we inzoomen op de linkerkolom, van vacaturesite B, zien we dat deze website de meeste gendergekleurde zoektermen voor 100% neutraliseert. Sommige van deze zoektermen leveren echter een vertekend beeld op, vanwege de gangbaarheid van de onderzochte zoekterm: hoe vaak vacatures met die functie überhaupt voorkomen op de vacaturewebsite. Hoe minder vacatures aanwezig zijn, hoe minder betrouwbaar het beeld dat in bovenstaande figuur wordt geschetst. Een zoekterm als 'actrice' levert weinig resultaten op, want er zijn niet veel vacatures voor deze term. Hetzelfde geldt voor 'brandweerman' (tussen de nul en 53 vacatures), maar ook 'politieman' (tussen de nul en negentien vacatures) en 'stewardess' (tussen de twee en acht vacatures). Daarnaast moet er vermeld worden dat er ook functietitels worden gebruikt anders dan de onderzochte benamingen voor deze functie. Bijvoorbeeld 'gastvrouw/-heer' wordt ook aangeduid als 'host', 'service medewerker', 'hospitality medewerker', en 'kassière' onder andere als 'medewerker klantenservice'. Daarom kunnen we geen verregaande conclusies trekken uit de resultaten die gebaseerd zijn op zoekacties met deze termen.

De zoektermen secretaresse, schoonmaakster en verpleegster worden op alle vacaturesites weinig geneutraliseerd, terwijl er wel een (redelijk) ruime hoeveelheid vacatures voor deze beroepen aangeboden wordt. De zoekterm timmerman levert zelfs in geen van de gevallen een vacature op met een genderneutrale term (zoals houtbewerker), maar alleen vacatures die de term timmerman letterlijk herhalen. Zeker bij secretaresse en timmerman lijkt de inhoud van de advertentie ook wat meer te zijn geschreven voor een vrouwelijk dan wel een mannelijk publiek. Zo worden in de advertenties voor secretaresse regelmatig beschreven met welke secretaresses (met vrouwennaam genoemd) de sollicitant gaat werken. Dit ondersteunt de aanname dat secretaresse vermoedelijk nog veelal als een vrouwelijke functie wordt opgevat. Opvallend is dat een zoekfunctie met 'secretaresse' over de drie vacaturesites samengenomen ruim veertig keer zoveel vacatures opleverde als de neutralere zoekfunctie 'secretariaatsmedewerker'.

