

Probleemstelling

Risicoprofilering heronderzoek bijstandsuitkering

Samenvatting probleemstelling

Deze probleemstelling beschrijft ethische kwesties met betrekking tot handmatige en algoritmische profileringsmethoden die door Nederlandse gemeenten worden gebruikt om burgers die een bijstandsuitkering ontvangen te selecteren voor heronderzoek. Informatie over deze methoden is op onafhankelijke wijze verzameld en geanalyseerd door stichting Algorithm Audit. In dit document is feedback meegenomen van deskundigen met verschillende professionele achtergronden. Het startpunt van deze casus is een wob-verzoek (een verzoek tot openbaarmaking van overheidsinformatie) van journalisten, waarmee informatie over algoritmische risicoprofilering is verkregen, zoals broncode, modeldocumentatie, privacy impact assessment en geaggregeerde prestatiestatistieken. Gaandeweg is aanvullende juridische, statistische en context-gerelateerde informatie toegevoegd, onder andere door de gemeente Rotterdam naar aanleiding van specifieke vragen van Algorithm Audit. Het risicotaxatie-algoritme dat wordt besproken in deze probleemstelling wordt niet langer door de gemeente Rotterdam gebruikt. Desalniettemin blijven de onderliggende ethische kwesties relevant voor het bredere debat over de inzet van risicoprofilering door Nederlandse gemeenten. Aan de hand van dit document heeft een onafhankelijke auditcommissie advies uitgebracht over drie geïdentificeerde ethische kwesties. Dit advies kan worden gevonden in bijbehorend Auditcommissie Adviesdocument (AA:2023:02:A).

Inhoudsopgave

Introductie	3
Focus van case study	6
Juridische achtergrond van heronderzoek bijstandsuitkering	8
Het proces van heronderzoek en het concept fraude	9
Kwestie I: Uitdagingen rondom proxy-discriminatie en correlaties	10
Kwestie II: Ethisch (on)wenselijke criteria voor risicoprofilering	13
Kwestie III: Vergelijking van expert en algoritmische profilering	14
Appendix A – Juridische achtergrond van heronderzoek bijstandsuitkering	17
Appendix B – Dataverzameling	20
Appendix C – Variabelenselectie methoden	23
Appendix D – Resultaten per selectiemethode	25
Structurele partners van Algorithm Audit	30

Probleemstelling – Risicoprofilering voor heronderzoek bijstandsuitkering

Dit document beschrijft ethische vragen met betrekking tot risicogerichte profileringsmethoden om bijstandsgerechtigden te selecteren voor heronderzoek. Werkelijk toegepaste steekproefmethoden door de gemeente Rotterdam worden in deze casestudie besproken.

Introductie

Niet alle bijstandsuitkeringen worden rechtmatig toegekend. Onbekende verantwoordelijkheden, administratieve fouten, verwijtbare nalatigheid, en opzettelijke fraude leiden tot onrechtmatige uitkeringen. Ongeacht de achterliggende reden hebben Nederlandse gemeenten een wettelijke plicht om onrechtmatige bijstandsuitkeringen op te sporen en terug te vorderen. Nederlandse gemeenten voeren daarom regelmatig heronderzoek en uit bij bijstandsontvangers om na te gaan of bijstandsuitkeringen rechtmatig worden uitgekeerd. Deze heronderzoeken vinden plaats in de vorm van gesprekken en menselijke controles van verstrekte documenten. Het doel van deze heronderzoeken is niet alleen het opsporen van fraude, maar ook het opsporen van onbedoelde administratieve fouten en andere vergissingen die tot onrechtmatige uitkeringen kunnen leiden, bijvoorbeeld het bijwerken van achterhaalde informatie.

Bij selectie voor een heronderzoek worden bijstandsontvangers opgeroepen voor een gesprek met een medewerker van de gemeente om na te gaan of de informatie die de gemeente over de ontvanger heeft actueel is. De gemeente ziet een heronderzoeksgesprek als een natuurlijk contactmoment met burgers zonder voorafgaande verdenking. Desalniettemin beschouwen de ondervraagden het heronderzoeksgesprek en de periode

Over Algorithm Audit

Algorithm Audit is een Europees kennisplatform voor het testen van AI-bias en normatieve AI-standaarden. De activiteiten zijn drieledig:



Audit commissies

Adviseren over ethische kwesties in concrete algoritmische toepassingen door deliberatie, met *algotrudentie* als resultaat (zie Box 2)



Technische hulpmiddelen

Implementeren en testen van technische methoden voor detectie en mitigatie van vooringenomen data en algoritmes



Kennisplatform

Samenbrengen van kennis en experts om het collectieve leerproces over de verantwoorde inzet van algoritmes aan te jagen

die eraan voorafgaat als invasief, tijdrovend en stressvol, onder meer vanwege administratieve handelingen, zoals het overhandigen van identiteitsdocumenten of diploma's, of het leveren van een overzicht van persoonlijke uitgaven aan ambtenaren. In **Box 1** is een getuigenis ingevoegd van een inwoner van Rotterdam die een bijstandsuitkering ontvangt en is uitgenodigd voor heronderzoek. Van genodigden wordt verwacht binnen twee weken een afspraak te plannen met de gemeente na ontvangst van de uitnodigingsbrief.

Aangezien het onmogelijk is om alle bijstandontvangers te heronderzoeken, worden steekproefmethoden gebruikt om bijstandontvangers voor heronderzoek te selecteren. Om ambtelijke middelen efficiënt in te zetten en ontvangers van onrechtmatige uitkeringen in een vroeg stadium te helpen, voeren gemeenten risicogerichte steekproeven uit om bijstandontvangers voor een heronderzoek te selecteren. Het doel van risicogerichte steekproeven is om bijstandontvangers uit te nodigen waarbij een verhoogde kans op onrechtmatigheid wordt vermoed. Er bestaan verschillende methoden voor risicogerichte steekproeven. Zo kunnen ambtenaren die werkzaam zijn als persoonlijk klantmanager of inkomensconsulent bijstandontvangers op de heronderzoeklijst plaatsen als interactie verdenking wekt. Inwoners kunnen ook contact opnemen met de gemeente als zij anderen verdenken van frauduleus gedrag, wat door een medewerker nader wordt onderzocht¹. Dit type selectiemethode wordt een *signaal steekproef* genoemd.

