

Veelgestelde vragen over het rapport Xenofobe Machines – Discriminatie door ongereguleerd gebruik van algoritmen in het Nederlandse toelagenschandaal

AI Indexnummer: EUR 35/4686/2021

DISCRIMINATIE EN HET RISICOCCLASSIFICATIESYSTEEM

WAT LEGT HET RAPPORT BLOOT?

Het rapport legt bloot dat de Nederlandse overheid onvoldoende maatregelen heeft genomen om discriminatie door het gebruik van een algoritme in de toelagenaffaire te voorkomen of om herhaling ervan in de toekomst te voorkomen.

Het rapport van Amnesty International richt zich op een klein, maar belangrijk en onderbelicht aspect van het toelagenschandaal: het risicoclassificatie-model. Dat is een zelflerend algoritme dat wordt gebruikt bij de opsporing van onjuiste aanvragen van kinderopvangtoeslag en fraude.

Het toelagenschandaal omvat meerdere grote problemen, waaronder strenge regels en beleid, harde interpretaties van wetten en een meedogenloos terugvorderingsbeleid. Amnesty International richt zich maar op een klein onderdeel daarvan.

Het rapport van Amnesty International legt bloot dat ook sprake was van etnisch profileren door middel van het algoritme dat de Belastingdienst gebruikte om mensen te selecteren voor controles, het zogenaamde risicoclassificatie-model. Dat werkte als volgt.

In het risicoclassificatie-model gebruikte de Belastingdienst de nationaliteit van aanvragers. Mensen met een niet-Nederlandse nationaliteit kregen een hogere risicoscore. Het gebruik van nationaliteit in het algoritme geeft aan dat de ontwikkelaars en/of gebruikers van het systeem veronderstelden dat mensen van een bepaalde nationaliteit meer geneigd zouden zijn om fraude te plegen dan mensen van andere nationaliteiten. Het geeft ook aan dat de belastingdienst dacht dat er een verband bestaat tussen etniciteit en criminaliteit: nationaliteit werd gebruikt als aanwijzend element om bepaalde maatschappelijke groepen te identificeren, vanwege de veronderstelling dat deze groepen door hun gemeenschappelijke kenmerken meer geneigd zouden zijn tot fraude of crimineel gedrag. Deze behandeling van ouders en verzorgers met een niet-Nederlandse nationaliteit resulteerde in etnisch profileren.

HOE WERKT HET ALGORITME?

In 2013 voerde de Belastingdienst een algoritmisch systeem in voor de opsporing van fraude bij kinderopvangtoeslag. Dit systeem, dat bekend staat als het risicoclassificatie-model, leerde om risicoprofielen op te stellen van aanvragers van kinderopvangtoeslag. Het risicoclassificatie-model werd ontworpen door voorbeelden van goede en foute

aanvragen met elkaar te vergelijken. Het ontworpen model bestond uit meerdere risicofactoren, waaronder Nederlanderschap. Voor elke risicofactor kreeg de aanvrager een bepaald aantal punten, wat resulteerde in een totaalscore voor elke inkomende toeslagenvraag: de risicoscore. Ouders en verzorgers met een hoge risicoscore werden door ambtenaren onderzocht. De ambtenaar kreeg geen informatie over de reden voor de hoge risicoscore.

WAT BETEKENEN XENOFOBE MACHINES?

Met de naam 'xenofobe machines' doelt Amnesty International op algoritmische systemen of Kunstmatige Intelligentie die leiden tot discriminatie. Vaak wordt gedacht dat zulke systemen neutraal en objectief zijn, maar dat is niet zo. Zij worden gemaakt door mensen en kunnen de vooroordelen die deze mensen hebben, reproduceren of zelfs versterken. Wat algoritmes zijn en hoe zij kunnen discrimineren, leggen we uit in [deze video](#).

WAT IS ETNISCH PROFILEREN?