5.2.3 Kwantitatieve check

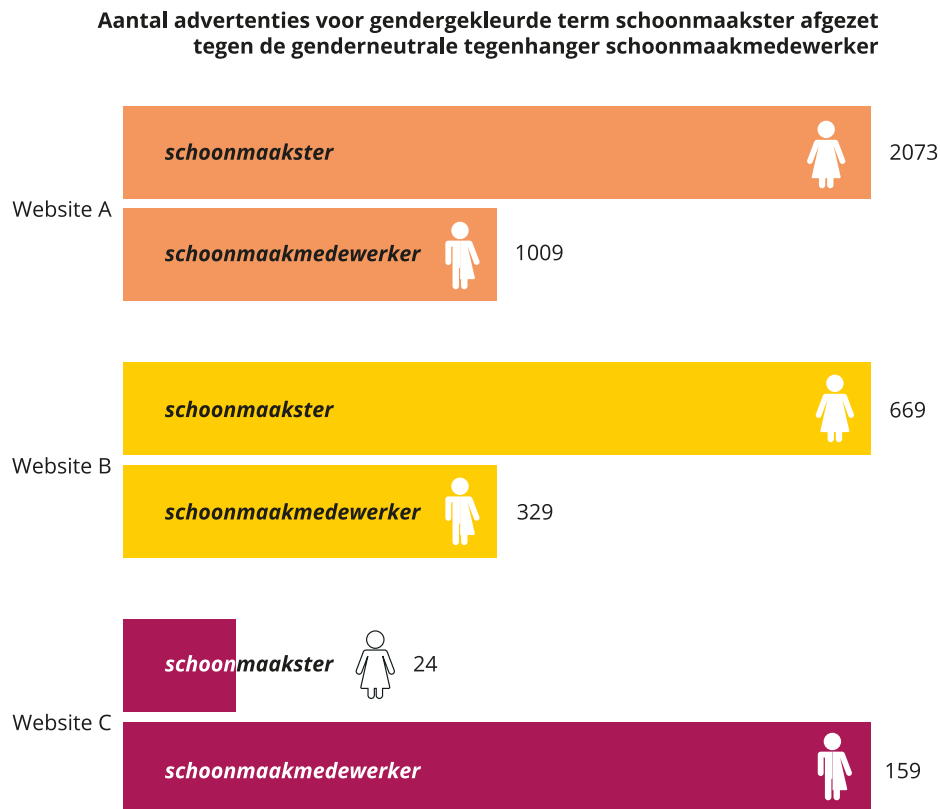
Vijftien gendergekleurde zoektermen en vijftien genderneutrale zoektermen zijn kwalitatief onderzocht. Om de bevindingen uit dit kleinschalige onderzoek verder te onderbouwen, zijn daarnaast ook voor alle gendergekleurde termen met hun genderneutrale tegenhanger het aantal resultaten per zoekterm genoteerd. Een vacaturesite, die geen groot verschil laat zien in resultaat voor een gendergekleurde term en een genderneutrale term, werkt hoogstwaarschijnlijk met een algoritme waarbij de gendergekleurde en genderneutrale term in één categorie in de taxonomie zijn opgenomen. Dit zorgt er dan voor dat bij de gendergekleurde term ook alle advertenties worden getoond waar de genderneutrale tegenhanger in voorkomt. Onze aanname daarbij is dat de vergelijking in het aantal zoekresultaten een indicator is van of er wel of niet verder gekeken wordt dan de exacte zoekterm. De resultaten van deze kwantitatieve check worden hieronder in visualisaties afgebeeld.

**FIGUUR 5****Aantallen advertenties voor gendergekleurde termen en hun genderneutrale tegenhanger**

Bovenstaande grafiek laat voor de gendergekleurde termen 'leraar' en 'lerares' en hun genderneutrale tegenhanger het aantal gevonden advertenties voor die zoekterm zien.

Figuur 5 laat zien dat vacaturesite B voor de gendergekleurde termen 'leraar' en 'lerares' en hun genderneutrale tegenhanger 'docent' vrijwel dezelfde aantallen teruggeeft. Bij vacaturesite A zien we ook weinig verschillen in aantallen. Vacaturesite C toont echter grote verschillen in aantallen.

Een tweede voorbeeld van de resultaten naar het aantal advertenties per zoekterm (figuur 6) laat grote verschillen zien in aantallen voor de termen 'schoonmaakster' en 'schoonmaakmedewerker', zelfs bij vacaturesite B. Verdere inspectie toont dat zowel A als B hier 'schoonmaakster' neutraliseren naar 'schoonmaker'. Aangezien dit vervolgens meer resultaten oplevert dan 'schoonmaakmedewerker' lijkt het zo te zijn dat deze term is genaturaliseerd tot de 'neutrale term'. Dit kan betekenen dat de van oorsprong mannelijke benaming (schoonmaker) voor een schoonmaakmedewerker, verschuift naar een meer als neutraal opgevatte benaming.

**FIGUUR 6****Aantallen advertenties voor gendergekleurde termen en hun genderneutrale tegenhanger II**

Bovenstaande grafiek laat voor de gendergekleurde term *schoonmaakster* en de genderneutrale tegenhanger *schoonmaakmedewerker* het aantal gevonden advertenties voor die zoekterm zien.