Wanneer geen signaal voorhanden is, worden diverse andere risicogerichte steekproefmethoden gebruikt. Zo worden bijstandontvangers bijvoorbeeld geselecteerd op basis van risicoprofielen. Deze risicoprofielen worden opgesteld volgens een reeks vooraf vastgestelde selectiecriteria, die geacht worden de kenmerken weer te geven van bijstandontvangers met een gemiddeld hoger risico op onrechtmatige uitkeringen. De criteria voor dergelijke risicoprofielen worden op verschillende manieren afgeleid. Op basis van beroepservaring wordt inhoudelijke deskundigheid van ambtenaren gebruikt om handmatig criteria voor risicoprofielen vast te stellen, bijvoorbeeld mannen die alleen wonen². Deze methode wordt *expert profilering* genoemd. Een thematisch profiel samengesteld door experts bestaat doorgaans uit 1-3 criteria die handmatig worden geselecteerd en meestal jaarlijks worden gewijzigd om te voorkomen dat te veel nadruk op een specifieke

¹ Bij de kliklijn van de gemeente Rotterdam kan verdacht gedrag gemeld worden <https://www.rotterdam.nl/loket/fraude-uitkering-doorgeven/>

² Geleerde Technologie, Rotterdamse Rekenkamer 2021 <https://rekenkamer.rotterdam.nl/onderzoeken/algoritmes/>

Omvang van bijstandsherziening

In de periode 2017-2021 was voor 38% van de circa 22.000 heronderzochte Rotterdamse bijstandontvangers actie nodig om de uitkering rechtmatig te maken, d.w.z. aanpassing, beëindiging of terugvordering. In 2019 rapporteerde de gemeente Amsterdam en gemeente Den Haag in totaal €4.9m en €5.5m aan onrechtmatige uitgekeerde bijstandsuitkeringen. Voor de gemeente Rotterdam zijn deze getallen niet bekend. Meer informatie over aggregatiestatistieken is beschikbaar in Appendix D – Resultaten per selectiemethode.

groep wordt gelegd. Meer details over het werkproces van expert profilering in Rotterdam is te vinden in [Appendix C – Variabelenselectie methoden](#).

Een andere methode om risicoprofielen op te stellen is **algoritmische profilering**. Bij deze aanpak wordt een grote hoeveelheid historische data over kenmerken van bijstandsontvangers en de resultaten van uitgevoerde heronderzoeken gebruikt om op geautomatiseerde wijze risicoprofielen te genereren. Het model wordt getraind op de doelvariabele 'uitkeringsrelatie is onrechtmatig' (meer details in de sectie [Juridische achtergrond van heronderzoek bijstandsuitkering](#)). Zodra een risicoprofiel is vastgesteld, door expert of algoritmische profilering, wordt het risicotaxatie-algoritme toegepast als een selectiefilter op alle bijstandsontvangers.

Ten slotte worden ontvangers van bijstand willekeurig geselecteerd via een *aselecte steekproef*, zonder selectiecriteria of risicoprofielen. Deze verschillende steekproefmethoden worden door Nederlandse gemeenten gebruikt om bijstandsontvangers te selecteren voor heronderzoek. Een weergave van deze steekproefmethoden wordt gegeven in [Figuur 1](#). Het aantal heronderzoeken per steekproefmethode verschilt per jaar en per gemeente. Meer informatie over steekproefverhoudingen in de gemeente Rotterdam in de periode 2017-2021 kan worden gevonden in [Appendix D – Resultaten per selectiemethode](#).

Box 1

Betrokkenen bij dit onderzoek

Voor dit onderzoek zijn de volgende belanghebbenden gehoord:



Individueel onderworpen aan algoritme



Onderzoekersjournalisten



Belangenvertegenwoordigers van benadeelden



Gemeentelijke instanties (Rotterdamse Rekenkamer en Ombudsman)



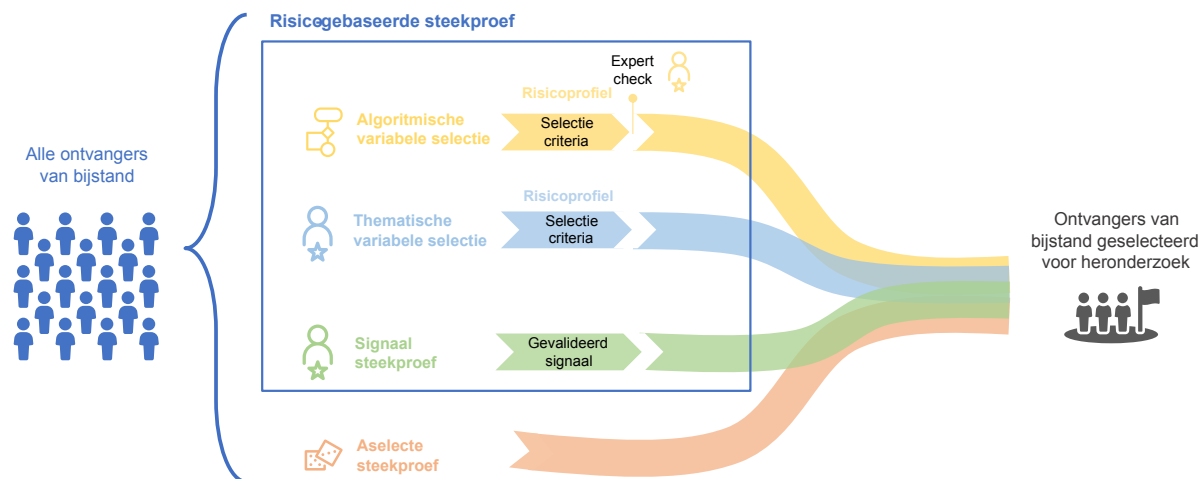
Gemeente Rotterdam



Juridische experts en wetenschappers

“Het is stressvol om uitgenodigd te worden voor een interview. In een onaangekondigde uitnodigingsbrief staat vermeld, dat als je niet komt opdagen, de uitkering automatisch wordt stopgezet. Dat aan mij, een geboren en getogen Nederlander, werd gevraagd om te bewijzen dat ik de Nederlandse taal beheers, was vernederend.”

Bijstandsgerechtigde uit Rotterdam (geanonimiseerd)



Figuur 1 – Steekproefmethoden om ontvangers van bijstand te selecteren voor heronderzoek

Focus van case study

In deze case study worden de ethische bezwaren van expert en algoritmische profilering besproken, met betrekking tot hun variabelenselectie en risicoprofileringsmethode. Een eerste ethisch bezwaar betreft vooringenomenheid van historische data waaruit selectiecriteria voor risicoprofielen worden gedistilleerd. Een tweede ethisch bezwaar relateert aan het selectieproces van variabelen, handmatig uitgevoerd door experts tijdens het vaststellen van een thema of automatisch uitgevoerd door een algoritme. Voor redenen hieronder toegelicht, focust deze case study zich niet op het eerste maar op het laatste ethische bezwaar.

Door middel van een wob-verzoek³ heeft een consortium van journalisten toegang verkregen tot de datasets die de gemeente Rotterdam heeft gebruikt voor het trainen van het risicotaxatie-algoritme. Dit onderzoek toont aan dat deze dataset vooringenomen is

³ Wet openbaarheid van bestuur (Wob) verzoek van VPRO Argos/Lighthouse Reports, 2017020 Privacy Impact Assessment pilotfase Project Uitkeringsonrechtmatigheid <https://www.vpro.nl/dam/jcr:c87f2d6c-3f9c-4498-9a9c-f3bc5483a437/Downloads%20Model%20Rotterdam.zip>

Box 2

Verloop van Algorithm Audit's case studies



Probleemstelling

- > Vaststellen ethische kwestie(s)
- > Verzamelen van relevante informatie
- > Feedback verwerken van experts



Auditcommissie

- > Samenstellen diverse auditcommissie
- > Initiële schriftelijke reactie op Probleemstelling
- > Deliberatief discussie over ethische kwesties