Etnisch profileren is het gebruik van ras, huidskleur of nationale of etnische herkomst door de politie of andere wetshandhavingsautoriteiten in plaats van gedrag als basis voor toezicht en opsporing of om te bepalen of iemand mogelijk betrokken is bij criminaliteit. Het is verboden. Etnisch profileren is niet altijd een uitgesproken praktijk. Het kan ook heimelijk gebeuren, bijvoorbeeld wanneer etnisch profileren geprogrammeerd is in algoritmische systemen. Etnisch profileren kan het gevolg zijn van een duidelijk discriminerende houding, maar het kan ook het gevolg zijn van onbewuste vooroordelen. Wat de omstandigheden ook zijn, etnisch profileren is altijd in strijd met het recht op om niet gediscrimineerd te worden. Het leidt tot criminalisering van bepaalde groepen, waardoor stereotypische associaties tussen bijvoorbeeld fraude en etniciteit worden versterkt. Etnisch profileren heeft schadelijke gevolgen omdat het bijdraagt aan de stigmatisering van groepen, wat een negatieve invloed heeft op de mentaliteit en het welzijn van de betrokken mensen. Dit soort ongelijke behandeling raakt aan de kern van menselijke waardigheid.

OP WELKE SCHAAL VINDT ETNISCH PROFILEREN PLAATS IN NEDERLAND?

Etnisch profileren is een veelbesproken kwestie in Nederland. Amnesty International bracht in 2013 twee rapporten uit over het probleem en de gevolgen van etnisch profileren door de Nederlandse politie: [Proactief politieoptreden vormt een risico voor mensenrechten](#) en [Gelijkheid onder druk](#). Over de schaal waarop etnisch profileren plaatsvindt, verschillen experts (enigszins) van mening. Uit Amnesty International's onderzoek blijkt dat etnische minderheden vaker onderworpen aan politiecontroles dan witte Nederlanders, terwijl daar geen objectieve reden voor is. In de rapporten staan aanbevelingen die de basis vormen voor Amnesty International's werk met betrekking tot etnisch profileren. De politie en het kabinet hebben maatregelen aangekondigd om etnisch profileren tegen te gaan. Amnesty International monitort de implementatie daarvan en is erover in gesprek met politie en autoriteiten.

In 2020 bracht Amnesty International een rapport uit over de mensenrechtenrisico's van 'predictive policing': [We sense trouble](#). In dit rapport legt Amnesty International uit hoe de Nederlandse politie gebruikmaakt van gegevens en risicomodellen om criminaliteit te 'voorspellen'. In een project in Roermond leidde dit tot geautomatiseerd etnisch profileren.

DE AP CONCLUDEERDE NIET ETNISCH PROFILEREN? HOE ZIT DAT NOU MET NATIONALITEIT EN ETNICITEIT?

Het rapport van Amnesty International is een verbreding en verdieping op het rapport van de Autoriteit persoonsgegevens.

De Autoriteit Persoonsgegevens (AP) concludeerde in [haar rapport](#) dat de Belastingdienst de *nationaliteit* van aanvragers van toeslagen onrechtmatig had verwerkt en dat deze verwerking discriminerend was op grond van nationaliteit. De Autoriteit Persoonsgegevens kon in haar onderzoek niet vaststellen of de Belastingdienst gegevens over *etniciteit* had verwerkt en stelde dat nationaliteit in het algemeen geen directe aanwijzing is voor etniciteit.

Amnesty International legt in het rapport uit dat nationaliteit in de praktijk een onderdeel kan zijn van etniciteit. Etniciteit is het idee dat maatschappelijke groepen kunnen worden onderscheiden op basis van hun gemeenschappelijke nationaliteit, geloofsovertuiging, taal, culturele en traditionele afkomst en achtergrond. Zo moet dat ook worden uitgelegd als je de internationale mensenrechtenverdragen volgt, wat Nederland zou moeten doen. In het risicoclassificatie-model van de Belastingdienst werd nationaliteit gebruikt als element om bepaalde maatschappelijke groepen te identificeren (niet-Nederlanders), vanwege de veronderstelling dat deze groepen vanwege hun gemeenschappelijke kenmerken meer geneigd zouden zijn tot fraude of crimineel gedrag. Dit leidde tot onderscheid op grond van etniciteit.

JULLIE HEBBEN HET OVER DISCRIMINATIE OP SOCIALE AFKOMST, WAT HOUDT DAT IN?

Sociale afkomst verwijst naar de status die iemand wordt toegekend op grond van een sociaal, cultureel of economisch kenmerk. Dit kan bijvoorbeeld betrekking hebben op armoede of dakloosheid. Mensenrechten verbieden discriminatie op grond van sociale afkomst. Mensen en groepen mogen niet ongelijk worden behandeld vanwege het behoren tot een bepaalde economische of sociale groep.