5.2.4 Vooringenomenheid tegengaan, reproduceren of versterken

Over het algemeen genomen zien we dat vacaturesite B het meest neutraliseert van de drie onderzochte vacaturesites. Wanneer een werkzoekende daar een gendergekleurde zoekterm invult, krijgt hij of zij ook de vacatures te zien met een genderneutrale functietitel. Dit proces gaat bestaande vooringenomenheid van de werkzoekende actief tegen. Vacaturesite A neutraliseert ook maar in minder gevallen dan B. Vacaturesite C neutraliseert het minst van de drie. Het kan zijn dat vacaturesite C een simpel zoekalgoritme gebruikt dat letterlijk zoekt naar de opgegeven zoekterm die is ingegeven door de gebruiker.

Daartegenover staan A en B, die een complexer zoekalgoritme lijken te gebruiken, dat genderneutrale en gendergekleurde functietitels kan groeperen wanneer zij een soortgelijke functie beschrijven. Daardoor komen deze functies in dezelfde *pool* terecht, ook al worden zij op verschillende manieren aangeduid, en dat verklaart waarom de getallen in het figuur hierboven voor de gendergekleurde en genderneutrale

zoektermen nagenoeg gelijk zijn. De manier waarop de zoekalgoritmen van A en B te werk zouden kunnen gaan is door (1) zoektermen te relateren aan een vooraf opgestelde taxonomie waarbij gebruik wordt gemaakt van woordstammen en woordfamilies, of door (2) semantische scores toe te kennen aan vacatures op basis van 'word embeddings' en op die manier inhoudelijk vergelijkbare resultaten terug te kunnen koppelen.

Wanneer wordt gezocht met een genderneutrale functietitel, laten geen van de websites een gendergekleurde functietitel zien in hun zoekresultaten. Dit lijkt erop te wijzen dat gendergekleurde functietitels wel worden geneutraliseerd, maar dat andersom, genderneutrale functietitels niet worden 'gegenderd', of 'gendergekleurd'.

6 Conclusie

Het onderzoek richtte zich op twee verschillende, doch gerelateerde processen waarbinnen ongelijkheid kan opduiken: het zoekproces van werkzoekenden naar vacatures in fase I en het proces van genderstereotypering in functieomschrijvingen dat naar voren komen op basis van zoektermen in fase II. In beide processen worden risico's van uitsluiting verondersteld door de werking van achterliggende zoekalgoritmen. Bij dit onderzoek is onder meer gekeken of zoekalgoritmen op vacaturesites seterotypen op basis van gender, tegengaan, reproduceren of versterken. Daarbij zijn vier vacaturewebsites onderzocht in de maand december 2020.

6.1 Discriminatie op basis van profiel

Wat betreft het zoekproces van werkzoekenden: wij hebben geen indicaties gevonden die wijzen op onderscheid door de vacaturewebsites dat we direct toe kunnen schrijven aan verschil in profielkenmerken. Werkzoekenden met verschillende profielen kregen geen verschillende zoekresultaten te zien die herleid konden worden naar profielkenmerken. Het geslacht, de etniciteit of leeftijd van een kandidaat lijken daarmee geen uitsluitende werking te hebben voor werkzoekenden bij het zoeken naar personeeladvertenties op vacaturesites.

6.2 Genderstereotypen en het zoekalgoritme

Tijdens fase II keken we naar de werking van het zoekalgoritme en de invloed daarvan op genderstereotypen. Fase II was exploratief van aard, omdat het niet mogelijk was om te achterhalen hoe de zoekalgoritmen precies werken (door bijvoorbeeld inzage in de technische beschrijving ervan). Daarom hebben we getracht dat te doen aan de hand van het geven van bepaalde input en vervolgens de output van het algoritme te analyseren.