Algoprudentie

- > Opstellen normatief advies n.a.v. commissiebijeenkomst
- > Instemming van alle auditcommissieleden
- > Publicatie van algoprudentie

met betrekking tot leeftijd en geslacht van bijstandsontvangers⁴. Het vinden van vooringenomen datasets is een belangrijk wapenfeit om de scheve statistiek te voorkomen. Maar een vervuilde dataset staat los van *kwalitatieve* bezwaren die betrekking hebben op risicoprofilering en bijbehorende selectie van variabelen, denk bijvoorbeeld aan expert profilering waarbij het kiezen van thema's niet door data of data-analyse wordt geïnformeerd. En: in het geval een dataset als onvervuild wordt beschouwd, zullen kwalitatieve vragen over het selecteren van criteria voor risicoprofielen zich blijven voordoen. Deze kwalitatieve vragen zijn de kern van deze case study. Een andere reden om de datakwaliteit links te laten liggen is de context-afhankelijkheid van de dataset in kwestie. Kwalitatieve uitspraken over de datakwaliteit van de ene gemeente zijn lastig te generaliseren naar de dataset van een andere gemeente. Daarnaast heeft de gemeente Rotterdam een interne ethische commissie ingesteld, dat mandaat heeft om de kwaliteit van de data te beoordelen, bijvoorbeeld op mogelijke vooringenomenheid van historische data.

Alvorens de onderzoeksvragen van deze probleemstelling te introduceren, is het belangrijk de huidige status van het risicotaxatie-model van de gemeente Rotterdam te duiden. Gebruik van het risicotaxatie-algoritme is in april 2021 door de wethouder stilgelegd naar aanleiding van vragen uit de gemeenteraad over het rapport *Gekleurde Technologie* van de Rotterdamse Rekenkamer⁵. Nadien vindt selectie voor heronderzoek alleen nog plaats door expert profilering, aselecte- en signaalsteekproeven. De gemeente Rotterdam heeft de afgelopen jaren een nieuw algoritme ontwikkeld voor de selectie van ontvangers van bijstand voor heronderzoek dat aan hogere standaarden dan het in vorige risicotaxatie-algoritme voldoet. De status van dit nieuwe algoritme is onbekend.

⁴ Zie <https://www.lighthousereports.com/suspicion-machines-methodology/>

⁵ Zie <https://www.ftm.nl/artikelen/algoritme-gemeente-rotterdam>

Algo prudentie: Publieke kennisopbouw voor ethische algoritmes

Algorithm Audit heeft geen mandaat om juridische uitspraken of officiële uitspraken te doen. In onze case studies geven wij niet-bindend ethisch advies. Ethisch advies gaat vaak verder dan dat wat nodig is om de wet na te leven. Toch dient, vanwege de ambiguïteit van wettelijke regelgeving en bij gebrek aan heldere normen van toezichthouders, ons onafhankelijke ethische advies als een wegwijzer voor organisaties. Onze algo prudentie kan ook helpen bij het uitwerken van officiële standaarden of het ondersteunen van toekomstige besluiten van rechtsorganen. In dat opzicht is ons ethisch advies relevant binnen het juridische domein.



Ethische kwesties

We introduceren nu de onderzoeksvragen van deze case study. Een eerste focus van deze casestudie is de kwestie van proxy-discriminatie bij het opstellen van risicoprofielen. Het gebruik van ogenschijnlijk neutrale kenmerken van burgers (zoals opleidingsniveau, postcode of taalvaardigheid) als selectiecriteria voor risicoprofielen kan discriminerende uitkomsten opleveren als gevolg van hun correlatie met beschermde kenmerken, zoals etniciteit. De vraag is dan, rekening houdend met hun potentieel proxy-karakter, welke kenmerken legitiem kunnen worden gebruikt, en welke moeten worden uitgesloten.

Een tweede ethisch bezwaar betreft variabelen waarvan het gebruik moreel dubieus of problematisch is. Zoals hieronder wordt beschreven, is een deel van de persoonsgegevens die voor profilering wordt gebruikt, gebaseerd op subjectieve beoordelingen door ambtenaren, of lijken het persoonlijke kenmerken te zijn die losstaan van selectie voor heronderzoek. Voorbeelden zijn de verzorgdheid en assertiviteit van ontvangers van bijstand tijdens een persoonlijk interview. Met betrekking tot het gebruik van deze criteria in profileringsmethoden kan worden betwijfeld of de criteria objectief, proportioneel en noodzakelijk zijn met betrekking tot het nagestreefde doel.

Een derde kwestie betreft het verschil tussen expert en algoritmische profilering. Verhoogt of verlaagt expert profilering de ethische risico's ten opzichte van een algoritmische profilering? Zijn de ethische risico's van het profileren van ontvangers van bijstand vooral gelegen in het gebruik van machine learning technieken voor variabelenselectie – of liggen de risico's besloten in het opstellen van risicoprofielen zelf? Wat zou betekenen dat het handmatig opstellen van profielen door experts dezelfde risico's kent als algoritmische samengestelde profielen.

De focus van deze casus is dus driedig. Wij stellen de volgende ethische kwesties vast die deze case study richting geven:

Kwestie I: Welke kenmerken van bijstandsontvangers kunnen worden beschouwd als een proxy-variabele voor beschermde gronden (zoals gedefinieerd in Artikel 14 van de Europees Verdrag voor de Rechten van de Mens) en welke van deze variabelen moeten worden uitgesloten van profileringsmethoden om discriminerende vooringenomenheid te mitigeren?

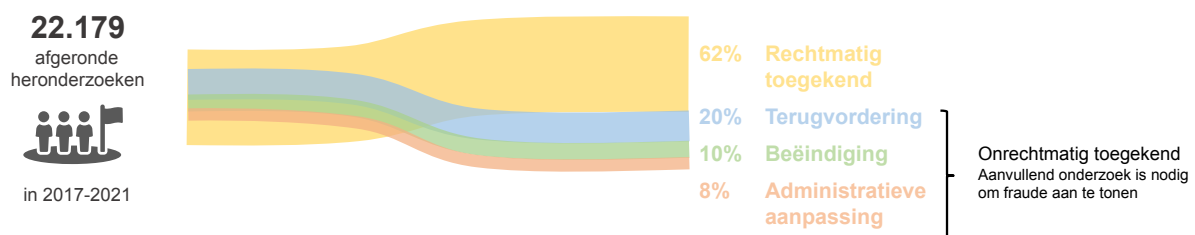
Kwestie II: Welke kenmerken zijn om andere redenen dan discriminerende vooringenomenheid ethisch onwenselijk om te gebruiken in profileringsmethoden?

Kwestie III: Onder welke voorwaarden is het wenselijk om ontvangers van bijstand voor heronderzoek te selecteren door middel van algoritmische profilering in plaats van expert profilering?

In de komende secties worden de genoemde ethische kwesties nader uitgewerkt. Daarnaast worden details over de juridische achtergrond, beschikbare selectiecriteria en specifieke selectiemethoden besproken.

