kennis en aanpak van sociale vraagstukken

“TikTok's algorithms knew I was bi before I did”
Jess Joho

Kunstmatige intelligentie en lhbti+ emancipatie: een verkenning van kansen en risico’s, stakeholders en mogelijke interventies in vier maatschappelijke deelgebieden

Onderzoeksrapport

Colofon

kennis en aanpak van
sociale vraagstukken

Auteurs:
Mayke Kromhout, Laura Jak, Lily van den Broek en Lotte Thissen
M.m.v. Charlot Pierik & Joey Poerwoatmodjo

Eindredactie: Danny Nomden

Dit onderzoek is in opdracht van het ministerie van Onderwijs, Cultuur en Wetenschap.

December 2023
© Movisie

Movisie: kennis en aanpak van sociale vraagstukken
Movisie is hét landelijk kennisinstituut voor een samenhangende aanpak van sociale vraagstukken. Samen met de praktijk ontwikkelen we kennis over wat echt goed werkt en passen we die kennis toe. De unieke rol van Movisie is het versnellen van leerprocessen. We zijn alleen tevreden als we een duurzame positieve verandering voor mensen in een kwetsbare positie realiseren.
Kijk voor meer informatie op www.movisie.nl.

Inhoud

SAMENVATTING 5
1) INLEIDING EN ACHTERGROND 6
1.1 Onderzoeksvragen en doelstelling 7
1.2 Onderzoeksmethode en uitvoering 8
2) BEGRIPSOMSCHRIJVING 11
2.1 Kunstmatige intelligentie 11
2.2 Algoritmen 11
2.3 Machine learning 12
2.4 Deep learning 13
2.5 Discriminerende algoritmen? 13
2.6 Algoritmen en de lhbti+ gemeenschap: risico's en kansen 14
3) KUNSTMATIGE INTELLIGENTIE OP VIER MAATSCHAPPELIJKE DEELGEBIEDEN 15
3.1 Onderwijs 15
3.1.1 Feiten en cijfers over lhbti+ in onderwijs 15
3.1.2 Relevantie van kunstmatige intelligentie 16
3.1.3 Risico's 16
3.1.4 Kansen 17
3.2 Arbeidsmarkt 18
3.2.1 Feiten en cijfers over lhbti+ op de arbeidsmarkt 18
3.2.2 Relevantie van kunstmatige intelligentie 18
3.2.3 Risico's 19
3.2.4 Kansen 20
3.3 Gezondheidszorg 20
3.3.1 Feiten en cijfers over lhbti+ en gezondheidszorg 20
3.3.2 Relevantie van kunstmatige intelligentie 22
3.3.3 Risico's 22
3.3.4 Kansen 23
3.4 (Online) Veiligheid 24
3.4.1 Feiten en cijfers over lhbti+ en (online) veiligheid 24
3.4.2 Relevantie van kunstmatige intelligentie 25
3.4.3 Risico's 25
3.4.4 Kansen 26
4) EXPERTS OVER DE RISICO'S VAN KUNSTMATIGE INTELLIGENTIE 32
4.1 Algemene risico's 32
4.2 Onderwijs 33

4.3 Arbeidsmarkt 34
4.4 Gezondheidszorg 35
4.5 (Online) Veiligheid 36
5) EXPERTS OVER DE KANSEN VAN KUNSTMATIGE INTELLIGENTIE 38
5.1 Algemene kansen 38
5.2 Onderwijs 39
5.3 Arbeidsmarkt 40
5.4 Gezondheidszorg 41
5.5 (Online) Veiligheid 42
6) EXPERTS OVER STAKEHOLDERS 44
7) CONCLUSIES 50
VERANTWOORDING 55
LITERATUURLIJST 58

SAMENVATTING

Het doel van dit onderzoeksrapport is achterhalen op welke wijze kansen en risico's van kunstmatige intelligentie voor de lhbti+ (on)gelijkheid kunnen worden benut c.q. beperkt. Dit wordt in kaart gebracht voor de deelgebieden *onderwijs*, *arbeidsmarkt*, *gezondheidszorg* en *(online) veiligheid*. Het onderzoek is een eerste verkenning over hoe lhbti+ en kunstmatige intelligentie zich tot elkaar verhouden, en is daarom beperkt tot deze vier deelgebieden. Per deelgebied worden de belangrijkste verantwoordelijke actoren (stakeholders) in beeld gebracht, evenals aanknopingspunten voor beleid die kunnen bijdragen aan gelijke behandeling van lhbti+ personen in en door kunstmatige intelligentie. Om de onderzoeksvraag te beantwoorden, is een literatuurstudie verricht, een interview afgenomen en een expertmeeting georganiseerd met experts op het gebied van kunstmatige intelligentie en/of lhbti+ vraagstukken.

Geconcludeerd kan worden dat er nog maar weinig onderzoek naar en kennis is van de crosssectie tussen kunstmatige intelligentie en lhbti+ emancipatie. De experts onderstreepten dan ook de noodzaak van onderzoek en een multi-stakeholder aanpak, waarbij de betrokkenheid van de lhbti+ gemeenschappen als noodzakelijke randvoorwaarde werd gezien om tot de ontwikkeling van inclusieve kunstmatige intelligentie te komen. Experts spraken hun zorgen uit over de opslag, vertrouwelijkheid, transparantie, representativiteit, heteronormativiteit en de mensenrechtentoets van bestaande datasets. Middels bewustwording en regelgeving moet het mogelijk worden om meer grip op te krijgen op deze terechte zorgen.

Deze verkenning onderstreept het belang van het betrekken van stakeholders vanaf het begin én in samenspraak om tot beleid, innovaties en handelingsperspectieven te komen die inclusief zijn en niet (on)bedoeld lbhti+ personen uitsluiten. Belangrijke stakeholders zijn investeerders en bedrijven, ministeries, wetenschap en kennisinstituten, de nationale AI-coalitie en bovenal lhbti+ gemeenschappen en -belangenorganisaties. Daarbij is het van belang om continu oog te hebben voor de kwetsbare positie van lhbti+ personen bij het ontwerp en de uitvoering van kunstmatige intelligentiesystemen. Zeker gezien de kwetsbare positie van lhbti+ personen, is het belangrijk de waarde van kunstmatige intelligentie niet zomaar voor lief te nemen. Uit zowel de literatuurstudie als uit de gesprekken met de experts blijkt dat het belangrijk is om de inzet van kunstmatige intelligentie kritisch langs de meetlat te leggen, samen met alle betrokken partijen, om na te gaan hoe het kan bijdragen aan de inclusie en, idealiter, de emancipatie van lhbti+ personen, zonder onbedoelde effecten en schadeberokkening.

Disclaimer over terminologie

De terminologie om naar regenbooggemeenschappen te verwijzen is continu in beweging. Waar Movisie eerder lhbti+ gebruikte, zijn we nu overgestapt op lhbtiqa+. Echter, we hebben ook teruggekregen uit de gemeenschappen dat het onterecht is om letters te gebruiken als deze niet specifiek in het onderzoek zijn meegenomen. In deze verkenning hebben we q en a niet meegenomen en daarom houden we het in deze verkenning bij lhbti+.

1) INLEIDING EN ACHTERGROND

Big data, algoritmen en kunstmatige intelligentie worden steeds vaker ingezet voor een betere dienstverlening, efficiëntie, (online) veiligheid of om een dienst of product op de markt te brengen. Ze kunnen patronen in een berg data (cijfers, foto's, video, geolocaties et cetera) herkennen die met behulp van technologie zijn verzameld en die wij zelf niet zouden zien.

De overheid gebruikt bijvoorbeeld voorspellende algoritmen om fraude op te sporen. De politie gebruikt algoritmen om bijvoorbeeld een risico-inschatting te maken van crimineel gedrag. Waar vroeger een agent of leerkracht die inschatting maakte, is het nu een computermodel. Sociale mediakanalen zoals Facebook en TikTok houden onder andere in de gaten welke video's we liken of delen, waar we reacties achterlaten en wie we volgen. Vervolgens worden algoritmen ingezet om te bepalen wat er op de tijdlijn van de gebruikers verschijnt en welke gerichte advertenties zij te zien krijgen. Netflix houdt kijkgedrag bij en het algoritme rangschikt en beveelt vervolgens films aan die de kijker waarschijnlijk leuk vindt, door scores toe te kennen aan bijvoorbeeld hetzelfde genre of dezelfde acteur.

Ook op andere terreinen zien we de inzet van kunstmatige intelligentie. In de zorg wordt een vorm van kunstmatige intelligentie gebruikt voor voorspellingen over virusverspreiding, het herkennen van tumoren, ondersteuning bij diagnoses of bij het identificeren van mogelijk suïcidale patiënten. In het onderwijs kan kunstmatige intelligentie het onderwijsbeleid ondersteunen om geïnformeerde beslissingen te kunnen nemen. Denk hierbij aan gepersonaliseerde lessen, gepersonaliseerde studietrajectbegeleiding of het herkennen van risicoleerlingen met een verhoogde kans op uitval of slechte resultaten.

In zijn algemeenheid kan men zeggen dat een algoritme wordt gemaakt om met behulp van trainingsdata een bepaalde uitkomst te genereren, patronen te ontdekken die we zelf niet kunnen zien. Die uitkomst moet dan vallen binnen een bepaalde bandbreedte van de maker van het algoritme. Voorbeeld: een algoritme dat melanomen moet opsporen, beperkt zich tot de huid. In dit geval is het doel van het algoritme duidelijk alsook de in te voeren trainingsdata. Maar wat als de maker niet stilstaat bij de diversiteit aan huidskleuren en het algoritmen niet in staat is om melanomen in donkere huidskleuren te signaleren? In het sociaal domein kan het eveneens complex worden. Als de eigenaar van het algoritme trainingsdata heeft gebruikt waarbij belangrijke informatie (onbewust) mist of foutief is, kunnen de uitkomsten onvoorspelbaar en ongewenst zijn. Dit was bijvoorbeeld het geval toen automatische selectie van cv's op basis van trainingsdata leidde tot de selectie van alleen (witte) mannen. Ook is het mogelijk dat de ontwerper van het algoritme dusdanig scherpe criteria heeft aangebracht dat het algoritme weliswaar goed werkt, maar de uitkomsten zeer ongewenst zijn. Dat bleek bijvoorbeeld bij de geruchtmakende kinderopvangtoeslagenaffaire, waarbij de Belastingdienst algoritmen gebruikte voor controle op foutieve aanvragen met de kinderopvangtoeslag. Het model dat de Belastingdienst gebruikte, ging uit van meerdere risicofactoren, waaronder het niet hebben van de Nederlandse nationaliteit. Mensen zonder de Nederlandse nationaliteit hadden een grotere kans om als fraudeur te worden bestempeld. De Rekenkamer Rotterdam wijst dan ook op ethische risico's bij de inzet van algoritmen door de gemeente Rotterdam voor het opsporen van uitkeringsfraudeurs. De landelijke overheid moest zelfs stoppen met het gebruik van het computersysteem SyRI (Systeem

Risico Indicatie) omdat onduidelijk was hoe het algoritme exact werkte en welke risico-indicatoren werden gebruikt voor het in beeld krijgen van potentiële fraudeurs.

De inzet van recruitmenttechnologie kan leiden tot discriminatie[1] van vrouwen en andere kwetsbare (of minderheids-) groepen. Zo bleek de inzet van gezichtsherkenningssoftware risicovol: vrouwen met een donkere huidskleur worden minder makkelijk herkend dan vrouwen met een lichtere huidskleur. Dat kan allerlei problemen voor hen opleveren, zoals bij toegang tot de arbeidsmarkt, kredieten of tot andere maatschappelijke domeinen. Voor etnische minderheden in het algemeen kan het leiden tot etnische profilering.

Er is relatief weinig bekend over de impact van kunstmatige intelligentie op de lhbti+ gemeenschap. Wel wijst onderzoek uit dat er problemen zijn met gezichtsdetectie van transgender personen, met name tijdens de transitie. Daarbij blijken kunstmatige neurale netwerken de seksuele oriëntatie van personen met veel grotere accuraatheid te kunnen detecteren dan mensen, wat belangrijke risico's oplevert rondom privacy en (online) veiligheid van lhbti+ personen, bijvoorbeeld als het in gezichtsherkenningssoftware zou worden toegepast. De Leidse onderzoeker Fosch-Villaronga wijst daarnaast op ongewenste effecten van het gebruik van filters door digitale platformen als Twitter en Facebook zonder oog te hebben voor de sociale context.

> Een voorbeeld is het onderling geaccepteerd taalgebruik van drag queens. Binnen de gemeenschap van drag queens is de term 'bitch' bijvoorbeeld normaal en wordt plagerig en/of positief gebruikt. Het algoritme van Facebook en Twitter ziet dit woord echter als ordinair of beledigend ('toxic') taalgebruik en verwijdert de content. Dit leidt onbedoeld tot een inperking van de uitingsvrijheid voor deze groep. (Fosch-Villaronga, 2020)

Op dit moment is nog onduidelijk op welke maatschappelijke deelgebieden kunstmatige intelligentie de lhbti+ gemeenschap raakt (en in de toekomst gaat raken). Dit is aanleiding voor het ministerie van OCW om een verkenning te doen. Deze bestaat uit twee delen: een literatuurstudie en een focusgroep met experts op het gebied van kunstmatige intelligentie en lhbti+ emancipatie. Het doel van de literatuurstudie is om aanwijzingen te vinden voor kansen en risico's van kunstmatige intelligentie voor lhbti+ (on)gelijkheid op een aantal potentieel risicovolle terreinen, namelijk de terreinen *onderwijs*, *arbeidsmarkt*, *gezondheidszorg* en *(online) veiligheid*. Tijdens de daaropvolgende focusgroep met experts worden *best practices* besproken, om vervolgens vanuit inzichten van stakeholders en de literatuur aanbevelingen te doen voor de ontwikkeling van beleid en interventies om te komen tot inclusieve kunstmatige intelligentie.

1.1 Onderzoeksvragen en doelstelling

Voor het onderzoek is de volgende hoofdvraag geformuleerd:

Op welke wijze kunnen kansen en risico's van kunstmatige intelligentie voor lhbti+ (on)gelijkheid voor verschillende deelgebieden worden benut c.q. beperkt?

[1] Deze zijn: godsdienst, levensovertuiging, politieke gezondheid, ras, geslacht, nationaliteit, seksuele gerichtheid, burgerlijke staat, handicap of chronische ziekte, leeftijd, arbeidsduur en soort contract.

De deelvragen zijn:

1. Welke kansen biedt kunstmatige intelligentie om lhbti+ gelijkheid op diverse deelgebieden te benutten?
2. Welke risico's van kunstmatige intelligentie voor lhbti+ ongelijkheid zijn voor diverse deelgebieden te beperken?
3. Welke *best practices* zijn beschikbaar voor het omgaan met kansen en risico's van kunstmatige intelligentie voor lhbti+ (on)gelijkheid op verschillende deelgebieden?
4. Welke stakeholders zijn verantwoordelijk om te komen tot algoritmen die bijdragen aan lhbti+ gelijkheid op verschillende deelgebieden?
5. Welke aanknopingspunten voor kunstmatige intelligentie-beleid en -interventies kunnen worden geformuleerd voor het bevorderen van lhbti+ gelijkheid op verschillende deelgebieden?

Doel van het onderzoek

Het doel van deze verkenning is inzicht te verkrijgen in de kansen en risico's van kunstmatige intelligentie met betrekking tot genderidentiteit en seksuele oriëntatie. Het gaat dus om het verkennen van potentiële gebieden waar kunstmatige intelligentie een kans, danwel een risico vormt voor lhbti+ (on)gelijkheid. Dit doen we voor de deelgebieden *onderwijs*, *arbeidsmarkt*, *gezondheidszorg* en *(online) veiligheid*. Per deelgebied worden de belangrijkste verantwoordelijke actoren (stakeholders) in beeld gebracht, evenals aanknopingspunten voor AI-beleid (interventies) die kunnen bijdragen aan gelijke behandeling in en door algoritmen.

1.2 Onderzoeksmethode en uitvoering

Het onderzoek bestaat uit drie methoden van dataverzameling: literatuuronderzoek, een interview en een expertmeeting. Het onderzoeksteam kwam regelmatig bijeen om te reflecteren op en te beslissen over de selectie van de literatuur en de criteria voor inclusie, de ordening en analyse van de literatuur en voor de interpretatie van de gesprekken met experts en de expertmeeting.

Voor het literatuuronderzoek hebben we een aantal zoekstrategieën ingezet. Ten eerste hebben we zoektermen en search strings geformuleerd ten behoeve van de onderzoeksliteratuur. Ten tweede hebben we literatuur verzameld op aanbeveling van experts. Een aantal onderzoeken zijn vanuit de experts-interviews en de expertmeeting ingebracht op basis van criteria, zoektermen en search strings. Ten slotte hebben we de sneeuwbalmethode toegepast voor de gevonden literatuur. We zochten peer-reviewed literatuur, maar bestudeerden ook overige literatuur, zoals rapporten en werkdocumenten. De literatuurstudie betreft voornamelijk buitenlands (Noord-Amerikaans) onderzoek, dat zich grotendeels richt op de algemene impact van kunstmatige intelligentie. Slechts in een aantal gevallen ligt de nadruk in de literatuur rondom dit thema op de impact op lhbti+ personen.

Het interview vond plaats met één van de zeer weinige onderzoekers naar de impact van kunstmatige intelligentie op lhbti+ emancipatie in Nederland. Tijdens dit interview werd gefocust op risico's en kansen, en werd dieper ingegaan op manieren om tot inclusieve kunstmatige intelligentiesystemen te komen.

De focus tijdens de expertmeeting lag op dezelfde zaken als in het interview. Gaandeweg het

literatuuronderzoek werd duidelijk dat er, buiten de geïnterviewde wetenschapper, weinig expertise op de crosssectie van kunstmatige intelligentie en lhbti+ in Nederland (en het buitenland) bestaat. Daarom werd de basis voor deze expertmeeting het bijeenbrengen van deze werelden, het inspireren van elkaar en het delen van kennis en ervaringen, met als doel het in kaart brengen van alle belanghebbenden voor het bijdragen aan inclusieve kunstmatige intelligentie.

De opbouw van dit rapport is als volgt:
De hoofdstukken 2) en 3) (deel 1) hebben betrekking op de literatuurstudie. Deel 2, met hoofdstukken 4) tot en met 6), gaat in op resultaten van de expertmeeting. Deel 3 (hoofdstuk 7) bevat overkoepelende conclusies en aanbevelingen waarna de verantwoording en de literatuurlijst volgt.

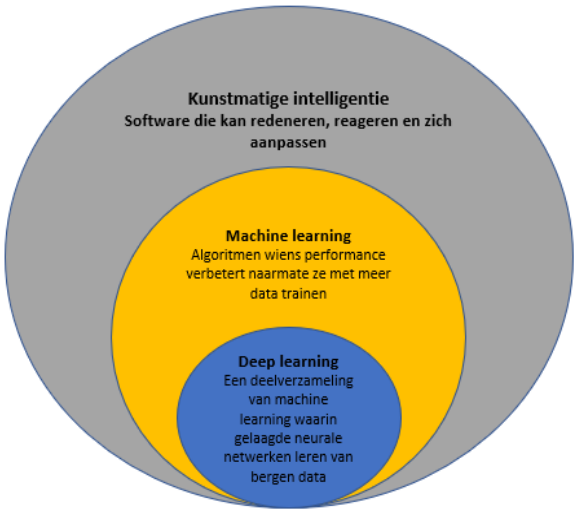
Deel 1

Literatuurverkenning

2) BEGRIPSOMSCHRIJVING

2.1 Kunstmatige intelligentie

Kunstmatige intelligentie is een verzamelnaam voor computersystemen die complexe taken uitvoeren die voorheen menselijke input vereisten. Een voorbeeld is de online communicatie met klanten of een potje schaken op en tegen de computer. Met kunstmatige intelligentie probeert men het denkvermogen van de mens na te maken: de computer wordt getraind om op basis van een berg data zelfstandig beslissingen te nemen én daarvan te leren. Hiervoor ontwikkelt men een stappenplan (algoritmen) om een rekenkundig of informaticaprobleem op te lossen. Het grootste voordeel van kunstmatige intelligentie is dat patronen in grote hoeveelheden data worden ontdekt. Ook patronen die wij als mensen nog niet eerder hebben ontdekt of kunnen ontdekken. In de zorg kan bijvoorbeeld vastgesteld worden welke patiënten het meeste risico lopen. Deze patiënten krijgen vervolgens prioriteit voor diagnostiek of behandeling.



Er is geen algemeen geldige definitie van kunstmatige intelligentie die consistent wordt gebruikt door alle belanghebbenden. In deze verkenning hanteren we de volgende omschrijving van de Europese Commissie: *'Systemen die intelligent gedrag vertonen door hun omgeving te analyseren en – met een zekere mate van zelfstandigheid – actie ondernemen om specifieke doelen te bereiken.*'

2.2 Algoritmen

Een algoritme is niets anders dan een stappenplan, een set regels of instructies die een computer uitvoert. Hiermee kunnen grote hoeveelheden data gecombineerd en geanalyseerd worden.

Voorbeeld: als je een serie van tien willekeurige cijfers op volgorde van klein naar groot wil leggen, dan kun je dat makkelijk handmatig doen. Bij 5.000 cijfers wordt het iets minder makkelijk. Dit kan dan door een computer gedaan worden, die geïnstrueerd moet worden. Die instructie bestaat uit een set regels, die in een bepaalde volgorde moeten worden doorlopen. De set regels noemt men algoritmen.

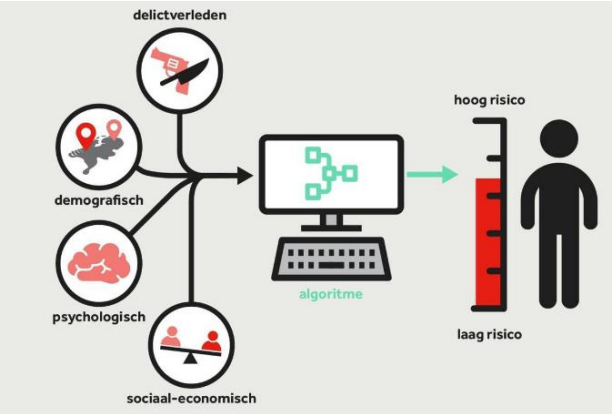
Wiskundige Hannah Fry beschrijft in haar boek Algoritmen aan de macht vier verschillende functies van een algoritme:

1. Prioriteren (bijvoorbeeld: Google beslist welke zoekresultaten of welke content goed te vinden is. Dat betekent dat Google gebruikmaakt van algoritmen die woorden als homo, lesbisch en

biseksueel verwijderen. Hierdoor zijn bijvoorbeeld video's van leden van de lhbti+ gemeenschap moeilijker vindbaar).