Wat betreft de werking van de onderzochte zoekalgoritmen zijn er grofweg drie mogelijkheden: het algoritme maakt gebruik van een woordstam taxonomie, het maakt gebruik van semantische scoring van nabijheid, of het zoekt simpelweg alleen naar exact het opgegeven zoekwoord. De eerste optie van een woordstam taxonomie houdt in dat het vacatureplatform gebruik maakt van een taxonomie, waarbij neutrale termen gehanteerd worden, en iedere vacature ingedeeld wordt in de categorie die het meest relevant is. Het zou bij een woordstam taxonomie ook kunnen dat de zoekterm zelf eerst vertaald wordt in een neutrale variant, alvorens deze gematcht wordt aan de vacatures in de database. Een andere manier om de zoek-input te koppelen aan relevante vacatures, is door het gebruik van taalmodellen die semantische nabijheid kunnen inschatten; met een dergelijk model zou de semantische nabijheid van begrippen met en zonder gender hoog kunnen zijn. Deze optie zou echter technisch gezien niets neutraliseren, maar zou in staat zijn om zowel de genderneutrale als gegenderde termen in de

resultaten te tonen. Wij kunnen echter niet met zekerheid zeggen welke van de bovengenoemde drie opties en/of andere opties gebruikt worden. Hoe de onderzochte zoekalgoritmen precies werken kunnen we niet met zekerheid vaststellen: het blijft een inschatting op basis van onderzochte output.

6.2.1 Het zoekalgoritme gaat genderstereotypen tegen

Vacaturesites A en B beschikken duidelijk over complexe zoekalgoritmen, die zouden kunnen werken met een woordstam taxonomie of met het score van semantische nabijheid. In de praktijk betekent dit dat zij ook zoekresultaten weergeven die niet één-op-één corresponderen met de opgegeven zoekterm. Vaak gaven zij genderneutrale zoekresultaten terug wanneer werd gezocht met een gendergekleurde zoekterm. Dit gebeurde echter niet andersom: wanneer werd gezocht met een genderneutrale zoekterm, werden geen gendergekleurde zoekresultaten getoond. Dit is interessant: het zou kunnen duiden op een semantische koppeling, of een taxonomie, in het algoritme die eenrichtingsverkeer is. Deze 'eenrichtingsverkeer-algoritmen' hebben de potentie om taal actief te neutraliseren en daarmee bestaande vooroordelen in de samenleving (pre-existing bias) tegen te gaan.

Van alle websites die onder de loep worden genomen, is het vacaturesite B die het beste lijkt te slagen in het neutraliseren van gendergekleurde termen. Wanneer er op deze website werd gezocht naar 'leraar' of 'lerares', faalde het zoekalgoritme er nimmer in om vacatures te laten zien waarin alleen een écht neutraal voornaamwoord zoals 'leerkracht' werd gebruikt. Als zodanig lijkt het zoekalgoritme van B het beste uitgerust om de naturalisatie van de gendergekleurde taal te bestrijden, evenals vooringenomenheid en genderstereotypen die in de samenleving circuleren tegen te gaan.

6.2.2 Het zoekalgoritme reproduceert genderstereotypen

We kunnen concluderen dat vacaturesite C hoogstwaarschijnlijk een simpel zoekalgoritme gebruikt dat slechts zoekresultaten laat zien die één-op-één de gebruikte zoekterm bevatten. Hoewel vacaturesite C er waarschijnlijk niet actief naar streeft, heeft de website wel de potentie om maatschappelijke vooroordelen te reproduceren. De website dient effectief als een direct doorgeefluik voor reeds bestaande vooroordelen: degenen die bijvoorbeeld geloven dat onderwijs een baan is voor mannen en dus zoeken met de term 'leraar' zullen alleen vacatures vinden waarin de vacatureomschrijving specifiek gebruik maakt van hetzelfde voornaamwoord. Vacatures met een genderneutrale functietitel worden niet weergegeven. Vooroordelen van werkzoekenden worden op deze manier herbevestigd – in dit geval dat lesgeven een taak is voor mannen, zonder een alternatief te bieden voor dergelijke ideeën. In deze hoedanigheid kan worden gesteld dat vacaturesite C bijdraagt aan de naturalisatie van gendergekleurde taal en daardoor genderstereotypen en vooroordelen over bepaalde beroepen reproduceert.