Juridische achtergrond van heronderzoek bijstandsuitkering

Voordat selectiemethoden in detail geëvalueerd worden, is het relevant om uit te diepen wat de juridische basis vormt voor risicogerichte profileringsmethoden voor bijstandsheronderzoek. Zoals beschreven door het Nederlands College voor de Rechten van de Mens in het rapport *Discriminatie door Risicoprofielen*⁶, is het vaste rechtspraak door Nederlandse en Europese rechtbanken dat overheidsinstanties bij opsporings- en handhavingsbevoegdheden, gebruik mogen maken van risicoprofielen en historische data⁷. Desalniettemin moeten de methoden voor het opstellen van profielen, indien zij worden gebruikt voor risicotaxatie, voldoen aan onder andere het recht op non-discriminatie, gelijke behandeling en gegevensbescherming. Een van de manieren waarop deze grondrechten worden gewaarborgd, is het gelijk behandelen van verschillend geselecteerde bijstandsontvangers, d.w.z. dat ambtenaren die heronderzoeksgesprekken voeren niet weten via welke methode (signaal-gedreven, expert, algoritmisch of aselect) de ondervraagde is geselecteerd. Toch kunnen risicoprofielen in strijd zijn met het discriminatieverbod, aangezien criteria gelinkt kunnen zijn aan beschermde gronden, zoals etniciteit of nationaliteit. Volgens de Algemene wet gelijke behandeling (AWGB) moeten criteria voor profilering geschikt, proportioneel en noodzakelijk zijn om het doel te bereiken. De beoordeling van deze bepalingen is echter normatief en contextafhankelijk. Met oog op de afwezigheid van jurisprudentie over het gebruik van algoritmische profilering door Nederlandse gemeenten, kunnen we ons voor deze casestudie (nog) niet baseren op duidelijke gerechtelijke uitspraken⁸. Normatieve duiding van (algoritmische) variabelenselectie voor profilering is echter dringend nodig. Dat is het doel van deze casestudie.



Figuur 2 – Uitkomsten van heronderzoeken in de periode 2017-2021

⁶ Discriminatie door Risicoprofielen, College van de Rechten van de Mens, 2021 <https://open.overheid.nl/documenten/ronl-c409ea31-2c00-4318-9a45-d47ad8a2ca7f/pdf>

⁷ Centrale Raad van Beroep (CRvB) Jun-5 2018, ECLI:NL:CRVB:2018:1541, r.o. 4.2; CRvB Sep-20 2016, ECLI:NL:CRVB:2016:4160, r.o. 4.4.5; CRvB Apr-14 2015, ECLI:NL:CRVB:2015:3249, r.o. 4.5

⁸ De Systeem Risico Indicatie (SyRI) uitspraak is geen inhoudelijk oordeel over de inhoudelijke werking van de algoritmische methode. De broncode van het algoritme is nooit met de rechter gedeeld. SyRI is verboden op grond van een 'vangnet constructie', zijnde een te grote inbreuk op priv sfeer (art. 8 EVRM)

De Participatiewet garandeert een minimuminkomen voor iedereen die volgens de wet in Nederland woont en onvoldoende middelen heeft om in levensonderhoud te voorzien. Als tegenprestatie formuleert de wet bepaalde verplichtingen voor bijstandsgerechtigden, zoals deelname aan een 'traject tot werk'. Meer details over de wettelijke basis die Nederlandse gemeenten verplicht om heronderzoeken uit te voeren in het kader van de Wet werk en bijstand uit te voeren, zijn te vinden in van [Appendix A – Juridische achtergrond van heronderzoek bijstandsuitkering](#). In deze appendix worden ook verplichtingen voor gemeenten onder de Algemene wet bestuursrecht (Awb) en de Algemene verordening gegevensbescherming (AVG) beschreven.

Het proces van heronderzoek en het concept fraude

De uitkomst van een heronderzoek is binair: de bijstandsuitkering wordt ofwel rechtmatig ofwel onrechtmatig toegekend. Onrechtmatig toegekende uitkeringen kunnen worden ontleed in drie categorieën: 1) terugvordering, 2) beëindiging en 3) administratieve aanpassingen. Een overzicht van de uitkomsten van heronderzoeksgesprekken afgerond in de periode 2017-2019 is weergegeven in [Figuur 2](#). Meer informatie over de hoeveelheid afgeronde heronderzoeken per selectiemethode voor deze periode is te vinden in [Appendix D – Resultaten per selectiemethode](#).

Op basis van de uitkomst van een heronderzoeksgesprek kan de gemeente besluiten een vervolgonderzoek in te stellen naar "opzettelijke misleiding" door de bijstandsontvanger "om een onrechtmatig voordeel te behalen", dat wil zeggen: het vaststellen van fraude⁹. Dit proces, het onderzoeken van fraude nadat het heronderzoeksgesprek heeft plaats gevonden, valt buiten de reikwijdte van deze case study. Het is belangrijk op te merken dat als gevolg van zekere ruimte binnen het juridische kader de inzet om onrechtmatige bijstandsontvangers te identificeren verschilt per Nederlandse gemeente. Sommige Nederlandse gemeenten treden "agressiever" op tegen mogelijk onrechtmatige verstrekte bijstand, terwijl andere gemeenten coulanter zijn.

Kwestie I: Uitdagingen rondom proxy-discriminatie en correlaties

Over algoritmische risicotaxatie in de context van bijstandsheronderzoek heeft de Rotterdamse Rekenkamer in het rapport *Gekleurde Technologie* gewezen op een link tussen laaggeletterdheid en etniciteit¹⁰. De Rekenkamer waarschuwde dat het opnemen van laaggeletterdheid als een selectiecriteria voor profilering een risico tot discriminerende vooringenomenheid met zich meebrengt, aangezien laaggeletterdheid fungeert als een proxy-variabele voor de beschermde eigenschap 'afkomst'. In het rapport ontbreekt echter een interpretatie of deze mogelijke proxy variabele geschikt, proportioneel en noodzakelijk is om te gebruiken in het risicoprofiel. Dit is problematisch, aangezien voor Europese en Nederlandse non-discriminatie wetgeving aanvullende bewijsvereisten nodig

⁹ Artikel 225, 227b en/of 326 van het Wetboek van Strafrecht

¹⁰ Gekleurde technologie, Rotterdamse Rekenkamer 2021 <https://rekenkamer.rotterdam.nl/onderzoeken/algoritmes/>

zijn om indirecte (proxy) discriminatie hard te maken¹¹. Volgens de redenering van de Rotterdamse Rekenkamer dragen *alle* gebruikte variabelen het risico een proxyvariabele te zijn, aangezien iedere variabele tot zekere hoogte (per definitie) gecorreleerd is met een beschermde grond. Door correlatie als criterium te nemen voor indirecte (proxy) discriminatie, zou geen enkele variabele rechtmatig gebruikt kunnen worden en zou het gebruik van profileringsmethoden als zodanig ernstig, zo niet volledig, worden beperkt. Dit lijkt niet de bedoeling te zijn van de Rotterdamse Rekenkamer, maar de gepresenteerde redenering is onvoldoende om een alternatief te bieden.