Wanneer een ouder of verzorger eenmaal geselecteerd was door het risicoclassificatie-model voor aanvullende controles, was de kans groter dat zij ook het slachtoffer zou worden van andere onbehoorlijke overheidsmaatregelen in verband met het toeslagenschandaal, zoals strenge regels en beleid, starre wetsinterpretaties, ongerechtvaardigde beschuldigingen van fraude en een meedogenloos terugvorderingsbeleid. Deze maatregelen hadden op verschillende manieren een onevenredig effect op mensen in lage-inkomensgroepen. Ten eerste zijn mensen met lage inkomens meer afhankelijk van financiële steun en toeslagen om rond te komen. Het verlagen of opschorten van toeslagen kan ouders en verzorgers uit deze groep onmiddellijk in financiële problemen brengen. Ten tweede, hoe lager het inkomen van iemand is, hoe hoger het totaalbedrag aan toeslagen dat diegene ontvangt. Het etiket

‘opzet & grove schuld’ werd vaker (automatisch) toegekend aan mensen uit huishoudens met een laag inkomen, simpelweg omdat zij meer dan 10.000 euro aan toeslagen ontvingen of meer dan 3.000 euro aan de Belastingdienst moesten terugbetalen. Ten derde hebben mensen met lage inkomens meestal weinig mogelijkheden om geld te reserveren om tijdelijke of permanente stopzetting van toeslagen te kunnen opvangen, laat staan dat zij de ruimte hebben om boetes te betalen of reeds ontvangen bedragen kunnen terugbetalen. Het terugbetalen van toeslagen en het betalen van boetes is voor deze groep vaak moeilijker.

WAT IS INTERSECTIONELE DISCRIMINATIE?

Sommige mensen worden geraakt door extra of unieke vormen van discriminatie vanwege meerdere kenmerken die onderdeel zijn van hun identiteit. Het gaat dan bijvoorbeeld om én sociale afkomst én etniciteit. Dit wordt intersectionele discriminatie genoemd. Het internationaal recht verplicht staten om maatregelen te nemen die alle vormen van discriminatie uitbannen. Het toeslagenschandaal moet door de lens van intersectionele discriminatie worden begrepen. Verschillende lagen van ongelijkheid werden veroorzaakt door discriminatie op etniciteit in het risicoclassificatie-model in combinatie met beleid en praktijk die vaker mensen met een lagere economische status troffen.

WITTE NEDERLANDERS ZIJN OOK SLACHTOFFER VAN DE TOESLAGENAFFAIRE, HOE ZIT DAT?

Dat klopt, ook witte Nederlanders waren slachtoffer van het toeslagenschandaal. Het rapport van Amnesty International geeft geen alomvattend overzicht van alle schade die verband houdt met het toeslagenschandaal. De volledige omvang van de gevolgen van het toeslagenschandaal voor de rechten van mensen is nog niet in kaart gebracht. Amnesty International beweert niet dat alle schade (rechtstreeks) verband houdt met de discriminerende praktijken en het discriminerende beleid die in het rapport worden beschreven. Amnesty International doet in het rapport aanbevelingen om alle mensenrechten te beschermen bij het gebruik van algoritmische systemen, onder meer door transparant te zijn over het gebruik van algoritmes door de overheid. Daarnaast roept Amnesty International de Nederlandse overheid op om, in aanvulling op de huidige compensatieregeling, compensatie te bieden voor discriminatie aan de ouders en verzorgers die mede op grond van hun niet-Nederlandse nationaliteit zijn geselecteerd voor controles.

AMNESTY INTERNATIONAL'S ONDERZOEK

HOE HEBBEN JULLIE HET ONDERZOEK GEDAAN?

Amnesty International analyseerde het algoritme aan de hand van parlementaire en ministeriële documenten die in de nasleep van het toeslagenschandaal zijn vrijgegeven, verslagen van parlementaire debatten en rapporten van toezichthouders. In het rapport

toetst Amnesty International het algoritmische systeem aan de hand van internationale mensenrechten, waaronder het recht om niet gediscrimineerd te worden.

Amnesty International heeft voorafgaand aan publicatie brieven gestuurd aan demissionaire minister-president Rutte, met daarin een beschrijving van het algoritme, het wettelijk kader en Amnesty International's mensenrechtenanalyse. De Nederlandse regering heeft op deze brieven gereageerd. Waar nodig heeft Amnesty International de reactie verwerkt in het rapport.