Hoewel dit probleem zeker het meest voorkomt bij de vacaturesite C, kan hetzelfde worden gezegd voor A en B; zoals vermeld, worden op deze websites de gendergekleurde zoektermen 'secretaresse', 'schoonmaakster' en 'timmerman' – beroepen die traditioneel zeer duidelijk gendergekleurd zijn – zelden gender-geneutraliseerd. Op die manier treedt hetzelfde effect als hierboven beschreven op bij zowel A als B.

6.2.3 Het zoekalgoritme versterkt genderstereotypen

Daarnaast zijn indicaties gevonden dat er in het geval van vacaturesite A sprake is van een versterking van genderstereotypen. Zoals eerder vermeld, gaf het zoekalgoritme van vacaturesite A altijd neutraal geformuleerde zoekresultaten terug wanneer gezocht werd met een gendergekleurde, vrouwelijke zoekterm – in sommige gevallen worden deze voornaamwoorden vertaald naar zogenaamd 'neutrale' voornaamwoorden die van oorsprong eigenlijk mannelijk zijn – terwijl mannelijke voornaamwoorden niet altijd worden teruggebracht tot een genderneutrale term. Effectief verbergt de zoekmachine daarmee vrouwelijke beroepsomschrijvingen zoals 'lerares' en behandelt (een aantal) mannelijke voornaamwoorden als neutraal. Daar komt nog bij dat, wanneer de zoekmachine wordt gevraagd om op basis van een vrouwelijk voornaamwoord zoals 'lerares' te zoeken naar vacatures, deze in staat is om met vacatures voor banen aan te komen die kunnen worden opgevat als 'lager' in de hiërarchie, zoals 'onderwijsassistent'. Dat gebeurde in onze steekproef niet wanneer werd gezocht met een mannelijke term als 'leraar'. Hieruit maken we op dat het zoekalgoritme van vacaturesite A de positie van mannen als 'de norm' in – in dit geval – de onderwijsarbeidsmarkt normaliseert, al dan niet onbewust en onbedoeld. Bij deze bevinding plaatsen we een grote kanttekening gezien de beperkte steekproef die we hebben gedaan: het ging namelijk slechts om één zoekresultaat dat een vacature weergaf voor 'onderwijsassistent'. Wel kan dit aanleiding geven om vervolgonderzoek te doen waarbij een grotere steekproef wordt genomen.

Er moet wel op worden gewezen dat, hoewel de hiervoor beschreven correlatie zeker aanwezig is, de zoekmachine van vacaturesite A wel degelijk ook bijdraagt aan het tegengaan/neutraliseren van gendergekleurde taal en genderstereotypen; immers, bij tal van gelegenheden vertaalt het mannelijke en vrouwelijke voornaamwoorden naar échte neutrale voornaamwoorden. Het zette bijvoorbeeld de termen 'lerares' en 'leraar' om in 'leerkracht'.

6.3 Overkoepelende conclusie en aanbevelingen

Dit onderzoek laat zien dat de manier waarop zoekalgoritmen op vacaturewebsites ontworpen zijn, bepalend is voor de manier waarop genderstereotypen al dan niet gereflecteerd worden in de zoekresultaten. Het zoekalgoritme heeft uiteraard invloed op de kansen die de werkzoekende krijgt, de werkzoekende krijgt immers pas de keuze om te reageren op een vacature, wanneer de vacature door het zoekalgoritme getoond wordt. Daarnaast weten we uit eerder onderzoek dat het gebruik van enkel het mannelijk voornaamwoord en mannelijke functietitels in vacatures nadelige gevolgen heeft voor iedereen die niet aan de mannelijke norm-positie voldoet. Om dat te voorkomen is het van belang dat het zoekalgoritme genderneutrale taal voorrang geeft boven gendergekleurde taal.