In het rapport van de Rotterdamse Rekenkamer ontbreekt een duidelijk kwalitatief argument waarom, en wanneer, de kwantitatieve vraag van correlatie met beschermde kenmerken onaanvaardbaar wordt vanuit het oogpunt van non-discriminatie. Voor sommige variabelen kan het risico dat zij als proxy fungeren voor beschermde kenmerken meer worden getolereerd dan voor anderen. Omdat het directe belang van sommige variabelen (bijvoorbeeld de samenstelling van het huishouden) voor de beoordeling van onrechtmatige uitkeringen anders weegt in een discriminatietoets dan dat van andere variabele (bijvoorbeeld laaggeletterdheid). Bovendien is het juridische criterium voor indirecte discriminatie – dat een variabele "kritisch verband" houdt met een beschermde grond – voor kwalitatieve interpretatie vatbaar¹¹. Hetzelfde geldt voor de eis dat de variabele geschikt, proportioneel en noodzakelijk moet zijn om het beoogde doel te bereiken. Kortom, een kwalitatieve evaluatie is nodig om het discriminerende karakter van een profileringsmethode en de gehanteerde selectiecriteria adequaat te kunnen vaststellen. Het doel van deze casestudie is een kwalitatieve evaluatie van het gebruik van proxy-variabelen: Welke variabelen zijn wel, en welke variabelen zijn niet gerechtvaardigd om te worden opgenomen in algoritmische en expert profilering in het kader van risicotaxatie voor bijstandsheronderzoek?

Nederlandse gemeenten hebben toegang tot een groot aantal datapunten over ontvangers van een bijstanduitkering. Voor deze case study worden de beschikbare selectiecriteria voor de gemeente Rotterdam gedeeld. [Appendix B – Dataverzameling](#) geeft de

¹¹ S. Wachter, B. Mittelstadt, C. Russell, Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI p.15 (2020). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3547922

Box 3

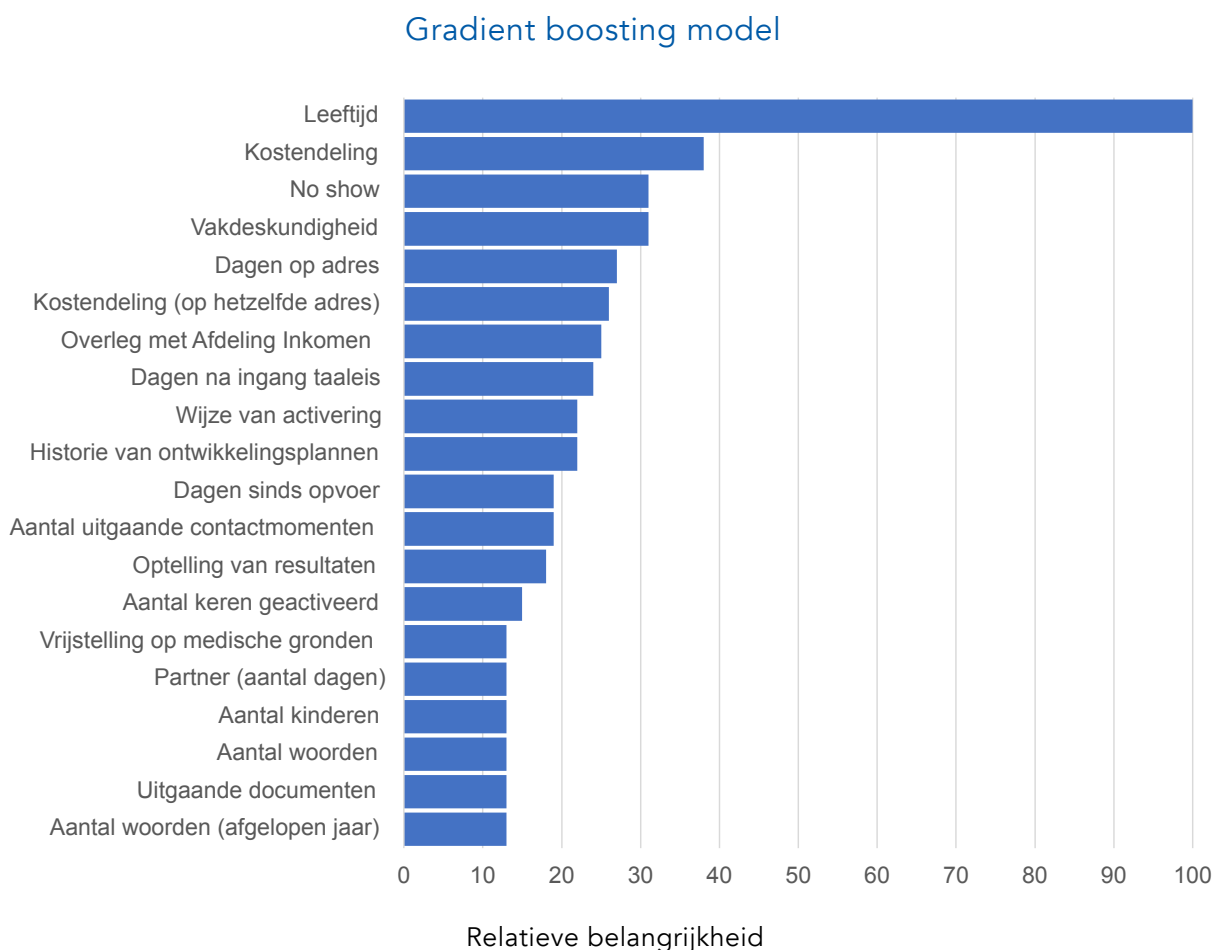
Wat is een gradient boosting model?

Gradient boosting is een machine learning methode die wordt gebruikt voor classificatie. Het geeft een voorspelling op basis van een verzameling van honderden of duizenden beslisbomen. Over het algemeen zijn boosting modellen nauwkeuriger dan simpelere modellen, zoals enkelvoudige beslisbomen of logistische regressiemethoden, al gaat dit ten koste van interpreteerbaarheid, aangezien de relatieve belangrijkheid van input variabelen gebaseerd is op een aggregatiestatistiek van het gehele ensemble van beslisbomen.

volledige lijst van 60 variabelen, onderverdeeld in 13 categorieën, die potentieel door de gemeente kunnen worden gebruikt voor zowel expert als algoritmische profilering. Uit deze lijst zijn de top-20 variabelen met het meeste voorspellende vermogen opgenomen in het gebruikte algoritmische risicotaxatie-model.

Het doel van deze casestudie is niet het bepalen welke van de 60 variabelen, of welke van de 20 geselecteerde variabelen, acceptabel zijn om als criterium opgenomen te worden in een risicoprofiel. Het gaat er vooral om met een redenering te komen welke (soorten) variabelen – rekening houdend met aspecten als proxy-risico, de link met beschermde gronden, en geschiktheid, proportionaliteit en noodzakelijkheid voor het beoogde doel – verantwoord kunnen worden gebruikt.

Figuur 3 toont de top-20 variabelen met de meest voorspellende waarde zoals aangegeven door de algoritmische *gradient boosting model* (gbm) profileringsmethode. De werking van de gbm-methode wordt toegelicht in [Box 3](#). Relatieve belangrijkheid is een



Figuur 3 – Top-20 best voorspellende variabelen volgens het gradient boosting model (gbm) profileringsalgoritme zoals gebruikt door de gemeente Rotterdam. De voorspellende waarde wordt uitgedrukt in termen van relatieve belangrijkheid (waar de variabele met de meest voorspellende kracht per definitie de waarde 100 krijgt). Een beschrijving van de beschikbare variabelen wordt gegeven in Appendix B – Dataverzameling.

maat die de voorspellende waarde van inputvariabelen berekent (waarbij aan de variabele met de meeste voorspellende kracht per definitie 100 wordt toegekend). De precieze numerieke waarde van de relatieve belangrijkheid is minder belangrijk dan de orde van grootte en de rangschikking van de verschillende variabelen. Een beschrijving van de top-20 van meest voorspellende variabelen staat in [Appendix C – Variabelenselectie methoden](#).