HOEVEEL MENSEN ZIJN SLACHTOFFER GEWORDEN VAN HET RISICOCCLASSIFICATIE-MODEL?

Dat weten we niet. Door het gebrek aan informatie over het risicoclassificatie-model weten mensen niet of zij op basis van het risicoclassificatie-model zijn geselecteerd voor controles. Het gebrek aan transparantie over algoritmische modellen is een groot probleem, omdat mensen zich hierdoor niet kunnen verweren tegen de uitkomsten. Amnesty International pleit voor een algoritme-register, waarin de overheid gedetailleerde en uitgebreide informatie bekendmaakt over de algoritmische systemen die zij gebruiken. Overheden moeten volledig transparant zijn en moeten informatie geven aan de mensen die geraakt worden door de algoritmische systemen.

HEBBEN JULLIE OOK SLACHTOFFERS GESPROKEN?

Nee, Amnesty International heeft geen slachtoffers geïnterviewd voor dit rapport. Veel media hebben ouders geïnterviewd over de manier waarop de strenge praktijken voor het opsporen van fraude hun leven hebben beïnvloed. In deze publicatie zijn dergelijke interviews niet opgenomen, noch zijn de dossiers van individuele ouders en verzorgers geanalyseerd om na te gaan op welke manieren zij zijn benadeeld en hoe zij daarvoor moeten worden gecompenseerd. Ouders en verzorgers en hun advocaten werden onwetend gehouden over het gebruik van algoritmen, aangezien er in hun dossiers geen informatie was over hoe en waarom zij door het systeem werden geselecteerd. Deze publicatie bespreekt de werking van het discriminerende algoritmische systeem dat werd gebruikt voor het opsporen van onjuiste en mogelijk frauduleuze aanvragen en doet aanbevelingen om soortgelijke mensenrechtenschendingen in de toekomst te voorkomen. Deze publicatie is bedoeld als leidraad voor beleidsmakers en praktijkmensen over de hele wereld bij het nemen van stappen om maatregelen en waarborgen in te voeren ter bescherming van de mensenrechten bij het gebruik van impactvolle algoritmische systemen door overheden.

WAT HEBBEN DE SLACHTOFFERS AAN DIT RAPPORT?

Amnesty International hoopt allereerst erkenning te bieden aan die slachtoffers die op basis van hun nationaliteit een hogere risicoscore waren toebedeeld en zo in de draaikolk van de fraudeaanpak terecht zijn gekomen. Amnesty International vraagt internationaal en nationaal aandacht voor het feit dat discriminatie al plaats vond in de selectie van het risicoclassificatie-model, en niet enkel in de latere behandeling. Daarnaast eist Amnesty International betere regels voor algoritmische systemen, zodat er in de toekomst geen nieuwe slachtoffers komen door discriminerende algoritmes. Overheden moeten waarborgen treffen zodat discriminatie in de toekomst wordt

voorkomen. Het rapport van Amnesty International is gericht op schade in de toekomst voorkomen.

IS DE GEAUTOMATISEERDE RISICOPROFILERING DE BELANGRIJKSTE OORZAAK VAN DE TOESLAGENAFFAIRE?

De toeslagenaffaire komt door een samenspel van onbehoorlijk overheidsoptreden waar door de overheid fout op fout lijkt gemaakt te zijn. Al die fouten bij elkaar en het feit dat die zo lang hebben kunnen doorgaan, zijn de oorzaak van de toeslagenaffaire. Risicoprofilering is een van die fouten. Het rapport van Amnesty International geeft geen alomvattend overzicht van alle schade die verband houdt met het toeslagenschandaal. De volledige omvang van de gevolgen van het toeslagenschandaal voor de rechten van mensen is nog niet in kaart gebracht. Amnesty International beweert niet dat alle schade (rechtstreeks) verband houdt met de discriminerende praktijken en het discriminerende beleid die in het rapport worden beschreven. Amnesty International doet in het rapport aanbevelingen om alle mensenrechten te beschermen bij het gebruik van algoritmische systemen om herhaling van de fout met het risicoprofileringssysteem te voorkomen. Daarnaast roept Amnesty International de Nederlandse overheid op om, in aanvulling op de huidige compensatieregeling, compensatie te bieden voor discriminatie aan de ouders en verzorgers die mede op grond van hun niet-Nederlandse nationaliteit zijn geselecteerd voor controles.