Kortom, de algoritmen die vacaturewebsites gebruiken om zoektermen te interpreteren, vervolgens vacatures te selecteren en te rangschikken kunnen vooringenomenheid in de vorm van genderstereotypen tegengaan, één-op-één reproduceren of versterken, afhankelijk van hun werking. Een simpel zoekalgoritme dat alleen letterlijk zoekt op de ingegeven zoekterm, reproduceert

genderstereotypen en geeft op die manier vooringenomenheid die al bestaat in de samenleving door. Een complexer algoritme dat bijvoorbeeld gebruik maakt van een woordstam taxonomie of semantische nabijheid scores is in staat om vooringenomenheid te neutraliseren, oftewel tegen te gaan. Daar staat tegenover dat in sommige gevallen een dergelijk algoritme ook genderstereotypen zou kunnen versterken, bijvoorbeeld in het geval waarin gezocht wordt op 'lerares' en het zoekalgoritme een vacature voor 'onderwijsassistent' toont, een functie voor een lager opleidingsniveau.

De drie onderzochte zoekalgoritmen hebben alledrie de fenomenen laten zien. Vacaturesite B slaagde er het beste in om gendergekleurde taal om te zetten in genderneutrale taal en gaat daarmee genderstereotypen tegen. Vacaturesite C slaagde hier het minste in. Alhoewel de ontwikkelaars van de algoritmen wellicht niet stil hebben gestaan van het effect van hun algoritme op genderstereotypen en daarmee op de kansen van werkzoekenden op de arbeidsmarkt, blijken de algoritmen wel degelijk invloed te hebben. Het is daarom van groot belang om hierbij stil te staan tijdens het ontwikkelproces van het algoritme, en eventueel testen uit te voeren om de invloed op genderstereotypen te identificeren en waar nodig te mitigeren. Daarbij kunnen we stellen dat een simpel zoekalgoritme dat de zoektermen slechts letterlijk herhaalt, in ieder geval niet zal voldoen.

7 Referenties

- Alessandro, Brian d', Cathy O'Neil, & Tom LaGatta. 2017. "Conscientious Classification: A Data Scientist's Guide to Discrimination-Aware Classification." *Big Data* 5 (2): 120–34. <https://doi.org/10.1089/big.2016.0048>.
- Alexander, Michelle. 2020. "Introduction." In *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*, 1–23. New York: The New Press. <http://ebookcentral.proquest.com/lib/uunl/detail.action?docID=5651869>.
- Blok, P.H., Hans Lammerant, & Paul De Hert. 2018. "Big Data Besluitvormingsprocessen En Sluipwegen van Discriminatie." *Nederlands Tijdschrift Voor de Mensenrechten* 43 (1).
- Bollegala, D., Matsuo, Y., & Ishizuka, M. (2007). Measuring semantic similarity between words using web search engines. *www*, 7(2007), 757-766.
- Bowker, Geoffrey C., & Susan Leigh Star. 1999. "Some Tricks of the Trade in Analyzing Classification." In *Sorting Things Out: Classification and Its Consequences*, 33–50. Cambridge, MA: MIT Press.
- Buolamwini, Joy, & Timnit Gebru. 2018. "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification." In *Proceedings of Machine Learning Research* 81, 81:1–15.
- Chatard-Pannetier, Armand, Serge Guimont, & Delphine Martinot. 2005. "Impact of the Lexical Feminization of Professions on Student Self-Efficacy: A Questioning of Male Universalism?" *Psychological Year* 105 (2): 249–72.
- College voor de Rechten van de Mens. 2020. "Als computers je CV beoordelen, wie beoordeelt dan de computers? Algoritmes en discriminatie bij werving en selectie". <https://mensenrechten.nl/nl/publicatie/5f4f81af1e0fec037359c49a>
- College voor de Rechten van de Mens. 2013. "De juiste persoon op de juiste plaats - De rol van stereotypering bij de toegang tot de arbeidsmarkt." <https://mensenrechten.nl/nl/publicatie/5b46fd03748c2212a54518b5>
- Dastin, Jeffrey. 2018. "Amazon Scraps Secret AI Recruiting Tool That Showed Bias against Women." *Reuters*, October 10, 2018. <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>.
- Day, Matt. 2016. "LinkedIn Changes Search Algorithm to Remove Female-to-Male Name Prompts." *The Seattle Times*. 2016. <https://www.seattletimes.com/business/microsoft/linkedin-changes-search-algorithm-to-remove-female-to-male-name-prompts/>.
- D'Ignazio, Catherine, & Lauren F. Klein. 2020. "1. The Power Chapter." In *Data Feminism*. The MIT Press. <https://data-feminism.mitpress.mit.edu/pub/vi8obxh7>.
- Dourish, Paul. 2016. "Algorithms and Their Others: Algorithmic Culture in Context." *Big Data & Society* 3 (2): 1–11.
- Dyer, Richard. 2002. "White." In *The Matter of Images: Essays on Representation*, 126–48. London: Routledge.
- Formanowicz, Magdalena, Sylwia Bedyńska, Aleksandra Cislak, Friederike Braun, & Sabine Sczesny. 2013.