2. Classificeren (bijvoorbeeld: een medisch algoritme onderscheidt kankercellen van gewone cellen).
3. Associëren (bijvoorbeeld: OkCupid, BlackPeopleMeet, Grindr en andere datingapps voor lhbti+ personen bevelen een geschikte partner aan op basis van bepaalde kenmerken. Mainstream datingapps zoals Tinder maken het steeds meer mogelijk iemands seksuele oriëntatie te selecteren of bouwen een veiligheidsfunctie in voor lhbti+ gebruikers wanneer ze naar landen reizen waar homo zijn nog steeds illegaal is).
4. Filteren (bijvoorbeeld: met Alexa is het tegenwoordig mogelijk om informatie over lhbti+ personen op te vragen. Deze filtermogelijkheid kan op andere gebieden van pas komen, bijvoorbeeld bij het beter beoordelen van claims van asielzoekers dat ze lid zijn van de lhbti+ gemeenschap).

Een algoritme combineert vaak meerdere van deze vier functies. Het eerder besproken politiealgoritme zoekt naar associaties tussen bepaalde kenmerken en crimineel gedrag. Vervolgens classificeert het mensen op basis daarvan als hoog of laag risico.



Bron: NOS, 2019

2.3 Machine learning

De term kunstmatige intelligentie wordt vaak door elkaar gebruikt met onderliggende gerelateerde termen, zoals *machine learning* en *deep learning*. *Machine learning* is gericht op het bouwen van systemen die kunnen leren of hun prestaties kunnen verbeteren op basis van de data en instructies waarmee de systemen worden gevoed.

De applicatie SimpledCard is een voorbeeld van machine learning, waarbij internationaal geaccepteerde betaalpassen uitgegeven kunnen worden zonder tussenkomst van een bank. De applicatie monitort alle transacties en leert zelf frauduleuze transacties herkennen. Hiermee wordt tijd bespaard die benodigd zou zijn voor handmatige controle van transacties en bonnetjes.

Een voorbeeld uit de zorg is het trainen van enorme hoeveelheden radiologische foto's om borstkanker te detecteren. Hiervoor worden algoritmen van Google ingezet. Vele (honderd)duizenden beelden worden gelabeld aangeboden aan de computer, totdat deze een algoritme heeft gevormd waarop de computer zelfstandig nieuw aangeboden beelden als zodanig kan herkennen.

2.4 Deep learning

Deep learning is een discipline van machine learning. Hierin leert een zogenaamd neuraal netwerk van grote hoeveelheden data. Het kan het beste worden omschreven als een techniek waarbij op basis van grote hoeveelheden data patronen kunnen worden herkend. Het netwerk bouwt een model van een stukje werkelijkheid, zoals menselijke gezichten, en in dat model zitten 'gewichten'. Dat zijn knoppen waar het netwerk in de trainingsfase zelf aan draait om de herkenning te optimaliseren. De eerste doorbraak van deep learning was op het gebied van spraaktechnologie. Het wordt omgezet in een andere vorm van taal, zonder te begrijpen wat het betekent. Denk bijvoorbeeld aan de mogelijkheid die bestaat om mondelinge instructies aan onze auto te kunnen geven. Of dat we opstaan met de stem van een digitale spraakassistent die ons goedemorgen wenst en het weer voorspelt.

2.5 Discriminerende algoritmen?

In principe zit er geen kwade geest achter algoritmen. Algoritmen hebben geen emoties. Ze doen wat hen gezegd wordt en ze voorspellen op basis van data. En daar zit de crux: de instructies worden door mensen gemaakt en de data wordt door mensen verzameld. Kunstmatige intelligentie-technologieën zijn afhankelijk van grote datasets en de instructies die hun makers ze geeft. Als die datasets en instructies (onbedoeld) beïnvloed worden door vooroordelen uit de praktijk en institutioneel beleid en normen, dan zullen de algoritmische modellen die ongelijkheden reproduceren, en in sommige gevallen zelfs versterken.
Het versterken van ongelijkheden kan bijvoorbeeld gebeuren wanneer algoritmen worden gebruikt door de politie om criminaliteit te registeren. Als politieagenten veel meer surveilleren in een buurt waar veel bewoners met een migratieachtergrond wonen, registreren ze ook meer criminaliteit, terwijl er in werkelijkheid misschien evenveel criminaliteit is als gemiddeld. Er kan dan door het algoritme een (foutief) verband worden opgepikt tussen het hebben van een migratieachtergrond en crimineel gedrag. Mensen die in een bepaalde buurt wonen worden dan in systemen als 'risicovol' gelabeld zonder dat ze het weten. Dit heet 'algorithmic bias', ofwel vooringenomen algoritmen (Zuiderveen Borgesius, 2018).

Verschillende organisaties, zoals het College voor de Rechten van de Mens (CRM) en het Bureau van de Europese Unie voor de grondrechten (FRA), maken zich steeds meer zorgen over eventuele schendingen van grondrechten bij kunstmatige intelligentie. Hierbij gaat het onder meer, maar niet uitsluitend, over privacy, gegevensbescherming, non-discriminatie en toegang tot de rechter (Bureau van de Europese Unie voor de Grondrechten, 2021). Zij vinden het belangrijk dat men juiste trainingsdata gebruikt zodat bias (vooringenomenheid) van de kunstmatige intelligentie kan worden voorkomen. De *European Data Protection Board* (EDPB) en de *European Data Protection Service* (EDPS) verwoorden dit als volgt:

'*Remote biometric identification of individuals in publicly accessible spaces poses a high-risk of intrusion into individuals' private lives, with severe effects on the populations' expectation of being anonymous in public spaces. For these reasons, the EDPB and the EDPS call for a general ban on any use of AI for an automated recognition of human features in publicly accessible spaces - such as of faces but also of gait, fingerprints, DNA, voice, keystrokes and other biometric or behavioural signals - in any context*.' (EDPB-EDPS, 18 juni 2021)

2.6 Algoritmen en de lhbti+ gemeenschap: risico's en kansen

In feite kunnen we bijna niet om kunstmatige intelligentie heen. Het wordt voor uiteenlopende toepassingen gebruikt, zoals gezichtsherkenning, werving en selectie, medische diagnoses, fraudebestrijding, digitale platformen et cetera. Afhankelijk van het doel biedt kunstmatige intelligentie ondersteuning bij het prioriteren, classificeren, associëren of filteren van data. Zogezegd kan kunstmatige intelligentie ons helpen op het gebied van gemak, gebruikerservaring en efficiëntie. Maar het biedt geen oplossing voor alle vraagstukken.

Auteurs wijzen op het gegeven dat menselijk ingrijpen nodig blijft, bijvoorbeeld om kunstmatige systemen te trainen en draaiend te houden. Verder is menselijk ingrijpen belangrijk in het controleren, schiften en labelen van informatie, het bepalen van randvoorwaarden en het inschatten van de impact op de sociale omgeving. Daar waar blind vertrouwd wordt op kunstmatige intelligentie, rijzen vragen op over de ethische kanten. Er zijn steeds meer aanwijzingen dat het gebruik van kunstmatige intelligentie kwetsbare groepen in de samenleving ernstig kan benadelen, zelfs wanneer er geen expliciete bedoeling is om te discrimineren (zoals we zagen met de kinderopvangtoeslagenaffaire).

Een complicerende factor is het feit dat termen in het sociale domein niet onwrikbaar zijn gedefinieerd, maar onderhevig zijn aan soms intense discussie. Nu zal het duidelijk zijn dat een gewenste uitkomst als 'is dit vlekje wel of geen melanoom' op basis van honderdduizenden foto's trainingsdata, hoewel arbeidsintensief en mits gebaseerd op een diversiteit aan foto's, klip en klaar is. Bij trainingsdata uit het sociale domein is dit veel lastiger. Welke data gebruikt men hiervoor? Zijn de data actueel? Welke criteria gebruikt men voor de gewenste uitkomst? Men kan zich voorstellen dat data van tien jaar geleden de huidige lhbti+ gemeenschap niet bepaald correct weergeven. Bijvoorbeeld, de termen 'genderexpressie' en 'genderidentiteit' zullen in deze data niet rijkelijk en correct voor handen zijn.

Maar ook actuele data zijn niet per definitie integer en zullen moeten worden beoordeeld op hun bruikbaarheid voor het leerproces van kunstmatige intelligentie. Alleen een binaire indeling in gender is uiteraard niet voldoende. Geautomatiseerde genderherkenningssoftware neemt het concept 'geslacht' bijvoorbeeld als uitgangspunt zonder rekening te houden met de binaire aard van dit concept en zonder andere relevante concepten in overweging te nemen. Ten gevolge hiervan wordt de software getraind aan de hand van data die genderstereotype kenmerken bevatten (denk aan lichaamsbewegingen, fysiologische en gedragskenmerken, gelaatsuitdrukkingen en taalgebruik).

Daarnaast is het belangrijk na te denken over het doel van het algoritme. Soms wordt vergeten om de 'waarom' vraag te stellen. In de regel is het verstandig om niet uit te gaan van de gedachte: 'we hebben data, we moeten hier iets mee', maar van een duidelijk en zinvol doel. Blind vertrouwen op algoritmen brengt risico's met zich mee. Bij een positieve inzet biedt kunstmatige intelligentie belangrijke kansen. Het doel moet duidelijk zijn, correct geformuleerd, conform privacy afwegingen en het zinvolle karakter ervan.

In het volgende hoofdstuk presenteren we de resultaten van de literatuurstudie. We hebben ons gericht op aanwijzingen voor risico's en kansen van de inzet van kunstmatige intelligentie voor lhbti+ personen. We bakenen de verkenning af tot vier maatschappelijke deelgebieden: het *onderwijs*, de *arbeidsmarkt*, de *gezondheidszorg* en de *(online) veiligheid*. De verwachting was dat we voldoende aanwijzingen zouden vinden over de intersectie van kunstmatige intelligentie met lhbti+ emancipatie.

3) KUNSTMATIGE INTELLIGENTIE OP VIER MAATSCHAPPELIJKE DEELGEBIEDEN

3.1 Onderwijs

3.1.1 Feiten en cijfers over lhbti+ in onderwijs

Lhbti+ leerlingen staan voor extra uitdagingen in het onderwijs, met name in het praktische onderwijs, op het (v)mbo en op religieuze scholen. De meest recente lhbt monitor van het SCP (Huijnk et al., 2022) maakt duidelijk dat lhbti+ leerlingen kwetsbaar zijn: ze hebben een groter risico om gepest te worden en ook een groter risico om slachtoffer te worden van geweld. Hoewel het veiligheidsbeleid op veel scholen is verbeterd, laat het onverlet dat tussen de 12 en 15 procent van leerlingen aangeeft liever niet bevriend te zijn met lhbti+ medeleerlingen. Dit percentage was het hoogste wanneer het over vriendschappen met transgender meisjes gaat (15 procent). Onder andere deze cijfers tekenen de extra kwetsbare posities van lhbti+ leerlingen in het algemeen en transgender jongeren in het bijzonder.

Het percentage leerlingen dat wordt gepest om hun seksuele oriëntatie, is de afgelopen jaren toegenomen. Die houding komt niet alleen van medeleerlingen, maar ook van leerkrachten en ondersteunend personeel**.** Transgender jongeren van middelbare schoolleeftijd in Nederland ervaren veel stigmatisering (Keuzenkamp & Kuyper, 2013). Waar van alle leerlingen in het v(s)o 10 procent aangaf één keer per maand of vaker gepest te worden op school, gaf 26 procent van de transgender leerlingen dit aan (Scholte, Nelen, De Wit & Kroes, 2016). Eén op de tien cisgender jongeren gaf aan voor hun zestiende (heel) vaak te worden gepest. Bij transgender leerlingen is dit één op de vijf (Van Beusekom & Kuyper, 2018). Transgender leerlingen maken ook significant vaker verbaal geweld, materieel geweld, sociaal geweld, lichamelijk geweld en seksueel geweld mee dan hun cisgender leeftijdsgenoten (Scholte et al., 2016)**.** Lhb scholieren zijn minder positief over hun leven dan heteroseksuele scholieren.

Uit onderzoek (Huijnk et al., 2022) blijkt dat lhbti+ leerlingen een relatief negatief zelfbeeld hebben. Zij hebben een lagere mate van zelfwaardering dan heteroseksuele jongeren. Pesten (door medeleerlingen, leerkrachten en ondersteunend personeel) blijkt de grootste voorspeller te zijn van zelfmoordgedrag onder lhb jongeren. Suïcidepogingen komen onder hen 4,5 keer vaker voor dan onder andere jongeren (Broekroelofs, 2021). Lhb scholieren voelen zich ook minder gelukkig. Heteroseksuele scholieren geven gemiddeld een 7,7 aan hun leven, terwijl lhb scholieren hun leven een 6,8 geven. In schooltermen uitgedrukt geven heteroseksuele scholieren dus bijna een ‘goed’ aan hun leven, terwijl lhb scholieren hun leven beoordelen met bijna een ruim voldoende. Een groot deel van de lhb scholieren, bijna één op de drie (32 procent), geeft aan zich minstens wekelijks ongelukkig te voelen (Huijnk et al., 2022).

Het televisieprogramma Pointer onthulde eerder (2020) dat het toelatingsbeleid van scholen voor bijzonder onderwijs ten aanzien van lhbti+ leerlingen en leerkrachten te wensen overlaat. Een op de

vijf reformatorische scholen spreekt openlijk hun afkeuring uit over seksuele relaties tussen twee mensen van hetzelfde geslacht.

3.1.2 Relevantie van kunstmatige intelligentie

In het onderwijs komen steeds vaker toepassingen van kunstmatige intelligentie voor. Het wordt bijvoorbeeld gebruikt voor het overbrengen en toetsen van kennis, en bij het informeren van docenten over prestaties van leerlingen. De inzet van kunstmatige intelligentie kan docenten op verschillende manieren ondersteunen. Denk hierbij bijvoorbeeld aan gepersonaliseerde lessen, gepersonaliseerde studietrajectbegeleiding of het herkennen van risicoleerlingen met een verhoogde kans op uitval of slechte resultaten (Van der Vorst, Jelicic, De Vries, & Albers, 2019). Hiervoor worden data uit een leerlingvolgsysteem gekoppeld aan persoonlijkheidskenmerken zoals leeftijd, geslacht (jongen-meisje), thuistaalgebruik, sociaaleconomische situatie van het gezin van de leerling et cetera. Het computersysteem biedt ook de mogelijkheid voor het noteren van docentobservaties ten aanzien van bijvoorbeeld de schoolveiligheid, oudergesprekken en absentie van leerlingen.

Een leerlingvolgsysteem is een computersysteem waarmee de school de resultaten van leerlingen bijhoudt.

In de literatuur kwamen we maar één voorbeeld tegen van risico's en/of kansen voor lhbti+ leerlingen vanwege de toepassingen van kunstmatige intelligentie. Dit – overigens buitenlands – voorbeeld gaat over de beperkte toegang van leerlingen tot internet inzake lhbti+ thema's (zie verder onder risico 2).

Op basis van de verzamelde literatuur en de ontwikkelingen die we in *3.1.1 Feiten en cijfers* bespreken, schetsen we enkele risico's en kansen. Waar mogelijk, werken we die uit voor lhbti+ leerlingen.

3.1.3 Risico's

Risico 1: beperkte toegang tot onderwijs door gebruik algoritmen bij toelating. Steeds meer onderwijs-instellingen gebruiken *machine learning*-algoritmen om studenten toe te laten of te weigeren. Dit brengt (ten minste) de volgende twee problemen met zich mee:

1) Er is geen transparantie in de wijze waarop dat proces van toelaten en afwijzen tot stand komt. Het algoritme wijst op basis van de voorkeuren van leerlingen een school toe. Het blijft voor de buitenwacht onduidelijk met welke criteria het algoritme de toe- of afwijzing voor de voorkeurschool bepaalt.
2) De onduidelijkheid kan mogelijk leiden tot benadeling van bepaalde kansarme groepen. Uit een studie in België komt bijvoorbeeld naar voren dat het huidige algoritme dat moet zorgen voor een betere spreiding van kansarme leerlingen over de scholen, pas beter kan presteren als de capaciteitsproblemen op de diverse scholen tegelijkertijd worden aangepakt.

Uit de literatuur komt niet naar voren of de 'toetscriteria' van het algoritme achter het aanmeldsysteem nadelig uitpakken voor lhbti+ personen. Het risico is daarbij dat dit algoritme eventueel wel discriminerend werkt, maar dat dit door gebrek aan die kennis niet wordt aangepakt. Wel wijst UNESCO erop dat voorzichtigheid geboden is bij het gebruiken van machine learning-algoritmen die zijn getraind op een bepaalde dataset (bijvoorbeeld met studenten uit een West-Europees land): dan is het resultaat mogelijk niet direct toepasbaar op studenten uit andere delen van de wereld. De

trainingsdataset kan vooringenomen zijn ten opzichte van een bepaalde groep en kan daarom oneerlijk discrimineren bij gebruik voor een andere groep (UNESCO, 2019). Vooringenomenheid komt ook voor in het geautomatiseerd beoordelen van de tekstkwaliteit van schrijfproducten of de schrijfvaardigheid van leerlingen (Bridgeman et al., 2012) en zelfs het opsporen van emotie bij studenten (Ocumpaugh et al., 2014). In deze studie is de lhbti+ achtergrond van de leerlingen niet meegenomen. Uit ander onderzoek komt naar voren dat diverse achtergrondkenmerken van leerlingen, zoals de sociaaleconomische status, van invloed kunnen zijn op de schrijftalenten van leerlingen. Er zijn nog geen studies bekend waarin een link gelegd wordt tussen de prestaties op schrijfvaardigheid en lhbti+ achtergrond.

Risico 2: Beperkte toegang tot informatie over lhbti+ door gebruik filters op schoolcomputers. In Nederland zijn er nog geen voorbeelden van, maar onderzoek in het buitenland laat zien dat schoolcomputers lhbti+ websites als porno classificeren (Richardson et al., 2002; Stapleton, 2013). Hierdoor wordt de content van deze websites niet getoond, terwijl deze websites een belangrijke functie kunnen hebben voor lhbti+ jongeren. Zo helpen ze met het verspreiden van informatie over het bespreekbaar maken van seksuele diversiteit, de stereotype beeldvorming van lhbti+ personen, en bieden de mogelijkheid om ervaringen uit te wisselen over de lhbti+ identiteit, waar dat op de computer thuis minder makkelijker kan zijn of oncomfortabel kan voelen.

Risico 3: Ontoereikende mogelijkheden voor het ondersteunen van leerlingen. Er zijn tal van studies die bijvoorbeeld aantonen dat een onveilig schoolklimaat impact heeft op de onderwijsprestaties van lhbti+ leerlingen of studenten (Aragon et al., 2012; Kosciw et. Al, 2015). In de data die het leerling-volgsysteem van leerlingen bijhoudt is er geen mogelijkheid tot differentiatie naar seksuele oriëntatie, alleen tot het onderscheid jongen-meisje. Dat betekent dus dat lhbti+ leerlingen onzichtbaar blijven en niet goed ondersteund kunnen worden met behulp van de kunstmatige intelligentie zoals die momenteel vormgegeven is. Tegelijkertijd bestaat het risico dat, wanneer er wel een aantekening gemaakt wordt van seksuele of genderoriëntatie van een leerling, deze informatie ook in handen kan komen van docenten die minder open-minded zijn. Bovendien kan het bijhouden van gender in deze binaire indeling ervoor zorgen dat een leerling die non-binair of in transitie is, steeds wordt aangesproken op een manier die voor hen pijnlijk of respectloos kan zijn.

3.1.4 Kansen

Kans 1: Ondersteunen van onderwijsprofessionals. De inzet van kunstmatige intelligentie kan professionals in het onderwijs helpen met het overnemen van een deel van hun administratieve taken. Daardoor krijgen ze meer tijd en ruimte voor taken die lastig of niet over te nemen zijn door kunstmatige intelligentie, zoals meer sociale taken als begeleiding en coaching (Van der Vorst, Jelicic, De Vries, & Albers, 2019). Bij geïndividualiseerd lesgeven en autonoom leren kan kunstmatige intelligentie helpen bij het beter analyseren van studiegegevens. Hierdoor kunnen docenten gepersonaliseerde leerplannen maken voor hun studenten. De tijdwinst kan onder meer worden benut voor het bieden van extra ondersteuning aan lhbti+ leerlingen.

Kans 2: Interactie met menselijk ingrijpen. Kunstmatige intelligentie, mits goed ingezet, kan onderwijs-professionals helpen bij het ondersteunen van lhbti+ leerlingen. Maar niet alles kan worden opgelost met behulp van kunstmatige intelligentie: het zou hierbij vooral om een signaleringsfunctie kunnen gaan, die gevolgd wordt door menselijk handelen. Onderwijsprofessionals kunnen gebruikmaken van de mogelijkheid om in het automatische leerlingvolgsysteem handmatig de ondersteuningsbehoeften

van hun lhbti+ leerlingen bij te houden. Daarnaast kunnen onderwijsprofessionals gebruikmaken van bestaande infrastructuur van het GSA-netwerk op scholen om hun ondersteuning aan lhbti+ leerlingen te optimaliseren. Dit vergt dus ook menselijk ingrijpen.

Het bovenstaande laat zien dat er met name op het gebied van ondersteuning van lhbti+ leerlingen mogelijkheden liggen voor de inzet of verdere toepassing van kunstmatige intelligentie.

3.2 Arbeidsmarkt

3.2.1 Feiten en cijfers over lhbti+ op de arbeidsmarkt

Ook op de arbeidsmarkt kunnen lhbti+ personen met extra uitdagingen geconfronteerd worden. In het algemeen geldt dat lhbti+ personen minder makkelijk een baan vinden en houden. Transgender personen en bi+ personen hebben nog een lastigere positie op de arbeidsmarkt. Zij zijn minder vaak in vaste dienst, vaker arbeidsongeschikt en hebben vaker een uitkering dan de rest van de beroepsbevolking, blijkt uit de lhbt monitor van het SCP (Huijnk et al., 2022). Meer dan de helft van de transgender personen zit in de laagste inkomenscategorie.