Van vijf verschillende kandidaatmodellen werd het gbm-algoritme door het Onderzoeks en Business Intelligence (OBI) team van de gemeente Rotterdam, dat voor dit project samenwerkte met adviesbureau Accenture, beschouwd als de best presterende methode. Als gevolg daarvan is het gbm-algoritme voor daadwerkelijke variabelenselectie ingezet. Dat betekent dat, aangevuld met andere selectiemethoden, in 2017, 2018 en 2019 respectievelijk 10%, 17% en 22% van alle heronderzoek geïnterviewden zijn geselecteerd door middel van een algoritmische profilering. Meer statistieken over de omvang en de nauwkeurigheid van de besproken selectiemethoden kunnen worden gevonden in [Appendix C – Variabelenselectie methoden](#).

Kwestie II: Ethisch (on)wenselijke criteria voor risicoprofilering

Afgezien van de kwestie van het dienen als proxy-variabelen voor beschermde kenmerken, zijn er ethische bezwaren tegen het gebruik van bepaalde persoonskenmerken voor profilering. In dit deel van de case study beschouwen we de ethische overwegingen die te maken hebben met potentieel oneerlijke, irrelevante, subjectieve of anderszins problematische persoonlijke data die worden gebruikt voor de risicoprofilering.

Data over bijstandsontvangers die voor veel Nederlandse gemeenten beschikbaar zijn, omvatten het aantal en de duur van afspraken met een werkconsulent, het bijwonen van activiteiten ter bevordering van arbeidsparticipatie en de manier waarop bijstandsontvangers communiceren met vertegenwoordigers van de gemeente, bijvoorbeeld per e-mail, telefoon of post. In Rotterdam worden aanvullende data over persoonlijke kenmerken van bijstandsontvangers verzameld zodra bijstand wordt aangevraagd, zoals psychische zorgbehoeften en ontvangen psychologische hulp. Zodra de bijstand is toegekend, scoren getrainde ambtenaren de bijstandsontvangers tijdens een gesprek (niet een heronderzoeksgesprek) op competenties en werkvaardigheden zoals motivatie, flexibiliteit, doorzettingsvermogen, zelfstandigheid, initiatief, stressbestendigheid, presentatie, professionele uitstraling, leergierigheid en houding. De volledige lijst van alle datapunten is weergegeven in [Appendix B – Dataverzameling](#). Een voorbeeld van een scoreformulier voor kan online worden geraadpleegd¹².

Men kan zich afvragen of sommige kenmerken wenselijk zijn om te dienen als selectiecriteria in risicoprofielen om bijstandsontvangers te selecteren voor heronderzoek. Niet

¹² Zie https://algorithmaudit.sharepoint.com/:b:/s/CaseRotterdam/Ec6gNE1_PwJOieXYW0rG-oMB2LO2VyY-FOj-IeuF0xz0Ogg?e=clGa2z

vanwege het risico van proxy-discriminatie, maar omdat er geen objectieve en redelijke rechtvaardiging kan worden gegeven voor onderscheid op basis van deze gronden. Zo zou onderscheid op basis van mentale gezondheidszorg (geen klassieke beschermde grond in het non-discriminatierecht) toch kunnen worden gezien als oneerlijk, omdat het sociale ongelijkheid zou kunnen versterken. Of men kan zich afvragen of het eerlijk is om bij het vaststellen van risicoprofielen rekening te houden met externe kenmerken, zoals het "professionele voorkomen" van de bijstandsontvangers of de aard van het gesprek tussen een ambtenaar en de bijstandsontvanger. Een andere zorg is de subjectieve aard van sommige datapunten, zoals de mate van "flexibiliteit" of "zelfstandigheid", die door een ambtenaar subjectief worden beoordeeld.

Eenzijds kan het voor een gemeente van toegevoegde waarde zijn om te weten dat sommige burgers een bepaalde achtergrond hebben. Door deze informatie te digitaliseren en op te slaan zou de gemeente adequate hulp kunnen bieden. Anderzijds moet men zich kritisch afvragen of deze zeer persoonlijke, gevoelige, en mogelijk subjectieve variabelen in de eerste plaats moeten worden gedigitaliseerd, en vervolgens beschikbaar moeten worden gesteld als selectiecriteria voor risicoprofilering.

Ongeacht of criteria worden geselecteerd door expert of algoritmische selectiemethoden, is het in deze context een relevante vraag welke kenmerken ethisch onwenselijk zijn om te gebruiken in profileringsmethoden, om andere redenen dan indirecte discriminatie. De tweede vraag van deze casestudie is dus welke selectiecriteria legitiem in risicoprofielen kunnen worden opgenomen. Deze kwestie relateert aan de eerste, maar nu toegespitst op de ethische bezwaren tegen het gebruik van variabelen. Dit is een contextuele normatieve afweging waarvoor geen eenduidige oplossing bestaat.

Kwestie III: Vergelijking van expert en algoritmische profilering

In het derde deel van deze case study vergelijken we de ethische implicaties van expert- en algoritmische profileringsmethoden om risicoprofielen te creëren. Maakt het uit of criteria en profielen worden afgeleid door een expert of een algoritme? Deze vraag moet in deze case study worden meegenomen om een niet te rechtvaardigen vooringenomenheid ten aanzien van algoritmische methoden voor de selectie van variabelen te voorkomen. Expert profilering op basis van handmatig gekozen criteria is het meest directe alternatief voor algoritmische profilering. Het is belangrijk te benoemen dat expert profilering ook berust op risicoprofielen die mogelijk discriminerend, oneerlijk of stigmatiserend zijn. Daarom zou een te grote nadruk op het gebruik van machine learning voor het opstellen van risicoprofielen kunnen leiden tot een toenemend gebruik van handmatige expert profilering, ook al is het onduidelijk of dat de ethische risico's daadwerkelijk vermindert.

Expertchecks

Voor zowel expert als algoritmische profilering moeten de variabelen die mogelijk als selectiecriteria worden gebruikt, handmatig worden geëvalueerd. Beide methoden vereisen

daarom dat "humans in the loop" het profileringsproces en de uitkomsten controleren op mogelijke vooringenomenheid. In Rotterdam wordt hier op de volgende manier invulling aan gegeven. Voor expert profilering wordt een risicoprofiel handmatig gegenereerd door medewerkers van de gemeente die doorgaans 1-3 selectiecriteria kiezen tijdens een expertbijeenkomst. Een integraal onderdeel van dergelijke bijeenkomsten is een kwalitatieve evaluatie of de selectiecriteria relevant, wenselijk en mogelijk discriminerend zijn. Als medewerkers bijvoorbeeld genderongelijkheid waarnemen in uitgevoerde heronderzoeken de afgelopen jaren, kunnen zij ervoor kiezen om voor het volgende jaar handmatig een criterium op te nemen om de ongelijkheid tegen te gaan. Dergelijke menselijke ingrepen in het profilerings- en selectieproces worden *expert controles* genoemd.