- "Side Effects of Gender-Fair Language: How Feminine Job Titles Influence the Evaluation of Female Applicants." *European Journal of Social Psychology* 43 (February): 62–71. <https://doi.org/10.1002/ejsp.1924>.
- Foucault, Michel. 2007. "The Meshes of Power." In *Space, Knowledge and Power*, edited by Jeremy W. Crampton and Stuart Elden, translated by Gerald Moore, 153–62. Hampshire: Ashgate Publishing Limited.
- Friedman, Batya, & Helen Nissenbaum. 1996. "Bias in Computer Systems." *ACM Transactions on Information Systems* 14 (3): 330–47. <https://doi.org/10.1145/230538.230561>.
- Grommé, Francisca, Sophie Emmert, Noortje Wiezer, and Claartje Thijs. 2019. "Digitale Arbeidsmarktdiscriminatie: Inzicht in de Risico's Op Arbeidsmarktdiscriminatie Door de Inzet van Recruitment Technologieën in Werving En Selectie." TNO. <https://repository.tudelft.nl/islandora/object/uuid%3A4d78130c-5270-4c42-b321-6f1d8e380e00>.
- Hall, Stuart. 1992. "The West and the Rest: Discourse and Power." In *Formations of Modernity*, edited by Stuart Hall and Bram Gieben, 275–332. Cambridge: Polity Press.
- Inspectie SZW. 2021. "Eindrapportage discriminatiemogelijkheden online platforms." Publicatie. Ministerie van Sociale Zaken en Werkgelegenheid. <https://www.inspectieszw.nl/publicaties/publicaties/2021/01/22/eindrapportage-discriminatiemogelijkheden-online-platforms>.
- Morse, Jack. 2017. "Google Translate Might Have a Gender Problem." Mashable. 2017. <https://mashable.com/2017/11/30/google-translate-sexism/?europa=true>.
- Noble, Safiya Umoja. 2018a. "A Society, Searching." In *Algorithms of Oppression: How Search Engines Reinforce Racism*, 15–63. New York: New York University Press.
- . 2018b. "The Future of Information Culture." In *Algorithms of Oppression: How Search Engines Reinforce Racism*, 153–70. New York: New York University Press.
- O'Neil, Cathy. 2016. "Bomb Parts: What Is a Model?" In *Weapons of Math Destruction*, 20–33. New York: Crown.
- Peng, Andi, Besmira Nushi, Emre Kiciman, Kori Inkpen, Siddharth Suri, & Ece Kamar. 2019. "What You See Is What You Get? The Impact of Representation Criteria on Human Bias in Hiring." *ArXiv:1909.03567 [Cs]*, September. <http://arxiv.org/abs/1909.03567>.
- Perez, Caroline Criado. 2019. "Introduction: The Default Male." In *Invisible Women: Data Bias in a World Designed for Men* | Caroline Criado Perez | Download, 14–41. New York: Abrams Press. <https://b-ok.org/book/3711694/3b938d>.
- Pous, Irene. 2020. "Vanaf Nu Een 'Redactrice.'" OnzeTaal.Nl. 2020. <https://onzetaal.nl/nieuws-en-dossiers/weblog/vanaf-nu-een-redactrice>.
- Sánchez, D., & Moreno, A. (2004, December). Automatic Generation of Taxonomies from the WWW. In *International Conference on Practical Aspects of Knowledge Management* (pp. 208–219). Springer, Berlin, Heidelberg.
- Seaver, Nick. 2017. "Algorithms as Culture: Some Tactics for the Ethnography of Algorithmic Systems." *Big Data & Society* 4 (2): 205395171773810. <https://doi.org/10.1177/2053951717738104>.
- Stahlberg, Dagmar, & Sabine Sczesny. 2001. "Effects of the Generic Masculine and Alternative Forms of Language on the Mental Inclusion of Women." *Psychologische Rundschau* 52 (3): 131–40.