Maar liefst 36 procent van de lhbti+ personen ervaart discriminatie op de werkvloer. Van heteroseksuele personen wordt vaker geaccepteerd dat zij over hun privéleven vertellen dan lhbti+ personen. Het verschil in ervaren discriminatie zit vooral in de bejegening door externen met wie lhbti+ personen door hun werk in contact komen, zoals klanten, patiënten of opdrachtgevers (Huijnk et al., 2022). 16 tot 17 procent van de homoseksuele en lesbische, 50 procent van de biseksuele en 60 procent van de transgender geneeskundestudenten meldt dat ze niet openlijk konden spreken over hun seksuele oriëntatie of genderidentiteit in hun professionele omgeving. De meest genoemde redenen hiervoor waren angst voor discriminatie, bezorgdheid over carrièremogelijkheden en de mening dat hun identiteit 'niemand iets aangaat' (Fallin-Bennett, 2015).

De verschillen in loopbaanbeleving, sociale omgang, burn-outklachten, verzuim en werktevredenheid tussen lhbti+ en heteroseksuele werknemers op de meeste terreinen zijn talrijk. De grootste verschillen doen zich voor tussen biseksuele en heteroseksuele werknemers. Biseksuele werknemers hebben vaker een tijdelijk contract, ervaren meer conflicten en ongewenst gedrag, hebben meer burn-outklachten en een hoger verzuim. Ook zijn ze minder tevreden. Dit komt overeen met eerder onderzoek. Bij lesbische en homoseksuele werknemers doen de problemen zich met name voor op het terrein van burn-out klachten, maar ook zij ervaren meer conflicten, meer ongewenst gedrag en zijn minder tevreden over hun werk.

In de literatuur kwamen we geen voorbeelden tegen van risico's en/of kansen van kunstmatige intelligentie op de arbeidsmarkt specifiek voor lhbti+ personen. Op basis van de verzamelde literatuur en de ontwikkelingen die we in *3.2.1 Feiten en cijfers* bespraken, schetsen we enkele risico's en kansen. Waar mogelijk, werken we die uit voor lhbti+ personen.

3.2.2 Relevantie van kunstmatige intelligentie

De roep om een lhbti+ inclusieve werkvloer klinkt steeds harder door. Critici wijzen op diverse mechanismen die werkgevers hiervoor kunnen inzetten. WOMEN Inc. pleit bijvoorbeeld voor de inzet van recruitmentsoftware die vrijgesteld is van vooroordelen tegen vrouwen. Daarnaast wijst Women

Inc. op transparantie en nut van de selectiecriteria van wervings- en selectiebureaus. Een ander mechanisme dat zou kunnen werken volgens critici is dat werkgevers al aan de voorkant van het wervings- en selectieproces uitstralen dat lhbti+ werknemers waardevol en welkom zijn bij het bedrijf. Er zijn geen voorbeelden bekend van hoe kunstmatige intelligentie dit proces kan ondersteunen. Wat betreft de werving zouden recruiters zich veel meer moeten gaan bezighouden met doelgroep-targeting, waardoor lhbti+ personen doelgericht worden gevraagd om te solliciteren. Hier is ook niet van bekend hoe precies kunstmatige intelligentie voor zou kunnen worden inzet.
Daarnaast zou gebruik kunnen worden gemaakt van manieren om lhbti+ personen die openlijk gender divers zijn, te laten participeren in bedrijfsvideo's. Een goede ontwikkeling is dat het ministerie van SZW algoritmen inzet om alle vormen van discriminatie in vacatureteksten op te sporen. De uitkomsten moeten input worden voor de bestrijding van discriminatie op de arbeidsmarkt. Er zijn zelfs landen die het gebruik van wervingssoftware verbieden.

3.2.3 Risico's

Risico 1: (on)bedoelde impact op de positie van lhbti+ personen op de arbeidsmarkt. Acemoglu en Restrepo (2019) onderzochten de invloed van kunstmatige intelligentie op de toekomst van werk. Zij stellen dat kunstmatige intelligentie op de verkeerde manier als vervanger van arbeid wordt ingezet en dat verdergaande automatisering van bedrijfsprocessen consequenties kan hebben voor de toch al zwakke arbeidsmarktpositie van kwetsbare groepen in het algemeen. Het gaat ook om werknemers die hun banen zien verdwijnen als gevolg van automatisering. Voor lhbti+ personen zou het bijvoorbeeld betekenen dat hun kansen op werk nog kleiner worden omdat ze relatief gezien een minder gunstige arbeidsmarktpositie hebben.

Risico 2: kunstmatige intelligentie werkt nog niet vlekkeloos in het wervings- en selectieproces. Omdat er ook bij de keuze voor kandidaten sprake kan zijn van (onbewuste) uitsluiting door recruiters, kan het algoritme die eerdere bias meenemen en overnemen, en kunnen kwetsbare groepen, waaronder lhbti+ personen, onbedoeld worden uitgesloten. Dit gebeurt door het trainen van algoritmen met cv's van gekozen kandidaten, waardoor de computer steeds betere kandidaten gaat herkennen. Wanneer de originele bias van de recruiter in de trainingsdata zit, wordt dit door de computer overgenomen en is de uitkomst welke kandidaat beter is dus ook (deels) gebaseerd op bias.

Risico 3: mogelijke discriminatie in geautomatiseerde online vacatures. Dalenberg (2018) onderzocht welke vereisten er zijn om te voldoen aan de Europese non-discriminatiewetgeving om discriminatie in geautomatiseerde online vacatures te voorkomen. Een voorbeeld is *hypertargeting,* dat zoveel betekent als het communiceren met een zeer specifieke doelgroep. Stel: een bedrijf is op zoek naar een junior accountmanager. Junior betekent in dit geval dat de ideale kandidaat zijn of haar studie commerciële economie bijna heeft afgerond of één tot twee jaar werkervaring heeft bij een commercieel bedrijf. Ook woont de ideale kandidaat binnen een straal van maximaal dertig kilometer van het bedrijf. Aanvullend gaat het bedrijf gegevens verzamelen over deze doelgroep, over hun interesses, de kanalen waarop zij zich bevinden, wat zij belangrijk vinden in hun zoektocht naar een (nieuwe) werkgever en hoe ze aangesproken willen worden. Deze data worden verwerkt in het aanbod op vacaturesites of in vacaturevideo's. Deze wervingssoftware (ook geavanceerde chatrobots, video-interviewsystemen en online assessmenttools) kan tot digitale arbeidsdiscriminatie leiden. Het onderzoek van Dalenberg toont bijvoorbeeld aan dat werkgevers – in (video)interviews – nog te weinig laten merken dat ze inclusief zijn. Ook zouden ze niet of te weinig in kanalen zoeken naar kandidaten waar lhbti+ personen zich begeven. Het SZW (zie pag. 32) vacaturescanner laten ontwikkelen.

Vooralsnog zoekt de scanner naar sectoren die in hun online vacatureteksten verboden onderscheid maken op basis van leeftijd, sekse of afkomst.

Risico 4: mogelijke nadelen van het inzetten van video-interviewing. De inzet van video-interviewing tijdens het wervings-en selectieproces wordt steeds vaker gemeengoed. Vanuit het oogpunt van kosten efficiëntie en tegengaan van onderscheid is dat begrijpelijk. Echter is dit niet zonder (ethische) risico. Het algoritme achter dit proces screent niet alleen op de inhoud van de antwoorden van sollicitanten, maar ook op aspecten van hun presentatie. Voorbeelden zijn de woordkeuze, lichaamstaal, gezichtsexpressie en -herkenning. Sommige software speurt ook internet af voor het samenstellen van een profiel van de kandidaat. In feite is onbekend in hoeverre de trainingsdata representatief zijn voor lhbti+ personen. Een van de risico's van video-interviewing tijdens het wervings- en selectieproces is dus dat het nadelig kan uitpakken voor lhbti+ personen.

3.2.4 Kansen

Kans 1: betere inrichting van het wervings- en selectiesysteem zodat ze minder of niet discrimineren. Kunstmatige intelligentie kan een gamechanger zijn voor recruiters voor wie diversiteit en inclusie op de eerste plaats komt. Zo kan een methode van objectief werven en selecteren ervoor zorgen dat aanwijzingen over seksuele oriëntatie van de kandidaat worden weggefilterd, zodat werkgevers of recruiters geen (onbewust) vooroordeel kunnen meenemen in hun keuze. Hoewel deze aanpak al bekend is, werkt het nog niet geheel vlekkeloos. Niet alle organisaties werken namelijk met een e-assessment, waardoor recruiters toch nog vertrouwen op hun eigen inschatting van de juiste kandidaat. Daarnaast kan er in het wervingsproces ook juist actief op zoek worden gegaan naar lhbti+ kandidaten, door bijvoorbeeld met software te werken die getraind is op het herkennen en selecteren van geschikte lhbti+ kandidaten.

Om ervoor te zorgen dat diverse kandidaten (inclusief lhbti+ personen) interesse tonen in een vacature, is het van belang de website van de werkgever uitnodigender te maken met behulp van het plaatsen van foto’s en video’s, het gebruik van inclusieve taal in vacatureteksten of andere content op de website (zoals blogs). Hiermee wordt een signaal afgegeven dat de werkgever zich inzet voor diversiteit en inclusie. Dit kan een positief effect hebben op de arbeidsmarktpositie van lhbti+ personen. De Europese non-discriminatierichtlijnen die directe en indirecte discriminatie op het gebied van arbeid op grond van ras of etnische afkomst, geslacht, seksuele oriëntatie, religieuze overtuiging, leeftijd en handicap verbieden, moeten hierbij als wettelijk kader gebruikt worden (Dalenberg, 2018).

3.3 Gezondheidszorg

3.3.1 Feiten en cijfers over lhbti+ en gezondheidszorg

In dit onderdeel schetsen we de situatie van lhbti+ personen in binnen de gezondheidszorg. Lesbische en homoseksuele personen voldoen minder vaak aan de normen voor gezond bewegen, ze hebben vaker weleens harddrugs gebruikt, zijn psychisch minder gezond, hebben vaker een depressie gehad en hebben een slecht ervaren gezondheid of een langdurige aandoening/ziekte. Daarnaast hebben lesbische en homoseksuele personen ook vaker contact met een huisarts. Uit onderzoek blijkt echter dat er in de zorg weinig kennis aanwezig is of wordt toepast over aandoeningen bij lhbti+ personen en de behandeling hiervan. Hierdoor worden diagnoses soms gemist en behandelingen (te) laat gestart. In de spreekkamer is de hetero cis-man nog steeds het

uitgangspunt. Lesbische, homoseksuele en bi+ mensen krijgen regelmatig te maken met aannamen en stereotypen. Daardoor krijgen én vragen ze niet altijd de juiste zorg. Dat geldt ook voor transgender personen, en in het bijzonder voor transgender personen van kleur (Kromhout & Wolf, 2019).

Ook andere specifieke groepen binnen de lhbti+ gemeenschap, zoals jongeren en ouderen, zijn extra kwetsbaar in de gezondheidszorg. Zo blijkt uit onderzoek dat ouderen zich soms genoodzaakt voelen weer de kast in te gaan: in de reguliere ouderenzorg zijn lhbti+ personen vaak niet op hun plek. Roos Pijpers (2020) wijst erop dat lhbti+ ouderen die zorg aan huis ontvangen van teams van verzorgenden en verpleegkundigen, het soms lastig vinden om iedere keer opnieuw in te moeten schatten of je open kunt zijn over wie je bent. Verder blijkt dat zelfstandig wonende roze ouderen die kwetsbaarder worden meer afstand nemen van de lhbti+ gemeenschap in hun stad of regio.
Daarnaast schetst onderzoek (Movisie, 2021) de zorgelijke situatie van dak- en thuisloze lhbti+ jongeren. Dak- of thuisloos zijn is op zichzelf al enorm zwaar voor jongeren. Voor lesbische, homoseksuele, biseksuele, transgender of intersekse jongeren (lhbti+) die soms worstelen met zelfacceptatie en afwijzing door familie of omgeving komt dit er bovenop. Vervolgens zijn er weinig veilige plekken in de opvang en komen pesten, geweld en misbruik vaker voor. Dit maakt hen drievoudig kwetsbaar. Zowel voor jeugdhulpverleners als lhbti+ belangenorganisaties lijken deze jongeren nog een blinde vlek. Lhbti+ jongeren lopen een hoger risico op zelfmoord dan andere jongeren: suïcidepogingen binnen deze groep komen 4,5 keer zo vaak voor (Movisie, 2021).

Risico's op mentale problemen en suïcide bij lhbti+ personen worden veelvuldig genoemd door onderzoekers (McDermott, 2015; McDermott & Katrina, 2012). Opvallend is dat onderzoekers aan de hand van informatie die lhbti+ personen op (semi-) anonieme sociale media plaatsen inzicht konden krijgen in kenmerken als geografische spreiding en de aard van de gezondheidsklachten. Dit is een belangrijk gegeven om rekening mee te houden en ons te realiseren dat het internet een plek is waar veel lhbti+ personen 'uit de kast komen'. Daar gaan ze op zoek naar gelijkgestemden en hulp. De hulp kunnen zij niet (altijd) offline vinden vanwege vooroordelen en een gebrek aan kennis over lhbti+ gerelateerde thema's. In een ander onderzoek is er met behulp van de technologie achter *machine learning* gekeken naar de mentale gezondheid van lhbti+ pesonen op basis van posts op sociale media. Daar werd gekeken naar kenmerken van minderheidsstress-ervaringen. De resultaten kunnen bijdragen aan gezondheidsbeleid en het verbeteren van de kennis met betrekking tot verschillen in mentale gezondheid van lhbti+ personen. Ook kan het onderzoek een basis bieden voor het ontwikkelen van online tools die zijn afgestemd op de behoeften van lhbti+ personen (Saha, 2019).

In de literatuur kwamen we twee voorbeelden tegen van risico's en/of kansen voor lhbti+ personen vanwege de toepassingen van kunstmatige intelligentie. Het eerste voorbeeld betreft het kunnen voorspellen van suïcidaal gedrag bij jongeren met behulp van kunstmatige intelligentie. Het tweede voorbeeld betreft het ontwerp van veel kunstmatige intelligentiesystemen dat te weinig rekening zou houden met de genderdiversiteit van de patiëntpopulatie.

Op basis van de verzamelde literatuur en de ontwikkelingen die we in *3.3.1 Feiten en cijfers over lhbti+ en gezondheidszorg* bespreken, schetsen we enkele risico's en kansen. Waar mogelijk, werken we die uit voor lhbti+ personen.

3.3.2 Relevantie van kunstmatige intelligentie

Kunstmatige intelligentie biedt veel handelingsperspectief voor professionals in de zorg (VWS, 2021). Het gaat daarbij niet alleen om de financiële effecten voor direct betrokkenen, maar om alle mogelijke effecten van een maatregel die invloed hebben op de welvaart voor álle partijen in de samenleving, zoals effecten op de volksgezondheid en het milieu. In rapporten van de Wereldgezondheids-organisatie en TNO benadrukken onderzoekers ook deze meerwaarde.

De gezondheidszorg is een van de belangrijkste domeinen waar kunstmatige intelligentie haar meerwaarde heeft laten zien, onder meer in de farmacie en biomedische wetenschappen. Met behulp van kunstmatige intelligentie is het mogelijk om medicijnen sneller en goedkoper te ontwerpen, bijwerkingen van medicijncombinaties te voorspellen, nauwkeurige diagnoses te stellen, behandel-plannen voor patiënten te vereenvoudigen en de zorgkwaliteit te verbeteren. Daartoe beschikt men over relevante demografische patiëntkenmerken van groepen en hun ervaringen. Het is wel belangrijk dat men zich realiseert dat in het ontwerp van de meerderheid van algoritmen geen rekening is gehouden met de geslachts- en genderdimensie en de bijdrage daarvan aan gezondheids- en ziekteverschillen tussen individuen. Data zijn dus niet representatief voor specifieke sociale groepen, zoals lhbti+ personen, laat staan voor personen van kleur. Deze sociale groepen zijn namelijk ondervertegenwoordigd in gezondheidsdata. Op dit moment is er dus geen goed inzicht in de diversiteit en complexiteit aan zorgvragen binnen de lhbti+ populatie (Saha, 2012).

3.3.3 Risico's

Risico 1: gebrek aan transparantie en uitleg van medische beslissingen. Gebrek aan transparantie en uitleg van medische beslissingen levert problemen op voor de acceptatie, regulering en implementatie van (te vaak onverklaarbare) kunstmatige intelligentie in de gezondheidszorg (Bélisle-Pipon et al., 2021). Miller (2021) wijst erop dat het inbrengen van technologie nooit neutraal is. Het succes van kunstmatige intelligentie in de gezondheidszorg en de integriteit en reputatie van zorgprofessionals hangt af van de afstemming tussen de medische ethiek en het vermogen om te reageren op de uitdagingen die kunstmatige intelligentie met zich meebrengt voor het zorgsysteem (Reddy et al., 2019). Het wordt daarom breed erkend dat kunstmatige intelligentie aanzienlijke voordelen zal hebben voor de gezondheidszorg (geoptimaliseerd proces, hogere kwaliteit, lagere kosten en uitgebreide toegang) die ten koste zullen gaan van ethische kwesties die specifiek zijn voor de technologie (Abràmoff, Tobey, & Char, 2020). Zo kunnen kunstmatige intelligentie-technieken schadelijk zijn voor individuen, omdat er niet voldoende controle is op de beoordelingen die worden gemaakt door geautomatiseerde besluitvormingsprocessen. De belangrijkste juridische en ethische kwesties zijn discriminatie, vooroordelen en een gebrek aan transparantie. Mensen hebben geen toegang tot hun 'profielen' (Gupta, 2021).

Risico 2: beperkte beschikbaarheid van kwalitatief goede gegevens. In bestaand onderzoek naar de gezondheid van lhbti+ personen is vaak sprake van *oversampling*, waardoor er geen goed inzicht is in de diversiteit en complexiteit aan zorgvragen (Saha, 2019). Lhbti+ personen van kleur zijn bijvoorbeeld ondervertegenwoordigd in gezondheidsenquetes. Dit is een belangrijk aandachtspunt als het gaat over gelijkheid en gelijke toegang tot zorg. Het is daarom belangrijk om aandacht te hebben voor de heterogeniteit van patiënt en ziekte, rekening houdend met het feit dat door kunstmatige intelligentie gedetecteerde patronen grotendeels afgeleid worden uit kleinere historische datasets (Dzobo et al., 2020; Bélisle-Pipon et al., 2021). In de literatuur wordt verder onderscheid gemaakt

tussen *confirmation bias,* waarbij medische professionals meer waarde hechten aan bewijs dat hun veronderstelde diagnose ondersteunt en bewijs negeren dat iets anders zegt, en *automation bias*, waarbij medische professionals de begeleiding van een geautomatiseerd systeem accepteren en niet verder zoeken naar bewijs dat hun eigen idee bevestigt. In het laatste geval vertrouwen medische professionals blind op kunstmatige intelligentie (Challen et al., 2019).

Risico 3: bias en discriminatie in het ontwerp en bij de inzet van kunstmatige intelligentie. De COVID-19 pandemie illustreert het onevenredige effect van gezondheidsdreigingen op achtergestelde gemeenschappen. Hierbij spelen onder meer factoren als systemisch racisme, marginalisering en structurele ongelijkheid. Kunstmatige intelligentie wordt steeds meer ingezet in de zorg en bij het bestrijden van besmettelijke ziekten. Het is bekend dat kunstmatige intelligentie vatbaar is voor algoritmische vooroordelen die bestaande ongelijkheid kunnen verankeren en vergroten. Het kritiekloos inzetten van kunstmatige intelligentie in bijvoorbeeld de strijd tegen COVID-19 dreigt dus de nadelige effecten van de pandemie op kwetsbare groepen te versterken, waardoor de ongelijkheid op gezondheidsgebied nog groter wordt. Het gebruik van bevooroordeelde kunstmatige intelligentie-modellen kan onevenredig schadelijk zijn voor kwetsbare groepen die niet goed vertegenwoordigd zijn in trainingsdatasets (Fenech & Buston, 2020; Leslie et al., 2021).

3.3.4 Kansen

De inzet van kunstmatige intelligentie in de gezondheidszorg biedt aanzienlijke kansen om de zorg en resultaten voor patiënten te verbeteren, kosten te verlagen en de gezondheid van de bevolking te beïnvloeden. Kunstmatige intelligentie zal waarschijnlijk een belangrijke en aanvullende rol spelen voor de menselijke cognitie om de levering van gepersonaliseerde gezondheidszorg te ondersteunen (Matheny et al., 2019).

Kans 1: ruimte voor gepersonaliseerde adviezen. Bij goede zorgverlening gaat het steeds meer over het inspelen op individuele wensen en mogelijkheden van de persoon. Daarnaast is de context van de patiënt zeer relevant bij het bepalen van de behandeling of bij het afzien van een bepaalde behandeling. In de gezondheidszorg neemt de complexiteit van de zorgvraag toe; patiënten hebben vaak meerdere aandoeningen tegelijkertijd. Naast medische factoren spelen ook sociale en psychische factoren een belangrijke rol. De inzet van kunstmatige intelligentie biedt mogelijkheden om deze grote hoeveelheid informatie te verwerken en om te zetten in meer gepersonaliseerde adviezen (Mesko, 2017; Raad voor Volksgezondheid & Samenleving, 2019).