WAAROM EEN INTERNATIONAAL RAPPORT OVER IETS NATIONAALS?

Nederland is niet het enige land dat gebruikmaakt van geautomatiseerde processen in de hoop fraude of criminaliteit op te sporen. Wereldwijd maken steeds meer landen gebruik van gegevens en algoritmes voor het opsporen van fraude. Amnesty International maakt zich grote zorgen over de snelheid waarmee deze algoritmische systemen worden uitgerold, terwijl waarborgen om mensenrechten te beschermen achterblijven. Het toeslagenschandaal laat zien hoe het gebruik van risicoprofielen in fraudebestrijding kan leiden tot schendingen van mensenrechten. Zowel op internationaal, Europees als nationaal niveau schiet de bescherming van mensenrechten bij het gebruik van algoritmes tekort. Daarom presenteerde Amnesty International het rapport ook in een nevenbijeenkomst bij de Verenigde Naties tijdens de General Assembly in New York op 26 oktober 2021. Amnesty International roept alle staten op om de aanbevelingen uit het rapport in wetgeving om te zetten. Amnesty International zal de komende jaren meer onderzoek doen en actie gaan voeren om een einde te maken aan mensenrechtenschendingen met algoritmische systemen.

KOMT DIT IN ANDERE LANDEN OOK VOOR?

Ja, zeker. Maar het toeslagenschandaal is uniek, omdat het algoritme enigszins gedocumenteerd is door de Autoriteit Persoonsgegevens en de Parlementaire Onderzoekingscommissie Kinderopvangtoeslag. Het is voor slachtoffers van andere algoritmische systemen vaak onmogelijk om te achterhalen hoe het betreffende systeem werkte. In veel gevallen zullen zij niet eens weten dat zij geselecteerd zijn op basis van een risicosysteem dat hen discrimineert. Mensenrechtenschendingen van andere projecten, in andere landen of in Nederland, zijn vaak onmogelijk te achterhalen.

Daarom eist Amnesty International meer transparantie over het gebruik van algoritmische systemen.

DE OPLOSSING VOOR HET PROBLEEM

WAT IS DE OPLOSSING?

Alle staten, inclusief Nederland, moeten mensenrechten beschermen bij het gebruik van algoritmische systemen. Regeringen moeten regels opstellen zodat mensenrechtenschendingen bij het gebruik van algoritmische systemen worden voorkomen. Zo moeten zij transparant zijn over het gebruik van algoritmes en een mensenrechtentoets uitvoeren voordat algoritmische systemen worden gebruikt. Daarnaast moeten regeringen monitoring- en toezichtmechanismen instellen die controleren of de regels worden nageleefd. Degenen die verantwoordelijk zijn voor schendingen moeten ter verantwoording worden geroepen en er moeten effectieve rechtsmiddelen zijn voor de mensen en groepen van wie rechten zijn geschonden.

WIE MOET DIT OPLOSSEN?

In de eerste plaats overheden, dus de Nederlandse regering in Nederland. Het kan ook dat de nationale overheid ervoor kiest om dit met EU-wetgeving aan te pakken. De Nederlandse overheid moet zich dan inzetten om ervoor te zorgen dat de aanbevelingen voor mensenrechtenbeschermingen in die EU-wetgeving terechtkomen.

WORDT DIT NIET AL GEREGLD IN DE HET WETSVOORSTEL VAN DE EUROPESE UNIE OVER ALGORITMISCHE SYSTEMEN, DE ARTIFICIAL INTELLIGENCE ACT?

Nee. De Artificial Intelligence Act (AIA) regelt bijvoorbeeld een vorm van toezicht en een vorm van transparantie, maar geen mensenrechtentoezicht en mensenrechtentransparantie. Zo stelt de AIA voor dat de ontwikkelaar transparant moet zijn naar de gebruiker van een systeem maar niet dat de gebruiker, bijvoorbeeld de Belastingdienst, ook transparant is naar mensen, zoals de slachtoffers van toekomstige toeslagenaffaires. Die laatste vorm van transparantie is wel vereist om mensenrechten te beschermen. Amnesty International roept de Europese wetgever op om die bescherming op te nemen in de AIA en roept de Nederlandse wetgever op om zich binnen de EU hiervoor in te zetten. Zie bijvoorbeeld [Amnesty International's inzending bij de Europese Commissie hierover](#).