- Sturken, Marita, & Lisa Cartwright. 2009. "Images, Power, and Politics." In *Practices of Looking: An Introduction to Visual Culture*, 2nd ed., 9–48. Oxford University Press: Oxford.
- Tobin, Ariana, and Jeremy B. Merrill. 2018. "Facebook Is Letting Job Advertisers Target Only Men." ProPublica. 2018. <https://www.propublica.org/article/facebook-is-letting-job-advertisers-target-only-men?token=ZsVw5mrMeDYGjT1yYrG59hyUg0xqNSKY>.
- Vervecken, Dries. 2013. "The Impact of Gender Fair Language Use on Children's Gendered Occupational Beliefs and Listeners' Perceptions of Speakers." Doctoral Thesis Dissertation, Berlin: Freien Universität Berlin. https://refubium.fu-berlin.de/bitstream/handle/fub188/5473/Dries_Vervecken_Ph._D._The_Impact_of_Gender_Fair_Language_Use___For_UB.pdf?sequence=1.
- Zuiderveen Borgesius, Frederik. 2018. *Discrimination, Artificial Intelligence, and Algorithmic Decision-Making*. Strasbourg: Council of Europe.

8 Appendix

Twaalf profielen van werkzoekenden

Naam	Geslacht	Nationaliteit	Werkervaring	Leeftijd (in g-mail)*
Henk de Vries	M	NL	10 (irrelevant)	30
Johan de Ruiter	M	NL	10 (irrelevant)	30
Youssef Yilmaz	M	Buitenlands	10 (irrelevant)	30
Emir Nguyen	M	Buitenlands	10 (irrelevant)	30
Saartje de Vries	V	NL	10 (irrelevant)	30
Anniek van Dijk	V	NL	10 (irrelevant)	30
Amira Yilmaz	V	Buitenlands	10 (irrelevant)	30
Farah Nguyen	V	Buitenlands	10 (irrelevant)	30
Wilma Janssen	V	NL (irrelevant)	20	45
Gerda Pietersen	V	NL (irrelevant)	20	45
Petra de Jong	V	NL (irrelevant)	1	23
Nynke de Ruiter	V	NL (irrelevant)	1	23

Lijst met gendergekleurde functietitels en hun neutrale tegenhanger

Gendergekleurde functietitel	'Genderneutrale' tegenhanger
Actrice	Acteur
Brandweerman	Brandwacht
Gastheer	Gastvrouw/-heer
Gastvrouw	Gastvrouw/-heer
Kassière	Kassamedewerker
Leidster kinderopvang	Kinderopvangmedewerker
Leraar	Docent
Lerares	Docent
Politieaan	Politieagent
Schoonmaakster	Schoonmaakmedewerker
Secretaresse	Secretariaatsmedewerker
Stewardess	Steward(ess)
Timmerman	Houtbewerker
Verpleegster	Verpleegkundige
Verpleger	Verpleegkundige