Kans 2: filteren van data uit digitale platforms. Onderzoekers ontwikkelden een geautomatiseerd classificatiesysteem om lhbti+ personen te identificeren op Twitter, met een nauwkeurigheid van 88 procent. Dit kan bijvoorbeeld gebruikt worden voor onderzoek naar de gezondheidssituatie van lhbti+ personen (Karami et al., 2021). Sociale media kunnen een nuttige epidemiologische bron zijn voor het verzamelen van gegevens over lhbti+ personen en het analyseren van de gegevens met behulp van machine learning. Hiermee kan online inhoud geanalyseerd worden. Dergelijke grote tekstuele datasets kunnen kwalitatief onderzoek ondersteunen. Dit kan nuttig zijn bij gezondheidsonderzoek naar lhbti+ personen. Het is moeilijk om via andere methodes aan dergelijke informatie te komen, ook vanwege bestaande stigma's in onderzoek en privacyregelgeving. Het gebruik van machine learning heeft potentie voor geautomatiseerde identificatie van personen met een verhoogd risico op suïcide (Meiselwitz, 2019), depressieve gevoelens (Li et al., 2020), minderheidsstressoren (Saha et al., 2019), negatieve emoties (Huang et al., 2019) en signalen van geestelijke gezondheid (Zhao et al., 2019)

onder de lhbti+ gemeenschap (Anasua Kundu, 2021). Naast dat dit kansen en mogelijkheden biedt, brengt deze ontwikkeling ook risico's met zich mee, bijvoorbeeld met betrekking tot privacy.

3.4 (Online) Veiligheid

3.4.1 Feiten en cijfers over lhbti+ en (online) veiligheid

De veiligheidsbeleving van lhbti+ personen manifesteert zich op een aantal deelterreinen: op school (zie ook 3.1.1), de werkvloer (zie 3.2.1), de openbare ruimte, de private sfeer en online. Op al deze deelterreinen is de veiligheidsbeleving van lhbti+ personen slecht (Van Beusekom en Kuyper 2018; CBS en WODC 2020; ResearchNed 2022). Als we specifiek kijken naar de veiligheid op scholen, zien we dat personeel met een lhbti+ achtergrond zich vaker onveilig voelt dan heteroseksueel cisgender personeel (ResearchNed 2022).

Uit de veiligheidsmonitor die in het primair en voortgezet onderwijs wordt uitgevoerd, blijkt dat lhbti+ leerlingen een groter risico lopen om gepest te worden, en ook om slachtoffer te worden van een geweldsincident (ResearchNed 2022). Ondanks dat het gevoel van veiligheid onder leerlingen in het voortgezet onderwijs over het algemeen niet veranderd is tussen 2018 en 2021, is deze wel verslechterd onder leerlingen die homoseksueel, lesbisch, biseksueel of transgender zijn. De mate waarin lhbti+ leerlingen worden uitgescholden of expres gestoord, is in 2021 niet verbeterd ten opzichte van 2018. Lhbti+ leerlingen zijn in 2021 vaker opzettelijk lastiggevallen door anderen dan in 2018.

Verder blijkt dat lhbti+ personen in vergelijking met heteropersonen naar verhouding vaak te maken hebben met huiselijk geweld. Deze doelgroep is echter niet goed in beeld bij hulpverleners. Lhbti+ personen voelen zich geremd om in de openbaar uiting te geven aan hun gevoelens. Bijna de helft (45 procent) van de Nederlandse transgender respondenten gaf in internationaal vergelijkend onderzoek aan dat zij nooit of zelden open zijn over hun genderidentiteit. Bij jongere transgender personen was dit zelfs meer dan de helft (EU Agency for Fundamental Rights 2020). Ruim de helft van de Nederlandse lhbti+ personen gaf aan dat zij bijna nooit of nooit hand in hand lopen met hun partner. En dit gaat alleen nog maar over bepaalde zichtbare vormen van intimiteit.

Voor veel lhbti+ personen is internet een plek geworden voor het ontmoeten van gelijkgestemden en het onderling uitwisselen van informatie. Dit wordt mogelijk gemaakt door specifieke eigenschappen van het internet, zoals het grote bereik van online berichten en de (gevoelsmatige) anonimiteit. Uit onderzoek dat het CBS op 10 januari publiceerde, blijkt dat 11,4 procent van de homo- of biseksuele internetgebruikers tussen de 12 en 25 jaar in 2018 te maken kreeg met online pesten (online haatzaaien). Bij heteroseksuele internetgebruikers in dezelfde leeftijdscategorie is dat 5 procent. Het gaat bijvoorbeeld om roddelen, pesten, stalken of dreigen met geweld, vaak met een seksuele (bij)bedoeling, op bijvoorbeeld WhatsApp en Instagram.

In de literatuur kwamen we twee voorbeelden tegen van risico's en/of kansen voor lhbti+ personen vanwege de toepassingen van kunstmatige intelligentie. Het eerste voorbeeld behandelden we reeds in 3.3.1, over het kunnen filteren/classificeren van gezondheidsproblemen bij online lhbti+ gebruikers. Het tweede voorbeeld betreft het filter dat door Twitter en Facebook voor specifiek taalgebruik waardoor drag queens zich niet meer kunnen uiten zoals ze dat elkaar gewend zijn.

Op basis van de verzamelde literatuur en de ontwikkelingen die we in *3.4.1 Feiten en cijfers* bespreken, schetsen we enkele risico's en kansen. Waar mogelijk, werken we die uit voor lhbti+ personen.

3.4.2 Relevantie van kunstmatige intelligentie

De problemen op het gebied van veiligheidsbeleving van lhbti+ personen zijn legio. Uit de literatuur komen voorbeelden en mogelijkheden naar voren op het gebied van online ondersteuning van lhbti+ personen. Voor lhbti+ personen kan online verbinding maken en delen een essentiële manier zijn om met elkaar om te gaan en oplossingen te vinden voor problemen die vrienden of familie misschien niet begrijpen. Er zijn echter ook risicogebieden voor lhbti+ personen bij online interactie. Op basis van locatiegegevens van een sociale media-app hebben onderzoekers gekeken naar de lhbti+ gemeenschap in Beijing. De onderzoekers stellen dat lhbti+ gerelateerde onderwerpen taboe zijn in China. Met hun onderzoek willen zij de publieke bewustwording vergroten over het bestaan, de omvang en de spreiding van de lhbti+ gemeenschap in Beijing. De onderzoekers brengen gebieden in kaart waar veel lhbti+ personen wonen, waaronder veel studenten. Ook brengen ze in kaart op welke tijdstippen lhbti+ personen zich op welke plekken bevinden en benoemen ze een hotspot van 'gay activiteiten'.

Online haatzaaien is een fenomeen dat leidt tot schadelijke psychologische effecten, en de aanwezigheid ervan op sociale media neemt toe. Transgender personen zijn het meest vatbaar voor online haatspraak en omdat de effecten in hun geval destructiever zijn, zou deze groep een centraal punt van zorg moeten zijn. Daarom moeten nieuwe onderzoeksinitiatieven zich richten op de specifieke manier waarop transgender personen reageren op en omgaan met online hate speech, samen met de bestrijdingsmaatregelen (Ștefăniță & Buf, 2021).

3.4.3 Risico's

Risico 1: inbreuk op grondrechten*.* Als de overheid algoritmen inzet in het veiligheidsdomein of deze gebruikt bij het differentiëren tussen groepen personen in het domein van sociale zekerheid, kan dit leiden tot strijdigheid met het recht op gelijke behandeling. Ook de rol van de overheid bij het reguleren van meningsuiting die plaatsvindt via een private, sterk technologie-gedreven infrastructuur, leidt tot nieuwe grondrechtelijke knelpunten, met name doordat de mogelijkheid van indirecte publieke censuur opdoemt. Aan de inzet van algoritmen bij het inperken van demonstratie- en verenigings-vrijheid zijn vergelijkbare risico's verbonden (Vetzo, Gerards, & Nehmelman, 2018). De onderzoekers benoemen specifiek dat zich bijzonder urgente problemen kunnen voordoen bij de uitoefening van rechten voor gelijke behandeling.

Risico 2: kwetsbaarheid*.* Het ministerie van Justitie en Veiligheid liet een inventarisatie doen naar mogelijkheden om het beoordelingsproces van asielaanvragen met een lhbti+ motief te verbeteren. Daarbij is ook gekeken naar de mogelijkheid om kunstmatige intelligentie in te zetten (Schans & Van Lierop, 2019). Onderzoek naar de geloofwaardigheidsbeoordeling van asielzaken met een lhbti+ motief laat zien dat het niet mogelijk en niet wenselijk is gebruik te maken van modellen en checklists. Dit omdat de persoonlijke identiteit uniek is en modellen die uitgaan van gedeelde ervaringen vaak voortkomen uit westerse ideeën over seksualiteit (Schans & Van Lierop, 2019).

Risico 3: delen van gevoelige informatie op mogelijk onveilige platforms en afhankelijkheid daarvan. Sociale media als Facebook en Tumblr kunnen volgens onderzoek een veilige plek zijn voor lhbti+ jongeren. Zij komen daar om te leren, participeren, communiceren met anderen en het vormen van hun identiteit en het uiten en ontdekken van hun seksualiteit en gender (Lucero, 2017). Deze behoefte aan specifieke online plekken voor lhbti+ personen onderschrijft daarnaast de kwetsbaarheid van lhbti+ personen en gemeenschappen in 'reguliere' digitale en gecommercialiseerde plekken (Cavalcante, 2018). Uit beide onderzoeken volgt een behoefte en noodzaak voor veilige online plekken voor lhbti+ personen om te leren, te ontdekken, elkaar te ontmoeten en uit te wisselen. Hiermee worden lhbti+ personen mogelijk bovengemiddeld afhankelijk van dergelijke plekken en de bedrijven die deze websites beheren. Daarnaast delen mensen persoonlijke en mogelijk gevoelige informatie online, in de veronderstelling dat het daar veilig is.

3.4.4 Kansen

Kans 1: benutten van online accounts van lhbti+ personen voor dataverzameling. In tegenstelling tot niet-lhbti+ personen hebben lhbti+ personen vaker accounts op sociale media en zijn ze dagelijks actief op sociale media (Karami, 2021; Tesfai 2019). Deze mediakanalen zijn een plek geworden waar lhbti+ personen informatie met elkaar delen en online zetten. Dit biedt interessante mogelijkheden voor onderzoek omdat men hiermee issues als privacy en stigma die in traditioneel onderzoek spelen en een belemmering vormen voor lhbti+ personen om informatie te delen, zou kunnen oplossen.

Kans 2: mogelijkheid tot online ondersteuning. Er is waardevol onderzoek uitgevoerd naar de lhbti+ gemeenschap op sociale media. Hierbij gaat het bijvoorbeeld over het definiëren van online identiteit, het verkennen van de maatschappelijke perceptie over de lhbti+ gemeenschap en het onderzoeken van het gebruik van social media door transgender adolescenten voor onderlinge steun. Ook het online gedrag van lhbti+ ouders is geanalyseerd, als het gaat om belangenbehartiging, het delen van ervaringen, het identificeren van gezondheidsproblemen van lhbti+ gebruikers en de correlatie tussen psychologisch welzijn en afhankelijkheid van sociale media. Verder benoemt het onderzoek de evaluatie van de lhbti+ ziekenhuiszorg op basis van berichten op sociale media en de gezondheids- en sociale behoeften van transgender personen. Kortom, sociale media bieden lhbti+ gebruikers een mogelijkheid om met privacy- en stigmaproblemen om te gaan.

Kans 3: beter inzicht in maatschappelijke houding ten aanzien van lhbti+ personen en thema's. Via kunstmatige intelligentie kunnen onderzoekers een beter beeld krijgen van de lhbti+ gemeenschap (Karami, 2021). Zo hebben onderzoekers met *machine learning* technieken gekeken naar de houding van online gebruikers ten aanzien van lhbti+ personen. Hierbij is met name gekeken naar het Twittergedrag en taalgebruik van gebruikers om een beeld te krijgen van hun houding. Er is een voorspellingsmodel ontwikkeld om die houdingen te onderzoeken en te identificeren. Hierbij vinden de onderzoekers dat woordanalyses een waardevol middel zijn om voorspellingen te kunnen doen (Bittner, Dettmar, Jaramillo, & Valta, 2020). Dergelijke technieken kunnen worden ingezet om de maatschappelijke houding te peilen ten aanzien van lhbti+ personen en lhbti+ gerelateerde thema's.

CONCLUSIE LITERATUURVERKENNING

Uit de literatuurverkenning komt naar voren dat kunstmatige intelligentie veelvuldig wordt ingezet voor het verhogen van de efficiëntie en het verbeteren van voorspellingen, dienstverlening en klantrelaties. Kunstmatige intelligentie omschrijven we als volgt: '*systemen die intelligent gedrag vertonen door hun omgeving te analyseren en - met een zekere mate van zelfstandigheid - actie ondernemen om specifieke doelen te bereiken'.*

De inzet van kunstmatige intelligentie dient vier doelen: prioriteren, classificeren, associëren en filteren van data waardoor we snel en goed geïnformeerd beslissingen kunnen nemen. Het helpt ons om besluiten te nemen door de verwerking van grote hoeveelheden informatie in een korte tijd. Een mogelijk doel is om maatschappelijke vraagstukken op te pakken of zoals in het geval van online platforms, een groter (doelgroep)bereik te realiseren. De essentie is het slimmer maken van software op basis van algoritmen. Die algoritmen (een set regels) worden gebruikt voor het herkennen van patronen (in data en datastructuren). Een belangrijke voorwaarde is dus dat het systeem zelfstandig moet kunnen leren. Kunstmatige intelligentie-systemen zijn zelflerend door het gebruik van neurale netwerken. Aan de hand van voorbeelden stellen zij trends en patronen in de data vast. De nauwkeurigheid neemt toe naarmate de machine meer ervaring en meer feedback krijgt.

Het doel van de literatuurverkenning was om te achterhalen welke mogelijke risico's en kansen van kunstmatige intelligentie voor lhbti+ personen aanwezig zijn op de terreinen *onderwijs*, *arbeidsmarkt*, *gezondheidszorg* en *(online) veiligheid*. De zoektocht leverde geringe informatie en aanwijzingen op over de kruising tussen kunstmatige intelligentie en lhbti+ personen. Ondanks dat kunstmatige intelligentie veelvuldig wordt ingezet op de door ons bestudeerde terreinen, lijkt dit erop te wijzen dat de inclusiviteit van de data waarmee kunstmatige intelligentiesystemen worden getraind te wensen over laat.

De weinige voorbeelden die we in de literatuur aantroffen, gaven ons wel enkele aanknopingspunten voor de inzet van kunstmatige intelligentie op de door ons bestudeerde terreinen. Hiervoor maakten we gebruik van cijfers over de actuele stand van zaken ten aanzien van lhbti+ personen en de toepassingsmogelijkheden van kunstmatige intelligentie. Daarbij lichtten we van elk terrein de risico's en kansen toe.

Onderwijs

Een belangrijk voordeel of een mogelijkheid van kunstmatige intelligentie in het onderwijs is dat het kan worden ingezet bij de ondersteuning van de docent, bij het beter laten aansluiten van het onderwijs bij de leerling, met zowel betere uitkomsten als een beter leerproces. De huidige toepassingen, bijvoorbeeld in de vorm van een leerlingvolgsysteem kunnen hiervoor worden benut. Het schoolsysteem maakt gebruik van een dergelijk systeem, onder meer om de prestaties van leerlingen te monitoren. Een belangrijk voordeel van dit volgsysteem is dat het inzicht geeft in prestatieverschillen tussen leerlingen, uitgesplitst naar hun achtergrondkenmerken. Seksuele oriëntatie van leerlingen wordt daar echter niet bij geregistreerd. Hoewel het vanuit een privacy-

oogpunt goed kan zijn om dergelijke informatie niet te registreren, zorgt het ook dat eventuele problemen van leerlingen die gerelateerd zijn aan hun seksuele oriëntatie minder makkelijk aan het licht komen. Daarbij zou het wel registreren een kans kunnen bieden om extra ondersteuning te bieden aan lhbti+ leerlingen. Uit de literatuurstudie komt naar voren dat lhbti+ leerlingen bijvoorbeeld behoefte hebben aan een juiste begeleidingsstructuur en een inclusief schoolklimaat. Scholen kunnen hieraan voldoen door een goede zorgstructuur te bieden. Een mogelijkheid daartoe is het benutten van de ruimte die het leerlingvolgsysteem biedt: leerkrachten kunnen handmatig informatie aanvullen in het kladblok. Dat geeft de mogelijkheid tot extra ondersteuning van lhbti+ leerlingen. Leerkrachten zouden ook het GSA-netwerk kunnen benutten voor het verzamelen van die informatie en het verder ondersteunen van lhbti+ leerlingen. Voor deze mogelijkheid om extra persoonlijke informatie over lhbti+ leerlingen in het 'leerling dossier' toe te voegen gelden dezelfde privacy regels als het leerlingendossier zelf.

Arbeidsmarkt

In het algemeen kan worden gezegd dat de nadruk van de toepassing van kunstmatige intelligentie veel meer ligt op tijdsbesparing en te weinig op het vinden van een divers samengesteld team. Women Inc. wees er bijvoorbeeld op dat tijdsbesparing ten koste kan gaan van kwaliteit als HR-managers zich te weinig verdiepen in mogelijke nadelen van de recruitment software die ze gebruiken voor bijvoorbeeld vrouwen.

Uit de voorbeelden in de literatuur leiden we af dat slimmer gebruik kan worden gemaakt van kunstmatige intelligentie bij het vinden van geschikte kandidaten. Voor de werving van lhbti+ personen noemt men diverse mogelijkheden, zoals de inzet van trainingsdata over geschikte lhbti+ kandidaten, het classificeren of filteren van persoonlijke data van kandidaten voor een 'eerlijke' kans op een baan. Bij deze voorbeelden lijkt de inzet op wat meer objectievere methodieken, maar ook hier geldt dat werkgevers zich bewust moeten zijn van mogelijke nadelen voor lhbti+ kandidaten. Er zijn geen voorbeelden bekend van de inzet van kunstmatige intelligentie bij promotiekansen voor lhbti+ personen op het werk.

Gezondheidszorg

Een belangrijk voordeel van de toepassing van kunstmatige intelligentie in de zorg is dat het patiëntenzorg van hogere kwaliteit kan bieden door verbeterde diagnose, behandeling en monitoring. Gezondheidsdata die voor dit doel worden verzameld, missen echter vaak informatie over gemarginaliseerde groepen zoals lhbti+ patiënten. Daarnaast hebben zij te maken met mechanismen binnen verschillende segmenten in de zorg die het belang van gender en seksuele diversiteit te weinig onderstrepen. Hierdoor missen lhbti+ patiënten mogelijk gerichte zorg en hebben zij mogelijk een lagere kans op goede gezondheid. Het is dus tijd voor gendersensitieve en lhbti+ sensitieve gezondheidszorg waarin 'verschil de nieuwe norm' wordt. Zodat alle patiënten de beste behandeling krijgen en zich begrepen en gehoord voelen.
Er zijn al mogelijkheden geïnventariseerd om data te ontsluiten over de gezondheid van lhbti+ gebruikers van online platforms. Dit is een stap in de goede richting. Onbekend is echter hoe deze data zijn verzameld en in hoeverre zij de privacyregels overtreden.

(Online) Veiligheid

Door de groei van het aantal gebruikers en de toename in het gebruik van sociale media, neemt ook de hoeveelheid ongestructureerde data toe. Voor de verwerking en analyse daarvan gebruiken steeds

meer platformen kunstmatige intelligentie. We hebben in de literatuur gezien dat kunstmatige intelligentie op een aantal manieren in het voordeel kan werken van lhbti+ personen, bijvoorbeeld bij de onzekerheid in beleid rondom asielzoekers die claimen lhbti+ te zijn. 'Een asielzoeker kan niet ‘bewijzen’ dat diegene lhbti+ is. Enkel de claim dat het zo is, is echter ook niet voldoende', merkt Severijns (2019) op. IND-medewerkers beschikken wel over instructies om gedurende een gesprek te beoordelen of een asielzoeker zich geloofwaardig kan verklaren over zijn of haar seksuele oriëntatie en/of genderidentiteit. De uiteindelijke afweging blijft echter voor een deel subjectief, zelfs met instructies. Een medewerker kan besluiten zijn vraagstellingen over mogelijk traumatische ervaringen te verzachten om een asielzoeker met angstklachten tegemoet te komen, terwijl een ander juist op strenge wijze doorvraagt met het idee om de waarheid boven tafel te krijgen. De kansen van een asielzoeker kunnen dus wezenlijk verschillen, afhankelijk van de IND-medewerkers die deze spreekt. De inzet van kunstmatige intelligentie zou deze afhankelijkheid van het ‘oordeel’ van de medewerker kunnen verkleinen. Dit is echter niet geheel zonder risico: het kan ook juist een valkuil zijn. Als je niet voldoet aan het plaatje van een lhbti+ asielzoeker zoals dat geprogrammeerd is (bijvoorbeeld omdat je niet uit de kast bent of niet aan bepaalde stereotype kenmerken voldoet), zou je door het systeem niet erkend kunnen worden en dus onterecht geen asiel krijgen.

Kunstmatige intelligentie kan daarnaast helpen bij het tegengaan van uitsluiting (door het gebruik van trefwoorden et cetera) op online sociale platforms waar lhbti+ personen zich begeven en het vergroten van hun online veiligheid. Tegelijkertijd rijst ook hier de vraag in hoeverre dit een nadeel kan zijn. Het onderzoek van Fosch-Villaronga (2020) laat immers zien dat juist deze digitale platforms filters gebruiken waardoor lhbti+ personen zich minder vrijelijk kunnen uiten.

Het is belangrijk dat we ons realiseren dat kunstmatige intelligentie werkt op basis van data(sets) en algoritmen – beiden worden door mensen geproduceerd. Met deze data en algoritmen kunnen machines menselijk gedrag nabootsen: zie het als een digitale medewerker die veel dingen heeft aangeleerd gekregen, maar zonder eigen (ethisch) denkkader. Kunstmatige intelligentie is nog in ontwikkeling, maar wat we merken, is dat relevante besluiten worden genomen die impact (kunnen) hebben op mensen. Volgens onderzoekers moeten we ons bewust zijn van ethische complicaties en transparantie in de totstandkoming van besluiten die op basis van kunstmatige intelligentie worden genomen. De overheid kan hierin een sturende rol innemen.

Deel 2

Expertmeeting

Dit tweede deel van het verkennend onderzoek naar kunstmatige intelligentie en lhbti+ emancipatie geeft de inzichten weer die op zijn gedaan middels een expertmeeting en een interview. Deze dienden een aantal doelen, namelijk 1) te verkennen waar kansen en risico's liggen wat betreft kunstmatige intelligentiesystemen voor lhbti+ personen, 2) onder welke voorwaarden veilige, privacy-vriendelijke en onbevooroordeelde kunstmatige intelligentiesystemen kunnen worden gebouwd en gebruikt opdat mensen uit de lhbti+ gemeenschap niet ongelijk behandeld worden, 3) de uitwisseling van praktijk-voorbeelden en veelbelovende aanpakken. De experts werden geselecteerd op hun kennis van kunstmatige intelligentie, lhbti+ vraagstukken en mensenrechten. De inzichten uit het interview werden als input gebruikt voor de expertmeeting.