Expert controles maken ook deel uit van de algoritmische variabelenselectie en vinden plaats op twee vaste momenten in het selectieproces. Allereerst wordt in de trainingsfase van het algoritme een expert controle uitgevoerd op de inputdata en het statistische model. Onderdeel van de controle in dit stadium is een beoordeling welke inputvariabelen om juridische en ethische redenen moeten worden uitgesloten. Over het statistische model moet een besluit worden genomen over het aantal te gebruiken selectiecriteria. In Rotterdam is besloten de top-20 meest voorspellende variabelen (zoals geïdentificeerd door het algoritme) te gebruiken voor het finale risicoprofiel. Ten tweede wordt, zodra het getrainde model is toegepast om ontvangers van bijstand te selecteren voor heronderzoek, een verplichte expert controle uitgevoerd op de steekproefpopulatie. Een expert controleert de geselecteerde populatie op een mogelijke vooringenomenheid of andere onjuiste kenmerken.

Zodoende zijn de uit te voeren expert controles dus vergelijkbaar voor expert en algoritmische profilering en brengen vergelijkbare ethische vragen met zich mee. Voor zowel expert als algoritmische profilering worden ambtenaren geconfronteerd met de vraag of hun risicoprofielen eerlijk zijn, en welke selectievariabelen ethisch verantwoord zijn om te gebruiken. Dit betekent, zoals gezegd, dat een eenzijdige kritiek op algoritmische profilering de fundamentele ethische kwesties rondom profilering niet wegneemt, aangezien die ook gelden voor expert profilering.

Verschillen tussen expert en algoritmische profilering

Toch zijn er ook verschillen tussen expert en algoritmische profilering die overwogen moeten worden. Het duidelijkste verschil is de grotere nauwkeurigheid van algoritmische profilering vergeleken met expert profilering. Zoals blijkt uit toepassing van het Rotterdam-algoritme, presteert algoritmische profilering in de periode 2017-2021 in de praktijk systematisch beter dan expert profilering om onrechtmatig toegekende bijstandsuitkeringen te voorspellen. In deze periode is de nauwkeurigheid van algoritmische profilering, geconditioneerd op terugvorderingen, circa 29,9% waar de nauwkeurigheid van expert profilering circa 16,5% betreft (n=4.388). Een hogere nauwkeurigheid betekent niet alleen een hogere efficiëntie voor het opsporen van onrechtmatige bijstandsaanvragen, wat gunstig is voor de gemeente, maar ook een afname betekent van het aantal "vals positieven"

– d.w.z. rechtmatige bijstandsontvangers die onnodig worden geselecteerd voor heronderzoek. Een daling van vals positieven is gunstig voor bijstandsontvangers, omdat geselecteerd voor heronderzoek als tijdrovend en stressvol wordt ervaren (zie [Box 1](#)). Resultaten per selectiemethode (expert en algoritmische profilering, signaal en aselechte steekproef) worden vermeld in [Appendix D – Resultaten per selectiemethode](#). In deze appendix worden ook andere relevante maten van de algoritmische profileringsmethode gegeven.

Een ander relevant verschil is het feit dat algoritmische methoden gebruik maken van veel meer data en persoonskenmerken. Dit is duidelijk in bij het Rotterdam-algoritme: expert profilering gebruikt typisch 1-3 selectievariabelen, terwijl algoritmische profilering 20 variabelen gebruikt. Het gebruiken van meer variabelen betekent een grotere kans op het gebruik van potentieel problematische variabelen: bij 20 variabelen is de kans groter dat een discriminerende proxy-variabelen, subjectieve of irrelevante persoonskenmerken wordt gebruikt. Het zorgt ook voor meer complexiteit. Algoritmische profilering selecteert persoonlijke kenmerken op een statistisch verfijnde manier. Het is vaak moeilijk om in niet-technische, toegankelijke taal uit te leggen waarom precies een individuele ontvanger van bijstand door een hoge risicoscore krijgt toegewezen. Voor algoritmische profilering kan een selectie worden uitgelegd aan de hand van de relatieve belangrijkheid van selectiecriteria, zoals weergegeven in [Figuur 3](#). De relatieve belangrijkheidsscore, die is gebaseerd op geaggregeerde data, is echter beperkt inzichtelijk om keuzes over individuen toe te lichten. Voor expert profilering is dit veel eenvoudiger. Aangezien het er simpelweg om gaat of een bijstandsontvanger voldoet aan een op voorhand vastgesteld risicoprofiel. Expert profilering is dus beter uitlegbaar dan algoritmische profilering. Ten slotte maken de relatief eenvoudige door expert gekozen risicoprofielen het gemakkelijker om profielen door de jaren heen te variëren en te diversificeren, om een historische vooringenomenheid te doorbreken. Zo kan een duidelijke vooringenomenheid van een profiel ten aanzien van bepaalde wijken in het ene jaar worden tegengegaan door in het volgende jaar handmatig te kiezen voor andere wijken. Alhoewel mogelijk, is het voor algoritmische profilering moeilijker deze variaties handmatig af te dwingen. Dit zijn enkele noemenswaardige, ethisch relevante verschillen tussen algoritmische en expert profilering.

Kortom, expert en algoritmische profilering zijn alternatieve methoden met hun eigen ethische voor- en nadelen. Voor een omvattende evaluatie van risicoprofilering door Nederlandse gemeenten is een afweging tussen expert en algoritmische profileringsmethoden onmisbaar. Beide profileringsmethoden zijn duidelijk onvolmaakt. De methoden bevatten allebei irreduceerbare vooringenomenheid en oneerlijkheid, en leveren beperkte prestaties. Toch doen ze dat op hun eigen manier. De vraag is wat het betere alternatief is in dit concrete geval en, wellicht, of beide profileringsmethoden onrechtvaardig zijn.

Appendix A – Juridische achtergrond van heronderzoek bijstandsuitkering

Populatie van bijstandontvangers die in aanmerking komen voor heronderzoek

Niet alle ontvangers van bijstand komen in aanmerking om geselecteerd te worden voor heronderzoek. Ontvangers die aan de volgende criteria voldoen mogen niet worden geselecteerd:

- > Een bijstandontvanger kan ten hoogste eens in de twee jaar worden geselecteerd voor een heronderzoeksgesprek;
- > Wie bijstandsuitkering heeft ontvangen voor een periode van 6 maanden of korter is uitgesloten van heronderzoek;
- > Bijstandontvangers ouder dan 64 jaar komen niet in aanmerking voor een heronderzoek;
- > Bijstandontvangers zonder een adres of wonend in een zorginstelling komen niet in aanmerking voor heronderzoek.

Algemene wet bestuursrecht

Op besluitvormingsprocessen van Nederlandse gemeenten is de Algemene wet bestuursrecht (Awb) van toepassing. Dit wettelijke kader stelt eisen aan gemeenten om zich te houden aan onder meer de Algemene beginselen van behoorlijk bestuur (Abbb):

- > Artikel 2:4 Awb (verbod op vooringenomenheid);
- > Artikel 3:2 - 3:4 lid 1 Awb (zorgvuldigheidsbeginsel);
- > Artikel 3:47 Awb (motiverings- en rechtszekerheidsbeginsel).