Dit tweede deel van de rapportage geeft de resultaten weer van de expertmeeting. Net als in deel 1 bespreken we de resultaten per deelgebied. In hoofdstuk 4 gaan we in op de risico's en in hoofdstuk 5 op de kansen. Hoofdstuk 6 bevat een overzicht van de stakeholders die volgens de experts nodig zijn om bij te dragen aan eerlijke en inclusieve kunstmatige intelligentiesystemen binnen de vier deel-gebieden.

4) EXPERTS OVER DE RISICO'S VAN KUNSTMATIGE INTELLIGENTIE

4.1 Algemene risico's

Voor de experts zit aan kunstmatige intelligentiesystemen een aantal risico's van algemene aard waarmee rekening gehouden zou moeten worden. Een belangrijk deel daarvan gaat over 1) representativiteit van data, ontwikkelaars en besluitvormers en 2) over transparantie, vertrouwelijkheid en ethiek. Experts maken zich bijvoorbeeld zorgen dat de datasets niet altijd uit een evenredige, representatieve verzameling van gegevens bestaat. In de praktijk tonen deze sets vaak een overwicht aan heteroseksuele mannen en vrouwen.

'Je traint een algoritme met een dataset over wat mannen en vrouwen zijn. De machine gaat zelf leren welke patronen die erin herkent, de machine bepaalt dit helemaal zelf. Dit wordt een black box genoemd. Uiteindelijk is dus niet te achterhalen waar de machine die herkenningspatronen in ziet.' (Respondent 1)

Verder maken zij zich zorgen over het feit dat ontwikkelaars en besluitvormers van kunstmatige intelligentiesystemen zelf ook niet representatief zijn voor de lhbti+ gemeenschap en dat onduidelijk is wat het doel en de herkomst zijn van datasets. Dit brengt volgens hen het risico met zich mee dat oplossingen voor vraagstukken met een heteronormatieve bril worden bekeken en dat de uitkomst niet per se ten gunste is van de lhbti+ gemeenschap.

'Wie is de persoon die de datasets aanlevert voor systemen? En wie zorgt ervoor dat die datasets niet bevooroordeeld zijn? Het is als een kind dat je aan het opvoeden bent.' (Respondent 10)

'Ja, als je een kind leert 'dit is goed, en dit is fout'. Het houdt een traditioneel beeld vast, want zo'n algoritme wordt getraind op basis van data van hoe het er tot nu toe uit zag. Dus de machine blijft heel erg hangen in een traditioneel beeld, want die krijgt geen beelden van hoe de toekomst eruit kan zien.' (Respondent 1)

Aan de orde komen ook de transparantie bij de inzet van kunstmatige intelligentiesystemen en over de inspraak van lhbti+ personen bij het ontwerp en gebruik daarvan.

'Zelfs als je de hele wereld verplicht om een mensenrechten impacttoets te doen, kom je op de vraag in hoeverre er bewustzijn is bij mensen en ontwikkelaars over mensenrechten en lhbti+ specifieke zaken.' (Respondent 4)

'En wat gebeurt er met de data die worden opgeslagen? Wordt dit ook voor andere dingen ingezet? En wat als de data in handen komt van andere partijen?' (Respondent 2)

'Mensenrechtentoets ontbreekt. Voor sommige dingen mag kunstmatige intelligentie gewoon niet gebruikt worden. Eerst de mensenrechtentoets uitvoeren voordat kunstmatige intelligentie gebruikt mag worden. Dit gebeurt nog niet genoeg.' (Respondent 5)

‘Wie doet die mensenrechtentoets? Heel vaak is dat self-assessment. Moet je dit overlaten aan bedrijven of instanties? Het kan hier gaan over leven en dood. Wat is de ondergrens? Wat willen we dat er niet gebeurt met kunstmatige intelligentie? Die gesprekken moeten gevoerd worden. Op de plekken waar het gemaakt wordt worden deze gesprekken niet gevoerd en tegelijkertijd hebben de mensen waar het over gaat (lhbti+) niet voldoende technologische kennis.’ (Respondent 5)

Een derde risico dat wordt genoemd, is dat zelfs inclusieve datasets niet per definitie ten gunste hoeven te zijn van lhbti+ personen. Die kunnen juist uitsluiten. Een van de experts noemt het volgende voorbeeld:

‘Als data helemaal inclusief zijn en er komen bepaalde verbanden uit, dat bijvoorbeeld transgender personen hogere zorgkosten hebben. Wat wil je daar dan mee doen? Wat doe je met die data als samenleving? Ik heb weinig vertrouwen dat daar oplossingen voor zijn.’ (Respondent 4)

Ten slotte – als vierde risico – wijst één van de experts op de (online) veiligheid. De expert herinnert eraan dat commerciële partijen kunstmatige intelligentie inzetten om het gedrag van gebruikers te sturen, bijvoorbeeld met commerciële doelen. (Online)Veiligheid is dan van ondergeschikt belang voor deze partijen.

‘De meeste platformen worden commercieel gebouwd waardoor veiligheid in het geding komt. De uitdaging zit hierin: de technologiebouwer moet de waarden hebben, en waar komt het geld vandaan?’ (Respondent 3)

4.2 Onderwijs

Experts zien twee grote risico’s aan het Nederlands onderwijsbestel:

Risico 1: de binariteit van geslacht in de datasets die worden verzameld om de prestaties van leerlingen te meten en monitoren. Op basis van de opgelegde binariteit van man of vrouw zijn, worden heel wat besluiten genomen en de vraag is of die besluiten terecht genomen zijn. In het moderne tijdsgewricht staat deze binariteit steeds meer ter discussie en dat vraagt erom om deze inclusiviteit ook mee te nemen in steekproefselectie, dataverzameling, vraagstelling en data-analyse door ook bewust ruimte te laten voor mensen die zich identificeren als non-binair, aseksueel of genderfluïde.

‘In het onderwijs zie je dat verschillen tussen jongens en meisjes wel worden meegenomen (in monitoren), maar gender wordt binair gezien, en het gaat nooit over lhbti+.’ (Respondent 2)

Een expert merkte terecht op dat deze inclusiviteit in sommige landen niet zonder gevaar is.

‘Je gaat nooit cijfers krijgen over lhbti+. Er staat de doodstraf op in sommige landen. Er is geen veiligheid voor die data. Dit is een aspect van de samenleving, je gaat geen veiligheid krijgen om deze data bekend te hebben zolang de samenleving zo is. Het moet eerst gaan om het herdefiniëren van gender.’ (Respondent 3)

Risico 2: onduidelijkheid over het opslaan van data en het effect van gedragssturing. De toepassingen van kunstmatige intelligentie kunnen op verschillende wijze inbreuk maken op het privacyrecht.

Betrokkenen kunnen het gebruik en misbruik van persoonsgegevens als een ongewenste inmenging in hun persoonlijke levenssfeer ervaren. Het effect van dataverzameling in het onderwijs zou daarnaast kunnen leiden tot het onderdrukken van bepaald gedrag van leerlingen en hen in een keurslijf dwingen.

'*Vaak worden heel veel data verzameld van studenten, en dan ook nog opgeslagen. Als die data eenmaal vastligt, kan het zo interessant zijn voor een school om die data te bekijken. Risico daarbij is het chilling effect. Dat wil zeggen dat mensen hun gedrag aanpassen als ze weten dat ze in de gaten worden gehouden. Ze worden voorzichtiger. Terwijl het onderwijs een plek is om te leren en experimenteren en waar je mag falen. Studenten gaan zich voorzichtiger gedragen, en dat bevordert het leren niet.'* (Respondent 8)

4.3 Arbeidsmarkt

De discussie van de experts over de risico's van kunstmatige intelligentie voor de positie van lhbti+ personen richtte zich vooral op de voorkant van de arbeidsmarkt, namelijk de werving en selectie. De risico's liggen volgens de experts in een aantal gebieden. Ten eerste, de algoritmen die voor recruitmentsoftware worden gemaakt. Volgens de experts dienen die slechts een doel, namelijk om mensen die qua profiel lijken op de trainingsdata, aan te nemen en dat zijn (voornamelijk) witte mannen. Ten tweede bestaan werving- en selectieprocedures grotendeels uit het maken van testvragen, assessments, persoonlijkheidstests, het doen van spelletjes et cetera. Het is daarbij van belang altijd na te gaan vanuit welk perspectief testvragen, assessments, tests en opdrachten zijn ontwikkeld en hoe deze als doel hebben om personen te werven met vergelijkbare profielen als degene in managementposities (voornamelijk witte, hoogopgeleide mannen).

'In de formulering van die vragen kan al heteronormativiteit zitten. Dan kunnen mensen al niet doorgaan met de sollicitatie want ze voelen zich dan al niet welkom.' (Respondent 6)

Twee experts brachten in dat er voorbeelden zijn van bedrijven die eigen databestanden en recruitmentsoftware hebben laten ontwikkelen:

Voorbeeld van bedrijven met eigen databestanden en recruitmentsoftware	Voorbeeld van een bedrijf met genderbias
'Grote bedrijven als Hema en Praxis hebben hun eigen datasets en eigen recruitmentsofware. Die software wordt ontwikkeld en gebouwd door recruitmentbedrijf Harver. Wat we echter niet weten, is wat er in die datasets zit en welke typologie werknemers daaruit voortkomt. Dat is creepy. Maar managers zitten ook omhoog. Verder heb je Randstad en ING. Die hebben enorme innovatieafdelingen die voortvarend bezig zijn en elke drie maanden nieuwe dingen verzinnen.' (Respondent 6)	'Amazon en andere platforms bleven maar vooral witte, heteroseksuele mannen aantrekken, omdat de manier waarop ze kandidaten wierven, gebaseerd was op woorden die traditioneel gezien vooral geassocieerd worden met mannen. Dit hinderde de werving van vrouwen. In theorie gebruik je dus een systeem dat dingen slimmer en beter maakt. Je kunt zeggen 'technologie maakt alles beter', maar als je geen manier hebt om te checken of dat wel zo is, dan heb je er geen idee van wat er eigenlijk écht gebeurt.' (Respondent 11)

4.4 Gezondheidszorg

Ook bij dit deelgebied zien experts risico's van kunstmatige intelligentie voor de toegankelijkheid van de zorg. Deze liggen op het vlak van de ethiek, representatie en medische urgentie.

Risico 1: gebrek aan representatie in de datasets. Volgens de experts is er een gebrek aan goede data over de gezondheid van lhbti+ personen. Hierdoor is kwaliteit van de zorg, specifiek op het gebied van de zorgveiligheid van lhbti+ personen, onvoldoende gewaarborgd.

'De data die er tot nu toe is binnen de zorg, zijn binair ingestoken. Op 't moment dat je kijkt naar transgender personen en medische hulp… die zijn niet goed gerepresenteerd in de data. Ergens is ook een norm, dat de biologische opmaak niet overeenkomt met 'de norm'. Hoe zorg je dat veiligheid van zorg gewaarborgd blijft?' (Respondent 5)

'*Het gebrek aan representatieve informatie over lhbti+ personen heeft desastreuze gevolgen in de gezondheidszorg. In hoeverre moet je het willen als het niet representatief is? Hoe kun je zo'n fluïde community vatten in iets als kunstmatige intelligentie? In een ideale wereld hebben we systemen waarin data representatief is, maar hoe gaan we dat voor elkaar krijgen?'* (Respondent 4)

'*De lhbti+ community wordt gekenmerkt door variatie. Kunstmatige intelligentie werkt het best met opties. Kan je daarmee de community vangen? Transgender personen hebben over het algemeen hogere zorgkosten. De lhbti+ gemeenschap is harder getroffen.'* (Respondent 4)

Risico 2: ethische aspecten aan dataverzameling. Dit risico ligt in het verlengde van risico 1, namelijk dat experts bang zijn dat niet kan worden geanticipeerd op de uitkomsten van inclusieve(re) datasets. Experts stonden tegenover elkaar in de discussie over dit risico. Enerzijds werd het belang van inclusieve datasets gezien, anderzijds werd juist benadrukt in hoeverre het nodig is om inclusie in datasets te willen nastreven en wat de gevaren van inclusieve datasets kunnen zijn.

Een voorbeeld: als je depressie vroegtijdig kunt opsporen met kunstmatige intelligentie, maar het blijkt relatief vaker voor te komen bij lhbti+ personen, kan dat negatieve gevolgen hebben voor hun zorgpremie en -kosten.

'Dit is wel een ethisch dilemma. Als je weet wie kampen met depressie, is dat een risico. Dit wil je niet koppelen aan iemands naam.' (Respondent 8)

'Hoe herken je de niet-waarneembare kenmerken? Maar willen we überhaupt die data verzamelen? Wie gaat het gebruiken? Maar als je het niet verzamelt, is je data niet inclusief.' (Respondent 4)

Risico 3: inclusie van genderidentiteit is van ondergeschikt medisch belang. Experts wijzen erop dat het belangrijk is dat iedereen de juiste zorg ontvangt, maar betwijfelen in hoeverre men in de zorg bewust is van wat dit betekent in termen van genderidentiteit omdat er vaak een nadruk op de biologische sekse-identiteit ligt.

4.5 (Online) Veiligheid

Bij dit deelgebied signaleerden de experts drie belangrijke risico's.

Risico 1: de heteronormativiteit in het ontwerp van kunstmatige intelligentiesystemen en het soms moedwillig uitsluiten van leden van de lhbti+ gemeenschap.

Voorbeelden van de introductie van technologie ter uitsluiting van transgender personen	
De Giggle app werd in 2021 speciaal geïntroduceerd als sociale app waarmee alleen cisgender vrouwen – vrouwen die bij geboorte het vrouwelijke geslacht kregen toegewezen – met elkaar contact kunnen zoeken. Het algoritme maakt gebruik van gezichtsherkenningssoftware en bepaalt – althans in theorie – of je als cisvrouw een profiel mag aanmaken op het platform. Deze app is ontwikkeld door zogenaamde TERFS, een acroniem voor Trans-Exclusionary Radical Feminist. Dit zijn radicaal feministen die transgender vrouwen niet beschouwen als vrouwen en hen daarom uitsluiten van de strijd voor vrouwenrechten. Hierover ontstond ophef op Twitter waardoor veel transvrouwen, die 'passing' waren, profielen gingen maken om het algoritme te omzeilen en manipuleren. '*Het manipuleren van algoritmen kan ook een vorm van emancipatie zijn. Maar dan moet je wel weten hoe die algoritmen worden gemaakt en door wie*.' (Respondent 9)	Voor hun hypothese 'Worden homoseksuele mannen meer femininized dan heteroseksuele mannen?' zetten Fosch-Villaronga en een collega een enquête uit op Twitter. Zij kwamen erachter dat Twitter de genders afleidde van de gebruikers, maar in een kwart van de gevallen ernaast zat. 'Toen we dit beter bestudeerden, zagen we dat slechts 8% van de heteroseksuele mannen 'misgendered' werden, terwijl dit voor heteroseksuele vrouwen 16% was, voor biseksuele personen 20%, voor homoseksuele mannen 25% en 100% voor non-binaire personen. We concludeerden dat heteroseksuele vrouwen en personen die zich identificeren als lhbti+ vaker misgendered worden dan heteroseksuele mannen. De onderliggende kwestie is: gender classificatiesystemen nemen geslacht (vrouw, man, intersex) als uitgangspunt – dus de hele lhbti+ gemeenschap is al verdwenen in dat systeem. Maar gender is een subjectieve impressie en er is iets fundamenteel fout als je probeert iets dat inherent subjectief is te objectiveren.' (Respondent 11)

Noot: Passing is **de mogelijkheid van een persoon om zich voor te doen als iemand van een andere sociale groep dan de eigen groep**, bijvoorbeeld qua geslacht of etniciteit.

Risico 2: de inbreuk op de privacy en de gevolgen voor de (online) veiligheid.

'Wij hebben bijvoorbeeld een app gecreëerd voor lhbti+ jongeren. Het veiligheidsaspect is moeilijk, want er wordt gevraagd naar een kopie van het identiteitsbewijs om te kijken of iemand ook echt 18 jaar is en dat kun je nu wel monitoren. Als je gaat opschalen, wordt het problematisch. Initiatieven werken zolang het klein is. Als je kunstmatige intelligentie daar een rol in gaat laten spelen kan dat ook andere risico's met zich meebrengen.' (Respondent 5)

'Er zitten veel gemeenschappen op Facebook. Veiligheid zit de technologie altijd in de weg. Veiligheid wordt bijna nooit gewaarborgd. Er vindt wel omzeiling plaats om toch dat contact mogelijk te maken

zonder dat de veiligheid in het geding komt, bijvoorbeeld door gebruik te maken van Telegram in plaats van WhatsApp. Het is voor de veiligheid voor lhbti+ personen nodig om via die manier te communiceren. Maar het is niet helemaal inclusief, want je moet kennis hebben over bijvoorbeeld VPN om hierin te komen. Inmiddels weten we dat je op alle checklists van de CIA voorkomt als je activistisch bent. Daar moeten we niet ook nog een algoritme op zetten.' (Respondent 3*)*

'In het systeem sta ik geclassificeerd als gay, als ik in Saudi-Arabië aankom, ik kan door de poortjes en ze scannen niet m'n paspoort maar mijn gezicht, met software die ziet dat ik gay ben, dan zou ik rechtstreeks naar de gevangenis gaan.' (Respondent 3*)*

Risico 3: financiële uitputting en sturing van online gebruikers.

CASUS: Cambridge Analytica
Cambridge Analytica was een consultancy bedrijf dat datamining, data-analyse en direct marketing combineerde met strategische communicatie voor verkiezingscampagnes. Via een quiz op Facebook verzamelde het bedrijf data van gebruikers om hen te profileren. Hoewel de quiz een paar honderdduizend gebruikers had, liet Facebook het toe om via bepaalde methodes data van 87 miljoen gebruikers zonder hun toestemming te verzamelen. Die data werden nadien gebruikt voor zeer gerichte advertenties door middel van 'targeting', onder meer voor de verkiezing van Donald Trump in de VS. Zij hebben die targeting gebruikt om mensen te manipuleren om op Trump te stemmen. De feiten van de dataverzameling zelf dateren van 2014. Cambridge Analytica handelde weliswaar tegen de voorwaarden van Facebook in, maar Facebook zelf maakte het op haar platform tegelijk wel mogelijk om die data te verzamelen. Zie ook de Netflix film 'The great hack.' (Bron: Respondent 5)

5) EXPERTS OVER DE KANSEN VAN KUNSTMATIGE INTELLIGENTIE

5.1 Algemene kansen

Naast risico's zien experts ook kansen voor kunstmatige intelligentiesystemen. Deze hebben we onderscheiden in algemene kansen en kansen specifiek voor de vier deelgebieden.

In het algemeen vinden de experts dat er meer aansluiting gezocht moet worden bij bestaande veelal kleinschalige initiatieven en start-ups die zich richten op kunstmatige intelligentie. Volgens de experts hebben deze steun in de breedste zin van het woord nodig om een volgende stap te kunnen zetten.

'Er zijn verschillende start-ups die zich bezighouden met het ontwikkelen van technologieën en oplossingen, nieuwe ideeën et cetera. Er komen vaak hele goede ideeën uit, maar op een gegeven moment moeten zij op zoek naar investeringen. Onder investeerders is nu 4% vrouw bijvoorbeeld. Dus hoe werkt dit dan door in het werk van die start-ups? In beleid kan ervoor gekozen worden om zulke initiatieven en ontwikkelingen/ ontwikkelaars te steunen.'

Voorbeeld van veelbelovend kleinschalige initiatief
Roos Groothuizen is multimediakunstenaar met wortels in grafisch ontwerp. Ze is lid van het kunstcollectief Telemagic, lid van de lhbti+ gemeenschap en doet onderzoek naar de onethische kant van de data-industrie. Bij het maken van haar kunstinstallaties laat ze zich leiden door de volgende vragen: wat zijn de consequenties van kunstmatige intelligentie? Hoe kun je kunstmatige intelligentie op de juiste manier inzetten en bewapenen? Wat kan wetenschap leren van kunst en andersom? Voor haar werk genaamd 'The Black Box Bellagio' ontving ze eerder de Icarus Award. Deze prestigieuze prijs wordt uitgereikt aan kunstenaars en nieuwe makers die een kruisbestuiving tussen kunst en technologie creëren en via de kunst aandacht vragen voor de soms dunne lijn tussen technologie en ethiek.

Andere experts zijn de mening toegedaan dat de overheid de kans moet benutten om het datamonopolie van bedrijven te stoppen en meer kansen te bieden aan kleinere spelers.

'Binnen een kapitalistisch systeem wordt uitgegaan van concurrentie, waardoor je nieuwe ideeën stimuleert. Maar eigenlijk hebben grote techbedrijven, zoals Google, Meta etc., een monopolypositie op de huidige markt. Ga daar eens iets aan doen.' (Respondent 7)

In kunstmatige intelligentie werkt men steeds vaker met fysieke systemen, *embodiments* genoemd. Denk bijvoorbeeld aan chatbots en robots. Deze embodiments bieden kansen voor lhbti+ representatie volgens een expert.

'*AI gaat ook over virtual assistance, robots. Zodra je werkt met embodiments, kun je kiezen voor bepaalde representatie. Je kunt kiezen voor queer representatie, met uiterlijk kun je juist veel.'* (Respondent 8)

Ten slotte wijst een expert op de mogelijkheid om met behulp van kunstmatige intelligentiesystemen bestaande vooroordelen te bestrijden.

‘*We kunnen in feite systemen creëren die erop gericht zijn om vooroordelen te detecteren. We kunnen zelfs systemen hebben die zoeken naar discrepanties tussen verschillende datasets, verschillende populaties, enzovoort. Deze kunnen gebruikt worden om te bepalen of vooroordelen een rol hebben gespeeld, bijvoorbeeld bij werving- en selectieprocedures. Het is geen simpele taak voor mensen, maar het gebruik van computationele krachten is iets wat we zouden kunnen doen. We zouden ook veel meer inclusieve surveys moeten hebben wanneer we data verzamelen: we zouden meerdere opties van man/vrouw kunnen geven en dan ook vragen naar gender met verschillende opties. Mettertijd zullen we veel meer kennis vergaren met die data*.’ (Respondent 11)

5.2 Onderwijs

Binnen het deelgebied onderwijs worden kansen gezien voor de specifieke inhoud van opleidingen die zich meer kunnen richten op kennis over kunstmatige intelligentie maar ook bewustzijn creëren voor lhbti+ kwesties. Daarnaast zagen experts ook kansen voor het inclusiever maken van onderwijsculturen en toegang tot onderwijs in het algemeen.

Kans 1: sommige experts geven aan dat het huidige opleidingsaanbod tekortschiet en dat noodzakelijkerwijs geïnvesteerd zal moeten worden in het vak kunstmatige intelligentie, een soort mediawijsheid. Dit geldt ook voor docenten. Hierbij moet de ethiek achter kunstmatige intelligentie een belangrijk onderdeel zijn. De mensen die we opleiden, maken straks deel uit van een arbeidsmarkt die vergeven is van kunstmatige intelligentie. Het zou daarom goed zijn om mensen daarop voor te bereiden en hen leert hoe daar kritische vragen over te stellen. Bekeken zal moeten worden in hoeverre aangehaakt kan worden bij de nationale AI-cursus, een platform dat zoveel mogelijk Nederlanders de basisprincipes wil bijbrengen over kunstmatige intelligentie.

Kans 2: een tweede kans is dat er betere en inclusievere content wordt gemaakt voor docenten. Hiermee creëer je meer bewustzijn voor lhbti+ kwesties en hoe het onderwijsmateriaal daaraan kan bijdragen. Hier zou Movisie een belangrijke rol in kunnen vervullen door bijvoorbeeld vaker langs te gaan in het onderwijsveld en de noodzaak van dat bewustzijn voor lhbti+ personen, ook in het onderwijsmateriaal, onder de aandacht te brengen.

Kans 3: in de nieuwe opleiding kunstmatige intelligentie zal aandacht gevraagd moeten worden voor normen en grensoverschrijdend gedrag binnen docententeams. Binnen de technische opleidingen bestaat bijvoorbeeld een sterke cultuur die niet inclusief is. Door deze cultuur haken vrouwen, mensen van kleur en ongetwijfeld ook lhbti+ personen sneller af.

Kans 4: ten slotte kunnen inclusievere opleidingen een aantrekkingskracht uitoefenen op lhbti+ personen zelf. Experts specificeerden hier de kansen voor zowel de werving van leerlingen en studenten als de cultuur en het leefklimaat binnen opleidingen. Kunstmatige intelligentie biedt hierin een kans om potentiële daders van pesten op te sporen en op die manier preventief de onderwijsomgeving inclusief en veilig te houden. Het voorspellingsmodel, dat eerder werd aangemerkt als kans in de literatuurverkenning bij het deelgebied (online) veiligheid, zou hier een rol in kunnen spelen. Bijkomstig effect hiervan zou kunnen zijn dat depressieve gevoelens en zelfmoordgedachten

en -pogingen van lhbti+ personen preventief bij oorzaken, zoals uitsluiting en discriminatie, worden aangepakt (zie ook kansen bij gezondheidszorg verderop). Deze kans geldt niet alleen voor het deelgebied onderwijs maar biedt ook preventiemogelijkheden op de andere terreinen waar groepen samenkomen en een dynamiek van pesten kan ontstaan.

Kans 5: het vergroten van de toegankelijkheid van onderwijs.

'Als je in een gemarginaliseerde groep zit, heb je dan überhaupt toegang tot onderwijs? Armoede speelt een belangrijke rol, en er is ook cross-over tussen lhbti+ en armoede. Als je op jonge leeftijd uit de kast komt en uit huis wordt gezet, kom je in armoede, kun je überhaupt geen opleiding doen, zij hebben niet eens een computer. Ethische kant, issues rondom onderwijs in combinatie met lhbti+ en kunstmatige intelligentie: het gaat om het verhogen van toegang. Toegang komt voordat we het überhaupt hebben over kunstmatige intelligentie. Dus eerst toegankelijkheid, maar dan ook inclusie. Soms komen mensen wel binnen, maar haken af op de cultuur bij de opleidingen.' (Respondent 8)

5.3 Arbeidsmarkt

Wat betreft de arbeidsmarkt zien experts zowel kansen voor individuen als voor beleid.

Kans 1: individuen, zoals lhbti+ personen, kunnen gebruikmaken van allerlei trucs om online wervingssystemen te omzeilen. Ondanks dat het wenselijker zou zijn dat organisaties zelf stappen ondernemen om inclusief te worden, helpen deze manieren van omzeiling verandering binnen organisaties te forceren. Manieren om de risico's van kunstmatige intelligentie bij werving en selectie te omzeilen zijn bijvoorbeeld:

'Als sollicitant kun je gebruikmaken van AI-ondersteuning om het systeem tegen te werken om je cv er toch doorheen te laten komen. Dit kan door onzichtbare watermerken toe te voegen aan het cv die opgepikt worden door recruitmentsoftware.'

'Op LinkedIn kun je aanbevelingen genereren. Dan kom je op recruiterslijstjes.'

Kans 2: een expert wijst op een aanpak die door meerdere partijen wordt omarmd en die sollicitanten, ook uit de lhbti+ gemeenschap, meer arbeidsmarktkansen biedt.

Voorbeeld van een veelbelovende aanpak om ongelijkheid op de arbeidsmarkt tegen te gaan
House of Skills is een publiek-privaat samenwerkingsverband van het bedrijfsleven, branche-, werknemers- en werkgeversorganisaties, kennisinstellingen, onderwijs en bestuurders. Doel is te werken aan een meer op vaardigheden 'skills' gerichte arbeidsmarkt. Hierdoor komen de skills van werkzoekenden, die nu onzichtbaar blijven, in beeld. Daarmee ontstaat een breder beeld van de werkzoekenden en kunnen al hun skills een rol spelen bij het vinden van werk. Het belangrijkste voordeel van deze methode is dat het anoniem is. Er wordt immers alleen gevraagd naar je vaardigheden en niet naar persoonlijke karakteristieken. De methode is experimenteel. Op dit moment zijn onderzoekers aan het uitzoeken in hoeverre deze aanpak betere resultaten oplevert dan de traditionele wervingsmethode met een cv.

'Risico daarbij is dan wel dat bepaalde vaardigheden geassocieerd worden met masculien en feminien. En de vraag is ook, waar leg je de scheiding: havo/vwo. Schifting in werkervaring en opleidingsniveau is heel belangrijk. Vanuit mensenrechten mag ongelijke behandeling soms, maar alleen als je het kunt rechtvaardigen.' (Respondent 6)

Kans 3: in lijn met kans 4 in deelgebied onderwijs, kan kunstmatige intelligentie een rol spelen in het opsporen van potentiële daders van pesten op de werkvloer. Een voorspellingsmodel kan inzicht geven in de bedrijfscultuur en het werkklimaat en mogelijke risico's op uitsluiting en discriminatie aanpakken.

5.4 Gezondheidszorg

Kans 1: experts zien de ontbrekende kennis over lhbti+ personen als een uitgelezen kans om algoritmen te ontwikkelen die gezondheidsprofessionals in hun werk kunnen ondersteunen, bijvoorbeeld bij het controleren van de hormoonwaarden van transgender personen. Een ander voorbeeld waar experts naar verwezen, betreft de onduidelijkheid die bij gezondheidsprofessionals kan bestaan over aandoeningen bij transgender personen. Een expert haalde het voorbeeld van Brand Berghouwer aan. Hij ging in 2017 langs bij de gynaecoloog omdat hij last had van buikpijn en menstruatieklachten. De gynaecoloog verwees hem naar een psycholoog met verstand van gender. Brand kreeg ook medicijnen voorgeschreven. Maar hij bleef de klachten houden. Volgens een arts in de transkliniek was dit een typisch geval van Trans Broken Arm Syndrome. Dat betekent dat mensen voor een bepaalde zorgvraag bij een dokter komen, maar omdat ze trans zijn, de aandacht niet naar de zorgvraag gaat, maar naar het trans zijn. Dat kan bijvoorbeeld door klachten niet serieus te nemen, maar ook door invasieve vragen te stellen over hormonen of geslachtsveranderende operaties, zonder uit te leggen waarvoor het relevant is.

Voorbeeld van een veelbelovende aanpak om ontbrekende kennis over vrouwen aan te pakken

Met de opkomst van de femtech-markt is de focus op onderbelichte gezondheidsissues komen te liggen. Femtech verwijst naar 'Female Health Technology' en wordt gedefinieerd als apps, apparaten en technologische diensten die zich richten op de gezondheid en welzijn van vrouwen. Het is onder meer in opkomst binnen de categorie zelftests. Op basis van fysiologische gegevens die met deze technologie wordt verzameld, krijgt de gebruiker een indicatie van de eigen gezondheid. De gegevens kunnen daarnaast ook worden gebruikt voor het opsporen van onregelmatigheden en aandoeningen.

'Het gevaar ligt in het opslaan van data. Femtech vergroot de opties op het vlak van gezondheid en zelfregie: door het monitoren met behulp van apps krijgen vrouwen meer inzicht in de werking van het eigen lichaam. Maar hiermee wordt data opgeslagen en daar kleven gevaren aan. Kijk naar de abortuswet in de Verenigde Staten. Als vrouw ben je dan wel heel kwetsbaar wanneer je je menstruatie in zo'n app hebt geregistreerd en je data wordt gedeeld met onbekende partijen of opgevraagd door de overheid.' (Respondent 6)

Kans 2a: cijfers laten zien dat lhbti+ jongeren vaker kampen met depressieve gevoelens en zelfmoordgedachten. De mentale gezondheid van lhbti+ jongeren is in vergelijking met andere jongeren slechter. Kunstmatige intelligentie kan ingezet worden bij het vroegtijdig opsporen van suïcide.

Kans 2b: echter, zoals reeds benoemd onder onderwijs, kan kunstmatige intelligentie eveneens een rol spelen in het peilen van de houding ten opzichte van lhbti+ personen in de gezondheidszorg (van ziekenhuizen tot ggz en oudereninstellingen) en in het opsporen van potentiële pesters. Dit zou volgens experts waardevoller zijn dan de focus alleen te richten op het opsporen van daders van pesten of het vroegtijdig signaleren van suïcidepogingen bij lhbti+ jongeren.

'*Als je via kunstmatige intelligentie daders opspoort, ben je dat probleem via een andere weg aan het tackelen. Ga kijken naar potentiële pesters, in the long run is dat effectiever. Met AI kun je bijvoorbeeld suïcidaliteit signaleren, wat veel speelt bij lhbti+ jongeren. Maar: het probleem van de jongeren is niet hun lhbti+ zijn, maar hun omgeving is het probleem. Dat is de oorzaak van het probleem. Dus dat moet je aanpakken, niet de jongeren. Gebruik AI bijvoorbeeld voor het identificeren van daders en ga daarop inzetten.'* (Respondent 5)

Ten slotte wijst een expert er in het algemeen op dat het opeisen van het recht van spreken en bemoeienis door lhbti+ personen zeer belangrijk is in de gezondheidszorg waar het een verschil kan maken tussen leven en dood:

'Alle UMC's zijn bezig met AI en data. lhbti+ communities moeten aanschuiven bij deze gesprekken. Alle UMC's zijn bezig met behandelingen en AI. Dus we moeten als lhbti+ community daar bij zijn en van ons laten horen. Ze zijn overal nu die programma's aan het optuigen en data verzamelen.' (Respondent 3)

5.5 (Online) Veiligheid

Kunstmatige intelligentie maakt het mogelijk dat mensen elkaar online opzoeken en zij zich onderdeel voelen van een lhbti+ community. Het kan een bron van inspiratie zijn en het helpt lhbti+ personen in hun zelfbeeldontwikkeling en het herkennen van genderidentiteit. Kunstmatige intelligentie maakt het mogelijk om gelijkgestemden te vinden, die je in de samenleving niet zo snel en gemakkelijk zou kunnen vinden. De experts benadrukken dat het belangrijk is om deze rol van kunstmatige intelligentie te blijven zien en te blijven faciliteren:

Kans 1: focus op wat veiligheid betekent voor de gemeenschap en hoe je die kan bewerkstelligen middels kunstmatige intelligentie:

'Online kan ook juist echt een safe space bieden voor mensen. Voor veel mensen is in de normale wereld heel lastig om een veilige plek te vinden. Als mensen elkaar vinden via sociale media, bijvoorbeeld non-binaire en trans personen, die veel te maken hebben met negativiteit en geweld en gevaar. Zij kunnen met elkaar online een veilige omgeving vinden. Dan kan het breken van zo'n filterbubbel ook juist heel slecht/heftig voor hen zijn. Er wordt vergeten hoe belangrijk sociale media is voor gemarginaliseerde groepen. Privacy experts, vaak witte cis heteromannen, die zijn vaak geprivilegieerd, die hebben die safe space en sociale media niet nodig.' (Respondent 9)

Kans 2: opzetten en onderhouden van veilige ruimten. Deze kun je veilig houden, monitoren en beschermen zolang het op kleine schaal blijft. Maar als het groter wordt en als je wilt groeien, dan moet je iets anders zoeken.

Kans 3: inzet van sociale media en algoritmen bij lhbti+ emancipatie:

'Communities zijn heel groot geworden door de algoritmen die al die mensen bij elkaar gebracht hebben, bijvoorbeeld Black Lives Matter en nu de trans beweging. Dit is echt een movement geworden door sociale media. De manier waarop nu over genderidentiteiten wordt gesproken, zou niet gebeurd zijn zonder sociale media en die algoritmen. Ik denk dat dat een hele grote rol ook heeft gespeeld in lhbti+ emancipatie, en dan vooral trans emancipatie en non-binaire emancipatie.' (Respondent 1)

Kans 4: zicht krijgen op houdingen ten opzichte van lhbti+ personen in verschillende online en offline omgevingen en opsporen van potentiële daders van pesten van lhbti+ personen.

6) EXPERTS OVER STAKEHOLDERS

Aan de experts is ook gevraagd welke partijen zij als verantwoordelijke zien voor kunstmatige intelligentiesystemen die bijdragen aan eerlijk en inclusief/inclusieve onderwijs, arbeidsmarkt, gezondheidszorg en (online) veiligheid. De experts noemen verschillende stakeholders zoals de overheid, IT-bedrijven, wetenschappers, werkgevers en lbhti+ gemeenschappen. Daarbij benadrukken de meeste experts de urgentie van de verbinding van diverse stakeholders in het ontwikkel- en uitvoeringsproces van kunstmatige intelligentie. De expertmeeting zelf bleek een waardevolle werkvorm om experts uit verschillende werelden bij elkaar te brengen en in gesprek te kunnen gaan in een veilige ruimte. Vervolgstappen zullen dan ook ten alle tijden stakeholders uit deze werelden dienen te betrekken, zodat verschillende invalshoeken en ervaringen meegenomen kunnen worden bij het ontwikkelen van handelingsperspectieven. De twee belangrijkste redenen die hiervoor werden genoemd, zijn:

1. De kwetsbaarheid van lhbti+ personen en de bewustwording daarvan bij stakeholders.

'De lhbti+ gemeenschap is al kwetsbaar. We moeten alert zijn. Niet wachten op onderzoeken, actie ondernemen. Rol OCW: stimuleren/financieren toegewerkt worden naar kritisch tegengeluid, dat kunnen ze nu al doen. Agenderen bij andere ministeries.' (Respondent 5)

'Als we zelf de community erbij betrekken, voorstel aan OCW: niet alleen formele vertegenwoordiging van de community, maar ook lhbti+ jongeren die gebruikmaken van de systemen. Daarmee rekening houden met onderlinge verschillen, nuances, specifieke ervaring, intersectionaliteit.' (Respondent 9)

2. De datasets houden geen rekening met genderidentiteit en met de consequenties daarvan.

'Het zit nu vooral op gender, maar vooral binair. Mensen zijn wel welwillend, maar niet bewust dat ze uitsluiten. Mensen moeten er bewust van worden gemaakt hoe anderen worden uitgesloten.' (Respondent 2)

In het interview haalde de expert het voorbeeld van genderherkenningssystemen aan. Dit zijn systemen die met kunstmatige intelligentie gezichtsscans analyseren om zo, bijvoorbeeld, reclames op billboards aan te passen aan de voorbijganger die langsloopt maar het kan ook ingezet worden in meer gevaarlijkere situaties waarbij de toegang wordt ontzegd tot sanitaire ruimten of kleedkamers. Deze systemen herkennen transgender personen vaak niet:

'Je kunt proberen iedereen ervan te overtuigen dat je genderherkenningssystemen wilt hebben, maar misschien is het beter om het de mensen zelf te vragen? Bijvoorbeeld, misgendering ondermijnt autonomie. Je bent dan niet vrij om te uiten wie je bent en dit bedreigt de veiligheid van mensen. Het kan bovendien gebruikt worden als surveillance instrument. En het kan afwijzing veroorzaken; het kan eigenwaarde en zelfrespect beïnvloeden en voor sociaal stigma zorgen. Mensen moeten gaan nadenken wanneer er iets ontwikkeld wordt: kunnen we mensen daadwerkelijk schaden hiermee?' (Respondent 11)

Dit alles maakt een bredere samenwerking noodzakelijk, tussen bedrijfsleven, academische wereld en lhbti+ gemeenschap. De lhbti+ gemeenschap is vaak onwetend op het gebied van kunstmatige intelligentie, kwetsbaar en wordt veelal uitgesloten. Het belang van het bedrijfsleven is dat zij vooroploopt op innovatiegebied. De academische wereld kan een rol spelen bij de vertaalslag van kunstmatige intelligentie naar de realiteit, zeker als het gaat om ethische aspecten. In de literatuur kwamen we juist dit belang van het betrekken van de lhbti+ gemeenschap niet of nauwelijks tegen. De rol van de overheid wordt door de experts meer gezien als toezichthouder op de ethische aspecten, maar kan ook bestaan uit het stimuleren van initiatieven vanuit de lhbti+ gemeenschap om het kritische tegengeluid mogelijk te maken. Een andere rol van de overheid kan volgens de experts bestaan uit het agenderen van deze vraagstukken bij ministeries. Sommige experts zijn echter minder overtuigd. Volgens hen zullen genoemde argumenten onvoldoende zijn om alle stakeholders mee te krijgen.

'Er zijn echter ook mensen die er gewoon geen zin in hebben.' (Respondent 8)

Ten slotte is er een klein aantal experts dat zich afvraagt of je je niet eerst moet afvragen in hoeverre je met kunstmatige intelligentie vraagstukken van de lhbti+ gemeenschap kunt oplossen, voordat je gaat praten over stakeholders. Voor lhbti+ emancipatie zijn waarschijnlijk eerst andere interventies nodig.

'Moet je kunstmatige intelligentie hierbij willen toepassen? We weten waar de oplossingen zijn, dus kijk: heb je daar kunstmatige intelligentie bij nodig? En zo niet, dan niet doen.' (Respondent 5)

De tabel op de volgende pagina's geeft zowel de algemene stakeholders als stakeholders binnen de vier deelgebieden weer.

Figuur 1: Overzicht door experts genoemde stakeholders en hun verantwoordelijkheid		
	Stakeholder	**Verantwoordelijkheid**
Algemeen	Lhbti+ gemeenschappen	Mensen uit lbhti+ gemeenschappen die in aanraking komen met kunstmatige intelligentie, bijvoorbeeld jongeren die gebruikmaken van die systemen. Daarmee rekening houden met onderlinge verschillen, nuances, specifieke ervaring, intersectionaliteit. Belangrijk: focus niet alleen op belangenorganisaties zodat superdiversiteit binnen groepen meegenomen wordt!
	Lhbti+ belangenorganisaties	Belangenorganisaties als COC en Transgender Netwerk Nederland vertegenwoordigen lhbti+ personen uit Nederland.
	Nederlandse AI Coalitie	Deze coalitie doet onderzoek naar kunstmatige intelligentie. Dit gebeurt op initiatief van VNO-NCW MKB-Nederland, het ministerie van EZK, TNO, Seedlink, Philips, Ahold Delhaize, IBM en Topteam Dutch Digital Delta met als doel om Nederland in een voorhoedepositie te krijgen op het gebied van kunstmatige intelligentie. Zij bieden eveneens de Nationale AI-cursus aan waarbij kennis, informatie en vaardigheden over kunstmatige intelligentie wordt verspreid.
	Ministerie van OCW	Dit ministerie stimuleert en financiert onderzoek en aanpak van kunstmatige intelligentie en kan agenderen bij andere ministeries.
	College voor de Rechten van de Mens	Het College voor de Rechten van de Mens is het mensenrechteninstituut van Nederland en kan optreden als onafhankelijke toezichthouder bij uitsluitingsrisico's van kunstmatige intelligentie, specifiek voor lhbti+ personen.
	Movisie en kennis- en onderzoeksinstituten	Kennisinstituten kunnen ervoor zorgen dat informatie makkelijk te vinden is en dat data gebundeld en toegankelijk zijn.
	Investeerders	Kunstmatige intelligentie is een aantrekkelijke investering gezien de kansrijke toekomst die het biedt. Investeerders kunnen het mogelijk maken om nieuwe technologieën te ontwikkelen waarmee uitsluitingsrisico's verminderd worden.
	Lesbians Who Tech	Deze internationale speler is interessant omdat zij de werelden van technologie en lbhti+ al bijeenbrengen op mondiaal niveau.

Onderwijs	Alle opleidingen op alle niveaus	Opleidingen dienen niet alleen zelf bewust te worden over de impact van kunstmatige intelligentie voor hun eigen processen en kennis te hebben over de relatie tussen kunstmatige intelligentie en lhbti+, zij dienen deze kennis ook te verspreiden bij leerlingen en studenten.
	Innovatieafdelingen grote bedrijven	Bedrijven achter e-learningsystemen en platforms dienen hun verantwoordelijkheid te nemen om kritisch te kijken naar deze systemen en mogelijke uitsluitingsmechanismen op te sporen en aan te passen.
	Grote bedrijven met eigen databases voor training met kunstmatige intelligentie	Bedrijven met databases voor training door middel van kunstmatige intelligentie dienen hun verantwoordelijkheid te nemen om kritisch te kijken naar deze systemen en mogelijke uitsluitingsmechanismen op te sporen en aan te passen.
Arbeidsmarkt	Workplace Pride	Stichting Workplace Pride heeft als doel het verbeteren van de levens van lhbti+ personen in werkplekken wereldwijd. Ze streven naar een wereld met inclusieve werkplekken waar lbhti+ personen zichzelf kunnen zijn, gewaardeerd worden en, door hun bijdragen, kunnen helpen de weg te leiden voor anderen.
	Ontwikkelaars en gebruikers recruitment software	Ontwikkelaars van recruitmentsoftware kunnen het mogelijk maken om nieuwe technologieën te ontwikkelen waarmee uitsluitingsrisico's verminderd worden. Samenwerking met gebruikers van deze software is daarvoor essentieel.
Gezondheidszorg	Universitair Medische Centra	Kennis opdoen en bewustwording over de rol van kunstmatige intelligentie in medisch onderzoek, medische behandelingen en ziekenhuizen en de behoeften van en impact op lhbti+ personen.
	Zorgopleidingen	Opleidingen dienen niet alleen zelf bewust te worden over de impact van kunstmatige intelligentie voor hun eigen processen en kennis te hebben over de relatie tussen kunstmatige intelligentie en lhbti+, zij dienen deze kennis ook te verspreiden bij leerlingen en studenten.
	Ontwikkelaars software/systemen in zorgdomein	Bedrijven achter software en systemen gebruikt in de zorg dienen hun verantwoordelijkheid te nemen om kritisch te kijken naar deze systemen en mogelijke uitsluitingsmechanismen op te sporen en aan te passen. Dit in samenwerking met zorgprofessionals.

	Zorgprofessionals	Professionals in het zorgdomein zouden zich bewust moeten zijn van welke kunstmatige intelligentie systemen zij gebruiken en welke vooringenomenheid daarin ingebakken zit. Zij zijn een belangrijke partner voor softwareontwikkelaars en opleidingen.
(Online) Veiligheid	Privacy experts	Zij hebben de technologische knowhow over hoe privacy en veiligheid van lhbti+ personen online te waarborgen.
	Sociale media	Sociale mediaplatforms zijn een belangrijke stakeholder op het gebied van (online) veiligheid. Door hen te verbinden aan algemene stakeholders én stakeholders uit de andere terreinen, kunnen deze sociale media kritisch hun platforms bekijken en uitsluitingsmechanismen voorkomen. In de huidige tijd is dit hun verantwoordelijkheid voor een rechtvaardige wereld voor iedereen.

Deel 3

Conclusies en aanbevelingen

7) CONCLUSIES

Deze verkennende studie was opgezet om te achterhalen op welke wijze kansen en risico's van kunstmatige intelligentie voor de lhbti+ (on)gelijkheid voor verschillende deelgebieden kunnen worden benut c.q. beperkt. Om deze vraag te beantwoorden werd een literatuurstudie verricht. Verder werd een interview afgenomen en een expertmeeting gehouden met experts op het gebied van kunstmatige intelligentie en/of lhbti+ vraagstukken. Uit de literatuurstudie, het interview en de expertmeeting kwamen vooral naar voren dat er nog maar weinig onderzoek en kennis is van kunstmatige intelligentie en lhbti+ emancipatie en dat dit zeker geldt voor de deelgebieden onderwijs, arbeidsmarkt, gezondheidszorg en (online) veiligheid. De experts onderstreepten dan ook de noodzaak van onderzoek, inclusieve kunstmatige intelligentie en samenwerking met diverse stakeholders, waarbij de betrokkenheid van de lhbti+ gemeenschappen als noodzakelijke randvoorwaarde werd gezien. Een expert zei bijvoorbeeld dat kunstmatige intelligentiesystemen met de beste intentie kunnen worden gemaakt, maar dat er altijd ongewenste bijvangst is. Het is de uitdaging om die ongewenste bijvangst te beperken. Dit kan het beste gedaan worden door lhbti+ personen of hun vertegenwoordigers in de gebruikerstoets mee te nemen. Zij vormen een erg kwetsbare groep in de samenleving en hebben juist ondersteuning nodig om een stap verder te komen.

Net als bij andere onderzoeken naar kunstmatige intelligentie spraken de experts hun zorgen uit over de vertrouwelijkheid, transparantie, representativiteit, opslag, heteronormativiteit en de mensenrechtentoets van bestaande datasets. Bovendien onderstrepen de experts dat er ook makers en gebruikers van AI-systemen zijn die het maken van winst als hoogste doel hebben, waarbij er alleen interesse is in het verbeteren van de systemen als het hun winst vergroot; ook als dat betekent dat het ten koste gaat van lhbti+ personen of andere gemarginaliseerde groepen. Regelgeving en een verantwoorde inzet van kunstmatige intelligentie zijn volgens hen een must.

Een belangrijke vraag die zowel tijdens de literatuurstudie als de gesprekken met de experts opdook, is dat kunstmatige intelligentie vaak wordt gepresenteerd als dé oplossing van alle vraagstukken. Sommige experts vroegen zich af of dit ook echt zo is. Zouden wij niet eerst met elkaar moeten nagaan in hoeverre kunstmatige intelligentie passend en wenselijk is voor een bepaald vraagstuk? Volgens een expert is het antwoord dat de meeste vraagstukken van lhbti+ personen niet op te lossen zijn met kunstmatige intelligentie, omdat het vaak gaat om structurele sociale ongelijkheden die niet met een technologische ‘quick fix’ kunnen worden verholpen. In onderstaande conclusie zullen we elk van de onderzoeksvragen beantwoorden.

Onderzoeksvraag 1:
Welke kansen biedt kunstmatige intelligentie om lhbti+ gelijkheid op diverse deelgebieden te benutten?

Onderwijs
Kansen uit literatuurverkenning:

- Ondersteunen van onderwijsprofessionals.
- Interactie met menselijk ingrijpen.

Kansen uit expertmeeting:

- Investeren in inclusievere opleidingen en onderwijsculturen.
- Verhogen van onderwijstoegankelijkheid voor lhbti+ personen.
- Betere aansluiting van het onderwijs op de behoefte aan kennis over ethiek in kunstmatige intelligentiesystemen.
- Betere onderwijscontent ten behoeve van bewustwording onder leerkrachten over kunstmatige intelligentie en lhbti+ kwesties.
- Zicht krijgen op houding ten opzichte van lhbti+ personen en opsporen van potentiële daders van pesten van lhbti+ personen.

Arbeidsmarkt

Kansen uit literatuurverkenning:

- Betere inrichting van het wervings- en selectiesysteem zodat ze minder of niet discrimineren.

Kansen uit expertmeeting:

- Kennis van kunstmatige intelligentie inzetten om discriminerende of uitsluitende (recruiter)software te omzeilen.
- Aansluiten bij experiment ‘House of Skills’ dat focust op vaardigheden van sollicitanten in plaats van (persoonlijke) capaciteiten.
- Zicht krijgen op houding ten opzichte van lhbti+ personen en opsporen van potentiële daders van pesten van lhbti+ personen door bevoegde instanties.

Gezondheid

Kansen uit literatuurverkenning:

- Ruimte voor gepersonaliseerde adviezen.
- Filteren van data uit digitale platforms.

Kansen uit expertmeeting:

- Ondersteunen van zorgprofessionals met gerichte kennis over sekse, gender, seksuele diversiteit.
- Vroegtijdig opsporen van suïcidaal gedrag.
- Zicht krijgen op houding ten opzichte van lhbti+ personen en opsporen van potentiële daders van pesten van lhbti+ personen.

(Online) Veiligheid

Kansen uit literatuurverkenning:

- Benutten van accounts.
- Mogelijkheid tot online ondersteuning.
- Beter inzicht in maatschappelijke houding ten aanzien van lhbti+ personen en thema’s.

Kansen uit expertmeeting:

- Focus op wat veiligheid betekent voor de community en hoe je die kan bewerkstelligen.
- Opzetten en onderhouden van online en offline veilige ruimten.
- Inzet van sociale media en algoritmen ten gunste van lhbti+ emancipatie.
- Zicht krijgen op houding ten opzichte van lhbti+ personen en opsporen van potentiële daders van pesten van lhbti+ personen.

Onderzoeksvraag 2:
Welke risico's van kunstmatige intelligentie voor lhbti+ ongelijkheid zijn voor diverse deelgebieden te beperken?

Onderwijs
Risico's uit literatuurverkenning:

- Beperkte toegang tot onderwijs door gebruik algoritmen bij toelating.
- Beperkte toegang tot informatie over lhbti+ door gebruik filters.
- Ontoereikende mogelijkheden voor het ondersteunen van leerlingen.

Risico's uit expertmeeting:

- De binariteit van geslacht in de datasets die worden verzameld om de prestaties van leerlingen te meten en monitoren.
- Onduidelijkheid over het opslaan van data en het effect van gedragssturing.

Arbeidsmarkt
Risico's uit literatuurverkenning:

- (On)bedoelde impact op de positie van lhbti+ personen op de arbeidsmarkt.
- Kunstmatige intelligentie werkt nog niet vlekkeloos in het wervings- en selectieproces.
- Mogelijke discriminatie in geautomatiseerde online vacatures.

Risico's uit expertmeeting:

- De algoritmen die voor recruitmentsoftware worden gemaakt, dienen slechts een doel: om mensen die qua profiel lijken op de trainingsdata aan te nemen, en dat zijn (voornamelijk) witte mannen.
- Heteronormativiteit in recruitmentsoftware voor werving- en selectieprocedures.

Gezondheidszorg
Risico's uit literatuurverkenning:

- Gebrek aan transparantie en uitleg van medische beslissingen.
- Beperkte beschikbaarheid van kwalitatief goede gegevens.
- Bias en discriminatie in het ontwerp en bij de inzet van kunstmatige intelligentie.

Risico's uit expertmeeting:

- Gebrek aan representatie in de datasets.
- Ethische aspecten aan dataverzameling.
- Inclusie van genderidentiteit is van ondergeschikt medisch belang.

(Online) Veiligheid
Risico's uit literatuurverkenning:

- Inbreuk op grondrechten.
- Kwetsbaarheid.
- Delen van gevoelige informatie op mogelijk onveilige platforms.

Risico's uit expertmeeting:

- De heteronormativiteit in het ontwerp van kunstmatige intelligentiesystemen.
- De inbreuk op de privacy en de gevolgen voor de veiligheid.
- Financiële uitputting en sturing van onlinegebruikers.

Onderzoeksvraag 3:
Welke *best practices* zijn beschikbaar voor het omgaan met kansen en risico's van kunstmatige intelligentie voor lhbti+ (on)gelijkheid op verschillende deelgebieden?

Beschikbare *best practices* lopen uiteen en variëren van elkaar. Zo zijn er experimenten met vaardigheden bij het werven van geschikte kandidaten, maar zijn er ook femtech oplossingen die als voorbeeld kunnen dienen voor technologische ontwikkelingen. Daarnaast is er technologie die inzicht geeft of kan geven in de gezondheid van lhbti+ personen en kunstinstallaties die duidelijk maken wat de impact kan zijn van kunstmatige intelligentie.

Onderzoeksvraag 4:
Welke stakeholders zijn verantwoordelijk voor algoritmen die bijdragen aan lhbti+ gelijkheid op verschillende deelgebieden?

Die zijn divers: bedrijven, overheid, wetenschap, de nationale AI-coalitie en -cursus, kennisinstituten zoals het College voor de Rechten van de Mens en Movisie, maar bovenal lhbti+ gemeenschappen (zie hoofdstuk 6 voor een overzicht). Enerzijds wijzen experts erop dat lhbti+ gemeenschappen zelf het recht van spreken moeten opeisen; anderzijds is het belangrijk dat invloedrijke stakeholders hen actief en vanaf het begin uitnodigen én betrekken bij de ontwikkeling van inclusieve kunstmatige intelligentie.

Onderzoeksvraag 5:
Welke aanknopingspunten voor kunstmatige intelligentie-beleid en -interventies kunnen worden geformuleerd voor het bevorderen van lhbti+ gelijkheid op verschillende deelgebieden?

- Inventariseer samen met stakeholders (zoals genoemd in hoofdstuk 6) welke knelpunten er bestaan op de vier deelgebieden onderwijs, arbeidsmarkt, gezondheidszorg en (online) veiligheid. Ga daarvan na welke knelpunten met en zonder kunstmatige intelligentie kunnen worden opgelost. Daarbij is het belangrijk een kritische discussie aan te gaan over de vraag of kunstmatige intelligentie altijd de oplossing is, zonder dit zomaar aan te nemen.
- Betrek stakeholders vanaf het begin én in samenspraak om tot handelingsperspectieven en innovaties te komen waarin verschillende perspectieven, ervaringen en in- en uitsluitingsgronden meegenomen worden. De samenkomst van en samenspraak tussen verschillende werelden kan mogelijk resulteren in technologische innovaties binnen kunstmatige intelligentie, zoals manieren om voorbij binaire algoritmen te gaan.
- Een verantwoorde inzet van kunstmatige intelligentiesystemen is uitermate belangrijk. Besef dat lhbti+ personen in vergelijking tot andere groepen in de samenleving een kwetsbare positie hebben en dat het van belang is hen bij het ontwerp en de uitvoering van kunstmatige intelligentiesystemen te betrekken.
- Zet in op regelgeving die het mogelijk maakt om meer grip te krijgen op vertrouwelijkheid, representativiteit, transparantie, opslag, heteronormativiteit en de mensenrechtentoets van of bij bestaande datasets.
- Waar kunstmatige intelligentie nu wordt genoemd in het vroegtijdig opsporen van suïcidaal gedrag onder lhbti+ jongeren, zou het goed zijn om ook onderzoek te doen naar de oorzaken hiervan. Een proefproject waarbij kunstmatige intelligentie ingezet wordt om potentiële pesters op te sporen in

combinatie met bewustwordingscampagnes om stigma, vooroordelen en pesten te voorkomen, zou lhbti+ gelijkheid kunnen bevorderen op de verschillende deelgebieden.

VERANTWOORDING

Movisie heeft één interview gehouden met een expert op het gebied van kunstmatige intelligentie en lhbti+ vraagstukken. Naast dit interview werd een expertmeeting gehouden met 12 experts met kennis van kunstmatige intelligentie of van lhbti+ vraagstukken. Het interview en de expertmeeting vonden plaats in september 2022. Movisie heeft met de volgende experts gesproken:

Dr. Eduard Fosch-Villaronga

Eduard Fosch-Villaronga is universitair docent en onderzoeksdirecteur bij eLaw - Center for Law and Digital Technologies aan de Universiteit Leiden (NL). Hij onderzoekt de juridische en regelgevende aspecten van robot- en AI-technologieën. Hij houdt zich daarbinnen ook specifiek bezig met de intersectie van AI met lhbti+ personen. Zo deed hij onderzoek naar 'misgendering' van personen, en vooral van personen die tot de lhbti+ gemeenschap behoren, door bias in algoritmen.

Dr. Francisca Gromme

Francisca Grommé is universitair docent Organisational Dynamics in the Digital Society bij Erasmus Universiteit Rotterdam. In haar onderzoek bestudeert ze de relatie tussen digitalisering en de dagelijkse uitvoering van staats- en bedrijfsautoriteit. Ze bestudeert dit vanuit een 'practice' benadering: hoe technologieën en methoden, zoals kunstmatige intelligentie en big data analytics, meedoen in de organisatie van het dagelijks leven en hoe dit de uitkomst is van hoe deze technologieën en methoden worden voorgesteld, weerstaan, onderhandeld en geïmplementeerd. Ze kijkt vooral naar misdaad, statistieken, beleidsontwikkeling, publieke diensten en werving.

Dr. Maaike Harbers

Maaike Harbers is lector bij Kenniscentrum Creating 010. Zij doet onderzoek naar de relatie tussen kunstmatige intelligentie en de samenleving. Haar werk richt zich enerzijds op de ethische en sociale impact van AI op de samenleving en anderzijds op hoe ideeën en structuren in de samenleving de ontwikkeling van AI beïnvloeden. Zij is in het bijzonder geïnteresseerd in de rol van gender in relatie tot de ontwikkeling van AI. Maaike werkt samen met verschillende praktijkpartners, zoals media-organisaties, overheden en zorginstellingen. Met deze partners onderzoekt zij of en hoe kunstmatige intelligentie op verantwoorde wijze ontworpen en ingezet kan worden.

Arthur Vankan

Arthur Vankan is partner en senior consultant bij Dialogic en docent Kunstmatige Intelligentie op Beeckestijn Business School. Hij helpt beleidsmakers bij het monitoren, evalueren en ontwikkelen van beleid. Hierbij biedt hij beleidsmakers zoveel mogelijk inzicht in hun beleidsveld door het strategisch inzetten van data en houdt zich bezig met diverse publieke beleidsterreinen waaronder digitalisering, innovatie, onderwijs-arbeidsmarkt en zorg. Hij is jarenlang docent 'Data en Digitale methoden' geweest bij de Utrecht Data School en is tweevoudig wereldkampioen Professional Gaming in 2004 en 2005.

Katina Yiannakas

Katina Yiannakas is beleidsadviseur en wetgevingsadviseur bij het College voor de Rechten van de Mens, waar zij zich onder andere bezighoudt met de mensenrechten van lhbti+ personen. Zij is tevens werkzaam binnen de (strategische) programma's van het College die zich richten op gendergelijkheid en het VN-verdrag Handicap. Daarnaast is zij actief op het thema nationale veiligheid en mensenrechten. Katina is jurist, met een specialisatie in internationaal recht en mensenrechten.

Joey Stofberg

Joey Stofberg is werkzaam als PhD-onderzoeker aan de Vrije Universiteit Amsterdam, waar die zich toespitst op maatschappelijke factoren op verschillende schaal-niveaus die van invloed zijn op het welzijn van lhbti+ jongeren van kleur in Nederland. Tevens is Joey actief als European Fellow bij Humanity in Action. Joey's onderzoek bevindt zich op het snijvlak van politiek, nieuwe media, technologie en cultuur. Deze onderwerpen komen samen in het thema 'belonging' in post-digitale en geglobaliseerde contexten. Eerder behaalde Joey een bachelorgraad in Geschiedenis aan de Universiteit van Amsterdam en voltooide die een onderzoeksmaster in de Internationale Betrekkingen aan de Rijksuniversiteit Groningen, beide *cum laude*. In hun masterscriptie verkende Joey de nieuwe mogelijkheden voor politiek-maatschappelijke zelfrepresentatie die sociale media (jeugdige) lhbti+ gemeenschappen hebben geboden sinds 2010.

Marie Ricardo

Marie Ricardo is strategisch directeur van COC Nederland. Marie Ricardo is sinds 2019 manager nationale programma's bij COC Nederland. Ze vervulde verschillende projectmanagement-rollen sinds ze in 2013 bij het COC in dienst kwam. Zo ondersteunde ze lhbti+ activisten in verschillende Afrikaanse en Caribische landen bij het versterken en verbinden van hun beweging. Daarnaast vervulde ze bestuursfuncties bij PlanetRomeo Foundation, Pride Amsterdam en tot op heden bij Atria. Marie Ricardo studeerde Public International Law aan de Universiteit van Amsterdam en de San Francisco State University, en daarvoor Management, Economy & Law aan de Hogeschool van Amsterdam.

Marion Mulder

Marion Mulder is specialist in technologische innovatie op het snijvlak van kunstmatige intelligentie en diversiteit. Sinds december 2021 is zij voorzitter van de Gaykrant. Ze is werkzaam geweest als Global Project Manager Diversity & Inclusion bij ING en mede-oprichter geweest van de Workplace Pride Foundation. Tegenwoordig is ze eigenaar en oprichter van MuldiMedia, een bedrijf dat mensen en organisaties ondersteunt bij design thinking en digitale technologie om een wereld te creëren waar zij in wil leven. Ze focust daarbij op de ethiek van kunstmatige intelligentie met als doel om beter te begrijpen hoe we ervoor kunnen zorgen dat opkomende technologieën een wereld creëren die goed en rechtvaardig is voor iedereen.

Dr. Jaswina Elahi

Jaswina Elahi is senior onderzoeker bij het Rathenau Instituut. Haar promotieonderzoek ging over het ontstaan van digitale etnische netwerken en instituties als gevolg van het internetgebruik onder verschillende categorieën jongeren. Als senior onderzoeker heeft ze een sterke oriëntatie op ethische dimensies van het leven en oog voor de bijzondere betekenis van details, maar wel met een brede blik voor allerlei stakeholders, maatschappelijke ontwikkelingen en 'governance' van de technologie. Zo doet zij onderzoek naar de impact van technologie en innovatie op de brede gezondheid en democratische betrokkenheid van kwetsbare, achtergestelde burgers, zoals migranten, kwetsbare ouderen, jongeren en vrouwen.

Roos Groothuizen

Roos Groothuizen is mediakunstenaar en doet onderzoek naar de onethische kant van de data-industrie. Ze is geïnteresseerd in de consequenties van kunstmatige intelligentie en maakt grote spelinstallaties over (de effecten van) kunstmatige intelligentie. Daarmee belicht ze de menselijke kant van deze systemen en hoe algoritmen ons systematisch discrimineren. Ze is bovendien lid van het kunstcollectief Telemagic, een open media-lab dat mysteries rondom digitale technologie probeert te ontrafelen. Voor haar werk Black Box Villagio ontving ze in 2017 de Icarus Award. Deze award wordt ieder jaar uitgereikt aan een jonge kunstenaar die via kunst aandacht vraagt voor de soms dunne lijn tussen technologie en ethiek.

Tieneke Sumter

Tieneke Sumter is manager bij COC Nederland. Ze is al 40 jaar een voorvechtster voor sociale rechtvaardigheid en gelijke rechten in Nederland en Suriname. Zij was al op jonge leeftijd betrokken bij de Surinaamse Homo-Organisatie (SUHO), Sister Outsider en later bij het LGBT Platform Suriname. Door haar brede werkervaring op het gebied van mensenrechten, seksuele reproductieve gezondheid en cultureel maatschappelijk vormingswerk is zij een echte verbinder en bruggenbouwer.

Ian Vassaire

Ian Vassaire is beleidsmedewerker bij het COC. Hij heeft zijn Bachelor Social Work en de master Culture, Organization and Management aan de Vrije Universiteit Amsterdam afgerond. Hij hoopt een bijdrage te leveren aan inclusiviteit binnen organisaties, waar iedereen, maar vooral lhbti+ personen uit Amsterdam en omgeving zich veilig en thuis voelen.

Eline Hustinx

Eline Hustinx is stagiair bij het College voor de Rechten van de Mens, waar zij zich vooral bezighoudt met mensenrechten en digitalisering. Daarnaast heeft ze een interesse in de rechten van lhbti+ personen. Ze heeft een achtergrond in de filosofie en het recht en volgt nu de master Staats- en Bestuursrecht aan de Radboud Universiteit Nijmegen.

LITERATUURLIJST

Abràmoff, M., Tobey, D., & Char, D. (2020). Lessons Learned about Autonomous AI: Finding a Safe, Efficacious, and Ethical Path through the Development Process. American Journal of Ophthalmology. Opgehaald van https://www.frontiersin.org/articles/10.3389/frai.2021.736697/full

Acemoglu, D., & Restrepo, P. (2019). Artificial Intelligence, Automation, and Work. The Economics of Artificial Intelligence: An Agenda. Opgehaald van https://www.nber.org/system/files/chapters/c14027/c14027.pdf

AI Index 2018. (2018). Artificial Intelligence Index 2018. Opgehaald van http://cdn.aiindex.org/2018/AI%20Index%202018%20Annual%20Report.pdf

Anasua Kundu, e. a. (2021). Machine Learning Applications in Mental Health and Substance Use Research Among the LGBTQ2S+ Population: Scoping Review. JMIR Med Inform. Opgehaald van https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8663464/

Anderson, D. (2017). Stanford University research under ethical review after falsely claiming that artificial intelligence could determine a person's sexual orientation. Opgehaald van GLAAD: https://www.glaad.org/blog/stanford-university-research-under-ethical-review-after-falsely-claiming-artificial

Aragon, S. e.a. (2012) The Influence of Peer Victimization on Educational Outcomes for LGBTQ and Non-LGBTQ High School Students, Journal of LGBT Youth, 11:1, 1-19, DOI: 10.1080/19361653.2014.840761

Badget, M., Waaldijk, K., & Van der Meulen Rodgers, Y. (2019). The relationship between LGBT inclusion and economic development: Macro-level evidence. World Development. Opgehaald van https://www.sciencedirect.com/science/article/pii/S0305750X19300695

Bélisle-Pipon, e. a. (2021). What Makes Artificial Intelligence Exceptional in Health Technology Assessment? Front. Artif. Intell. Opgehaald van https://www.frontiersin.org/articles/10.3389/frai.2021.736697/full#B27

Bittner, M., Dettmar, D., Jaramillo, D. M., & Valta, M. J. (2020). Virtual Tribes: Analyzing Attitudes Toward the LGBT Movement by Applying Machine Learning on Twitter Data. Digital Transformation of Collaboration. Opgehaald van https://link.springer.com/chapter/10.1007/978-3-030-48993-9_12

Blauw, S. (2019). Wat is een algoritme? Opgehaald van De Correspondent: https://decorrespondent.nl/10306/wat-is-een-algoritme/149980270484-745de161

Bridgeman, B. e.a. (2012) Comparison of Human and Machine Scoring of Essays: Differences by Gender, Ethnicity, and Country, Applied Measurement in Education, 25:1, 27-40, DOI: 10.1080/08957347.2012.635502

Bureau van de Europese Unie voor de grondrechten. (2021). Op het juist pad voor de toekomst. Artificiële intelligentie en de grondrechten. Luxemburg: Bureau voor publicaties van de Europese Unie. Opgehaald van https://fra.europa.eu/sites/default/files/fra_uploads/fra-2021-artificial-intelligence-summary_nl.pdf

Burke, e. a. (2016). Do Contact and Empathy Mitigate Bias Against Gay and Lesbian People Among Heterosexual Medical Students? Acad Med. Opgehaald van https://pubmed.ncbi.nlm.nih.gov/25674910/

Cavalcante, A. (2018). Tumbling Into Queer Utopias and Vortexes: Experiences of LGBTQ Social Media Users on Tumblr. Journal of Homosexuality. Opgehaald van https://www.tandfonline.com/doi/full/10.1080/00918369.2018.1511131

Challen, e. a. (2019). Artificial intelligence, bias and clinical safety. BMJ Quality & Safety. Opgehaald van https://qualitysafety.bmj.com/content/qhc/28/3/231.full.pdf

Chen, L., Chen, P., & Lin, Z. (2020). Artificial Intelligence in Education: A Review. IEEE Access. Opgehaald van https://ieeexplore.ieee.org/abstract/document/9069875

COC. (2017). Onderzoek gezichtsherkenning ‘doodeng’. Opgehaald van COC: https://www.coc.nl/homepage/onderzoek-gezichtsherkenning-doodeng
College voor de Rechten van de Mens. (2020). Rechter oordeelt dat fraudedetectiesysteem SyRI in strijd is met mensenrechten. Opgehaald van https://mensenrechten.nl/en/node/2650
COMEST. (2019). Preliminary study on the ethics of Artificial Intelligence. Opgehaald van https://unesdoc.unesco.org/ark:/48223/pf0000367823
Costanza-Chock, S. (2018). Design justice, AI, and escape from the matrix of domination. Journal of Design and Science. Opgehaald van https://jods.mitpress.mit.edu/pub/costanza-chock/release/4
Crawford, K. (2016). Artificial Intelligence’s White Guy Problem. The New York Times (Opinion). Opgehaald van https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html
Dalenberg, D. J. (2018). Preventing discrimination in the automated targeting of job advertisements. Computer Law & Security Review. Opgehaald van https://www.sciencedirect.com/science/article/abs/pii/S0267364917303758
Daugherty, P. R., Wilson, H. J., & Cho, R. (2019). Using Artificial Intelligence to Promote Diversity. MIT Sloan Management Review. Opgehaald van https://www.proquest.com/openview/7b528c41903fcfe0494b67d93f0e4447/1?pq-origsite=gscholar&cbl=26142
Dignum, V. (2017). Responsible Artificial Intelligence: Designing AI for Human Values. ITU Journal. Opgehaald van http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/2181/itu2017-1.pdf?sequence=1&isAllowed=y
Doorman, M. (2018). Innovatieperikelen over artificial intelligence in het onderwijs. In A. Romeijn, M. Doorman, & C. Muller, De vraag is wat wij willen, niet wat wij hadden gewild. Ministerie van OCW. Opgehaald van https://kennisopenbaarbestuur.nl/documenten/de-vraag-is-wat-wij-willen-niet-wat-wij-hadden-gewild/
Element AI. (2019). Global AI Talent Report 2019. Opgehaald van https://jfgagne.ai/talent-2019/
EU Agency for Fundamental Rights (2020) Fundamental Rights Report 2020. https://fra.europa.eu/en/publication/2020/fundamental-rights-report-2020
Fallin-Bennett, K. (2015). Implicit Bias Against Sexual Minorities in Medicine. Academic Medicine. Opgehaald van https://journals.lww.com/academicmedicine/fulltext/2015/05000/Implicit_Bias_Against_Sexual_Minorities_in.8.aspx
Fenech, M. E., & Buston, O. (2020). AI in Cardiac Imaging: A UK-Based Perspective on Addressing the Ethical, Social, and Political Challenges. Front. Cardiovasc. Med. Opgehaald van https://www.frontiersin.org/articles/10.3389/fcvm.2020.00054/full
Ge Huang, e. a. (2019). Inferring Opinions and Behavioral Characteristics of Gay Men with Large Scale Multilingual Text from Blued. Health and Wellbeing in Sexual Orientation and Gender Identity. Opgehaald van https://www.mdpi.com/1660-4601/16/19/3597
GLAAD. (2017). GLAAD and HRC call on Stanford University & responsible media to debunk dangerous & flawed report claiming to identify LGBTQ people through facial recognition technology. Opgehaald van GLAAD: https://www.glaad.org/blog/glaad-and-hrc-call-stanford-university-responsible-media-debunk-dangerous-flawed-report
GlobalOrange. (2022). Machine learning, kunstmatige intelligentie (A.I.) en deep learning. Opgehaald van GlobalOrange: https://www.globalorange.nl/artificial-intelligence-machine-learning-en-deep-learning
Gupta, S. e. (2021). Assessing whether artificial intelligence is an enabler or an inhibitor of sustainability at indicator level. Transportation Engineering. Opgehaald van https://www.sciencedirect.com/science/article/pii/S2666691X21000208

Hamidi, F., Scheuerman, M. K., & Branham, S. M. (2018). Gender Recognition or Gender Reductionism? The Social Implications of Automatic Gender Recognition Systems. CHI. Opgehaald van https://docs.wixstatic.com/ugd/eb2cd9_ff211774807946099e7e1dcd0023497d.pdf

Han, H., & Jain, A. K. (2014). Gender and Race Estimation from Unconstrained Face Images. MSU Technical Report. Opgehaald van http://biometrics.cse.msu.edu/Publications/Face/HanJain_UnconstrainedAgeGenderRaceEstimation_MSUTechReport2014.pdf

Hansson, M. (2018). New professor with focus on ethical and societal impact of AI. Opgehaald van WASP: https://wasp-sweden.org/new-professor-with-focus-on-ethical-and-societal-impact-of-ai/

Harbers, M. (2018). Verstand erbij: verantwoord ontwerp van toepassingen met kunstmatige intelligentie. Hogeschool Rotterdam.

Huijnk, e.a. (2022) LHBT-monitor 2022. De leefsituatie van lesbische, homoseksuele, biseksuele en transgender personen in Nederland. SCP Rapport. Opgehaald van: https://www.scp.nl/publicaties/publicaties/2022/07/05/lhbt-monitor-2022

Institute of Medicine. (2011). The Health of Lesbian, Gay, Bisexual, and Transgender People. Opgehaald van https://nap.nationalacademies.org/catalog/13128/the-health-of-lesbian-gay-bisexual-and-transgender-people-building

Karami, e. a. (2021). Automatic Categorization of LGBT User Profiles on Twitter with Machine Learning. Machine Learning Technologies for Big Data Analytics. Opgehaald van https://www.mdpi.com/2079-9292/10/15/1822

Keyes, O. (2018). The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition. Proceedings of the ACM on Human Computer Interaction. Opgehaald van https://dl.acm.org/citation.cfm?doid=3290265.3274357

Khatua, A., Cambria, E., Ghosh, K., Chaki, N., & Khatua, A. (2019). Tweeting in Support of LGBT?: A Deep Learning Approach. Proceedings of the ACM India Joint International Conference on Data Science and Management of Data. Opgehaald van https://dl.acm.org/doi/abs/10.1145/3297001.3297057

Klasen, S. (2002). Low Schooling for Girls, Slower Growth for All? Cross-Country Evidence on the Effect of Gender Inequality in Education on Economic Development. The World Bank Economic Review. Opgehaald van https://academic.oup.com/wber/article-abstract/16/3/345/1677348

Klasen, S., & Lamanna, F. (2009). The Impact of Gender Inequality in Education and Employment on Economic Growth: New Evidence for a Panel of Countries. Feminist Economics. Opgehaald van https://www.tandfonline.com/doi/full/10.1080/13545700902893106

Knowles, S. e. (2002). Are educational gender gaps a break on economic development? Some cross-country empirical evidence. Oxford Economic Papers. Opgehaald van http://refhub.elsevier.com/S0305-750X(19)30069-5/h0235

Kromhout, M. &. S. Wolf (2019) Verkennend onderzoek naar biculturele transpersonen in Amsterdam. Opgehaald van: Verkennend onderzoek naar biculturele transpersonen in Amsterdam - Het Onderzoekerscollectief.

Leavy, S. (2018). Gender bias in artificial intelligence: the need for diversity and gender theory in machine learning. Proceedings of the 1st International Workshop on Gender Equality in Software Engineering. Opgehaald van https://dl.acm.org/doi/abs/10.1145/3195570.3195580

Leslie, e. a. (2021). Does “AI” stand for augmenting inequality in the era of covid-19 healthcare? BMJ.

Li, e. a. (2020). Depressive Emotion Detection and Behavior Analysis of Men Who Have Sex With Men via Social Media. Front. Psychiatry. Opgehaald van https://www.frontiersin.org/articles/10.3389/fpsyt.2020.00830/full

Lucero, L. (2017). Safe spaces in online places: social media and LGBTQ youth. Multicultural Education Review. Opgehaald van https://www.tandfonline.com/doi/full/10.1080/2005615X.2017.1313482?src=recsys

Matheny, e. a. (2019). Artificial Intelligence in Health Care. A Report From the National Academy of Medicine. JAMA. Opgehaald van https://jamanetwork.com/journals/jama/article-abstract/2757958

McDermott, E. (2015). Asking for help online: Lesbian, gay, bisexual and trans youth, self-harm and articulating the ‘failed’self. Health. Opgehaald van https://journals.sagepub.com/doi/abs/10.1177/1363459314557967
McDermott, E., & Katrina, R. (2012). Youth on the virtual edge: Researching marginalized sexualities and genders online. Qualitative Health Research. Opgehaald van https://journals.sagepub.com/doi/abs/10.1177/1049732311425052
McKinsey Global Institute. (2017). Jobs lost, jobs gained: What the future of work will mean for jobs, skills and wages. Opgehaald van https://www.mckinsey.com/featured-insights/future-of-work/jobs-lost-jobs-gained-what-the-future-of-work-will-mean-for-jobs-skills-and-wages
Meiselwitz, G. (2019). Social Computing and Social Media. Design, Human Behavior and Analytics. Switzerland: Springer Nation. Opgehaald van https://books.google.com/books?hl=en&lr=&id=WKShDwAAQBAJ&oi=fnd&pg=PR10&ots=qF3F6xyZQp&sig=YBUx5atC4TKlDc0MUu5-gOXt8Qg
Mesko, B. (2017). The role of artificial intelligence in precision Medicine. Expert Review of Precision Medicine and Drug Development.
Miller, B. (2021). Is Technology Value-Neutral? Science, Technology, & Human Values, 46(1), 53-80. https://doi.org/10.1177/0162243919900965
Mishra, V., & Srikumar, M. (2017). Predatory Data: Gender Bias in Artificial Intelligence. In S. (. Saran, Digital Debates: CyFy Journal 2017. New Delhi: Observer Research Foundation.
Muller, C. (2018). Artificiële Intelligentie heeft een human-in-command benadering nodig. In A. Romeijn, M. Doorman, & Muller Catelijne, De vraag is wat wij willen, niet wat wij hadden gewild. Ministerie van OCW. Opgehaald van https://kennisopenbaarbestuur.nl/documenten/de-vraag-is-wat-wij-willen-niet-wat-wij-hadden-gewild/
Nkonde, M. (2019). Automated anti-blackness: facial recognition in Brooklyn, New York. Harvard Kennedy School Journal of African American Policy. Opgehaald van https://pacscenter.stanford.edu/wp-content/uploads/2020/12/mutalenkonde.pdf
Ocumpaugh, J. e.a. (2014) Population validity for educational data mining models: A case study in affect detection. https://doi.org/10.1111/bjet.12156
O'Neil, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens.
O'Neil, C. (2018). Amazon’s Gender-Biased Algorithm Is Not Alone. Bloomberg Opinion. Opgehaald van https://www.bloomberg.com/opinion/articles/2018-10-16/amazon-s-gender-biased-algorithm-is-not-alone
Parikh, R. B., Teeple, S., & Navathe, A. S. (2019). Addressing Bias in Artificial Intelligence in Health Care. JAMA. Opgehaald van https://jamanetwork.com/journals/jama/article-abstract/2756196
Pijpers, R. (2021) Hoe ondersteun je ouderen met lesbische, homo-, bi- of transgender (LHBT)gevoelens? Opgehaald van: https://www.movisie.nl/publicatie/hoe-ondersteun-je-ouderen-lhbt-gevoelens
Piketty, T. (2017). Capital in the twenty-first century. Harvard University Press.
ProPublica. (2016). Machine Bias. Opgehaald van https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing
Raad voor Volksgezondheid & Samenleving. (2019). Waarde(n)volle zorgtechnologie. Een verkennend advies over de kansen en risico's van kunstmatige intelligentie in de zorg. Opgehaald van https://kennisopenbaarbestuur.nl/documenten/waarde-n-volle-zorgtechnologie-een-verkennend-advies-over-de-kansen-en-risico-s-van-kunstmatige-intelligentie-in-de-zorg/
Raghavan, M., Barocas, S., Kleinberg, J., & Levy, K. (2020). Mitigating bias in algorithmic hiring: evaluating claims and practices. Proceedings of the 2020 Conference on Fairness, Account-ability, and Transparency. Opgehaald van https://dl.acm.org/doi/10.1145/3351095.3372828
Reddy, e. a. (2019). A Governance Model for the Application of AI in Health Care. Journal of the American Medical Informatics Association. Opgehaald van https://academic.oup.com/jamia/article/27/3/491/5612169

Reuters. (2018). Amazon ditched AI recruiting tool that favored men for technical jobs. The Guardian. Opgehaald van https://www.theguardian.com/technology/2018/oct/10/amazon-hiring-ai-gender-bias-recruiting-engine

Reuters. (2018). Amazon scraps secret AI recruiting tool. Opgehaald van https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G

Richardson, R., Schultz, J., & Crawford, K. (2019). Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice. NYU Law Review Online. Opgehaald van https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3333423

Richardson C. e.a. (2002) Does Pornography-Blocking Software Block Access to Health Information on the Internet? JAMA. doi:10.1001/jama.288.22.2887

Roll, I., & Wylie, R. (2016). Evolution and Revolution in Artificial Intelligence in Education. International Journal of Artificial Intelligence in Education. Opgehaald van https://link.springer.com/article/10.1007/s40593-016-0110-3

Saha, K. e. (2019). The Language of LGBTQ+ Minority Stress Experiences on Social Media. Proc ACM Hum Comput Interact. Opgehaald van https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7489301/

Samudzi, Z. (2019). Bots Are Terrible at Recognizing Black Faces. Let's Keep it That Way. The Daily Beast. Opgehaald van https://www.thedailybeast.com/bots-are-terrible-at-recognizing-black-faces-lets-keep-it-that-way

Schans, J. M., & Van Lierop, L. E. (2019). De geloofwaardigheidsbeoordeling van asielaanvragen met een LHBTI- of bekeringsmotief. Wetenschappelijk Onderzoek- en Documentatiecentrum. Opgehaald van https://repository.wodc.nl/bitstream/handle/20.500.12832/935/Mem_2019-2_Volledige_tekst_tcm28-391710.pdf?sequence=2&isAllowed=y

Scholte, R., Nelen, W., Wit, W. de & Kroes, G. (2016). Sociale veiligheid in en rond scholen. Nijmegen: Praktikon

Short, E. (2018). It turns out Amazon's AI hiring tool discriminated against women. Opgehaald van Siliconrepublic: https://www.siliconrepublic.com/careers/amazon-ai-hiring-tool-women-discrimination

Simonite, T. (2018). AI Is the Future—But Where Are the Women? Opgehaald van https://www.wired.com/story/artificial-intelligence-researchers-gender-imbalance/

Simpson, E., Hamann, A., & Semaan, B. (2022). How to Tame "Your" Algorithm: LGBTQ+ Users' Domestication of TikTok. Proceedings of the ACM on Human-Computer Interaction. Opgehaald van https://dl.acm.org/doi/10.1145/3492841

Stapleton, L. Lesbian, Gay, Transgender, and Questioning (LGBTQ) Web Resources: Are Positive, Appropriate Websites Available Behind School System's Filters? https://cdr.lib.unc.edu/concern/masters_papers/vt150n695

Ștefăniță, O., & Buf, D.-M. (2021). Hate Speech in Social Media and Its Effects on the LGBT Community: A Review of the Current Research. Romanian Journal of Communication and Public Relations. Opgehaald van https://www.journalofcommunication.ro/index.php/journalofcommunication/article/view/322

Tesfai, Y. (2019). YouTube and the LGBT+ Community: Demonitizing Pride. Masters of Media. Opgehaald van https://mastersofmedia.hum.uva.nl/blog/2019/10/24/youtube-and-the-lgbt-community-demonitizing-pride/

The Register. (2018). US gov quizzes AI experts about when the machines will take over. Opgehaald van https://www.theregister.co.uk/2018/06/26/

UNESCO. (2019). Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development. Opgehaald van http://repositorio.minedu.gob.pe/handle/20.500.12799/6533

Urbi, J. (2018). Some transgender drivers are being kicked off Uber's app. Opgehaald van CNBC: https://ainowinstitute.org/discriminatingsystems.pdf

Van Beusekom, G. & L Kuyper (2018) LHBT-monitor 2018. De leefsituatie van lesbische, homoseksuele, biseksuele en transgender personen in Nederland. Opgehaald van: https://www.scp.nl/publicaties/monitors/2018/11/21/lhbt-monitor-2018

Van der Vorst, T., Jelicic, N., De Vries, M., & Albers, J. (2019). De (on)mogelijkheden van kunstmatige intelligentie in het onderwijs. Dialogic. Opgehaald van https://www.dialogic.nl/wp-content/uploads/2019/04/Dialogic-De-onmogelijkheden-van-kunstmatige-intelligentie-in-het-onderwijs-v1.0.116.pdf

Vetzo, M. J., Gerards, J. H., & Nehmelman, R. (2018). Algoritmes en grondrechten. Den Haag: Boom Juridisch. Opgehaald van https://www.uu.nl/sites/default/files/rebo-montaigne-algoritmes_en_grondrechten.pdf

West, S. M., Whittaker, M., & Crawford, K. (2019). Discriminating Systems, Gender, Race, and Power in AI. AI Now Institute. Opgehaald van https://ainowinstitute.org/discriminatingsystems.html

WOMEN Inc. (2021). AI, gender en de arbeidsmarkt. Ministerie van OCW. Opgehaald van https://205rpe43adla3g1ggs4flrux-wpengine.netdna-ssl.com/wp-content/uploads/2021/04/AI-Gender-en-de-Arbeidsmarkt.pdf

Zhao, e. a. (2019). Assessing mental health signals among sexual and gender minorities using Twitter data. Health Informatics Journal. Opgehaald van https://journals.sagepub.com/doi/full/10.1177/1460458219839621

Zuiderveen Borgesius, F. (2018). Discrimination, Artificial Intelligence and Algorithmic Decision-Making. Raad van Europa. Opgehaald van https://kennisopenbaarbestuur.nl/documenten/discrimination-artificial-intelligence-and-algorithmic-decision-making/