Het verenigen van de Abbb met de inzet van algoritmische methoden voor selectiecriteria is een uitdaging. Algoritmische selectie van criteria voor profilering is geen zelfstandig besluit (Awb 1:3), omdat een medewerker van de gemeente na een heronderzoeksgesprek een beslissing neemt of de uitkering (on)rechtmatig wordt verleend. Variabelenselectie zou echter gezien kunnen worden als onderdeel van de zorgvuldige voorbereiding van dit besluit. Uitlegproblemen waarom bepaalde criteria op zijn genomen in een risicoprofiel kan ertoe leiden dat een overheidsorganisatie weliswaar 'rechtmatig' maar niet 'behoorlijk' handelt. Aanvullende organisatorische en juridische eisen hoe algoritmische profileringmethoden gerijmd kunnen worden met het behoorlijkheidsbeginsel is echter een open en context- afhankelijke vraag. Deze case study hoopt bij te dragen aan een antwoord op deze vraag.

Richtlijnen in Hoofdstuk 5 van de Awb vormen de wettelijke basis voor de handhaving van sociale voorzieningen. De Participatiewet valt onder het bijzonder bestuursrecht (zie hieronder).

Algemene verordening gegevensbescherming

De Algemene verordening gegevensbescherming (Avg) regelt de opslag en verwerking van gegevens over ontvangers van een bijstandsuitkering. Hoe de dataverwerkingsmethoden in deze casestudie zich verhouden tot de vereisten zoals gesteld in de GDPR is te vinden in de Privacy Impact Assessment (PIA), zoals uitgevoerd door de gemeente Rotter-

dam¹³. Bepaalde Avg bepalingen die relevant zijn voor deze case study worden hieronder vermeld:

> **Artikel 4 – Profilering definitie**

Profilering betreft “iedere vorm van geautomatiseerde verwerking van persoonsgegevens die bestaat uit het gebruik van persoonsgegevens voor de beoordeling van bepaalde aspecten van een natuurlijke persoon, met name voor het analyseren of voorspellen van aspecten die verband houden met de diens arbeidsprestaties, economische situatie, gezondheid, persoonlijke voorkeuren, interesses, betrouwbaarheid, gedrag, locatie, of verplaatsingen”.

> **Artikel 5 – Beginselen inzake verwerking van persoonsgegevens**

- **Doelbinding:** De gemeente Rotterdam meent zich te houden aan het beginsel van doelbinding aangezien de Participatiewet gemeenten verplicht onrechtmatige uitgekeerde bijstandsuitkeringen op te sporen. Alle gegevens worden verzameld ten behoeve van de naleving van de Participatiewet, zijnde persoonlijke klantgegevens, arbeidsverleden, traject naar werk, etc.;
- **Dataminimalisatie:** Zoals vermeld in de PIA, zegt de gemeente Rotterdam zich te houden aan het principe van dataminimalisatie, aangezien “voor dit project zo min mogelijk gegevens worden verzameld” en “alleen de meest voorspellende selectiecriteria, zoals aangegeven door algoritmische profilering, in aanmerking worden genomen voor selectiemethoden”. In hetzelfde document staat echter dat “zoveel mogelijk gegevensvariabelen in aanmerking worden genomen voor modeltraining om onverwachte patronen aan het licht te brengen”.

> **Artikel 6 – Rechtmatigheid van de verwerking**

Artikel 6 lid 1 sub e vormt de grondslag voor verwerking, gestoeld op de noodzaak voor de vervulling van een taak van algemeen belang. Deze taak vloeit voort uit de Participatiewet (artikel 53a, 64) en de Wet structuur uitvoeringsorganisatie werk en inkomen (Suwi, artikel 62);

> **Artikel 9 – Verwerking van bijzondere categorieën van persoonsgegevens**

In principe bepaalt lid 1 van dit artikel dat de verwerking van gezondheidsdata verboden is. Of deze verwerking onder een uitzondering valt is op moment van schrijven bij ons niet bekend;

> **Artikel 13-15 Te verstrekken informatie wanneer persoonsgegevens bij de betrokkene worden verzameld, -niet van de betrokkene zijn verkregen en Recht van inzage van de betrokkene**

Artikel 13.2.f, artikel 14.2.g en artikel 15.1.h stellen dat de betrokkene recht heeft op “zinnige informatie over de betrokken logica” met betrekking tot profilering;

¹³ FOI verzoek VPRO Argos/Lighthouse Reports <https://www.vpro.nl/dam/jcr:c87f2d6c-3f9c-4498-9a9c-f3b-c5483a437/Downloads%20Model%20Rotterdam.zip>

- > **Artikel 22 - Geautomatiseerde individuele besluitvorming, waaronder profilering**
Medewerkers van de gemeente nemen het uiteindelijke besluit of uitkeringen (on)rechtmatig worden toegekend op basis van een persoonlijk heronderzoek. Profileringmethoden dienen hierin als selectiemethode en worden derhalve niet beschouwd als volledig geautomatiseerde besluitvorming (fully automated decision-making) en zijn vrijgesteld van bepalingen zoals geformuleerd in dit artikel (zie ook de subsectie over het Algemene wet bestuursrecht).

Participatiewet

Artikel 11 van de Nederlandse Participatiewet garandeert een minimum inkomen voor iedereen die wettelijk gezien in Nederland woont en die onvoldoende middelen heeft om zichzelf te onderhouden. Wanneer er bijstandsuitkering wordt ontvangen, verplicht artikel 7 de ontvangers om deel te nemen aan een "traject naar werk". Gemeenten helpen bijstandsonvangers in dit traject. Bepaalde groepen, bijvoorbeeld alleenstaande ouders met een kind van tot 5 jaar oud, kunnen vrijstelling aanvragen van deze verplichting. Mensen zijn echter wel verplicht trainingscursussen bij te wonen. Artikel 53a en 64 bieden de rechtsgrond voor heronderzoek.

Wet structuur uitvoeringsorganisatie werk en inkomen

Artikel 62 van de Wet structuur uitvoeringsorganisatie werk en inkomen (Suwi) regelt de onderlinge gegevensverstrekking tussen uitvoeringsorganisaties en gemeenten.

Appendix B – Dataverzameling

Gegevens van bijstandsontvangers verzameld door de gemeente Rotterdam

Categorie	Variabele	Voorbeeld datapunt
Adres	Aankomstdatum, vertrekdatum	
	Nederlandse postcode	Vier getallen van postcode
	Stadswijk	"Charlois", "Binnenstad", "Zuidwijk"
Afspraak	Datum	
	Tekst over een afspraak met een job coach, klantmanager of heronderzoeksinterviewer etc	Omschrijving van een afspraak met een ambtenaar, bijv. "Begin bijstand", "Einddatum bijstand bereikt", "Heronderzoeksgesprek"
	Resultaat van de afspraak	"Aanvullende informatie vereist", "Deelname aan een sociale werkplek", "Uitkering beëindigd"
	Reden voor de afspraak	"Introducerend gesprek", "Administratieve controle", "Uitkering beëindigd"
Beschikbaarheid	Aanvullende systeemtekst	"Klant is mondeling gewaarschuwd", "Onderzoek naar verlenging van de bijstand", "Ingeschreven voor vervolgtraject"
	Begindatum, einddatum	
	Beschikbaarheid, titel	"Beschikbaar", "Beperkt beschikbaar wegens zorg <5 jaar oud"
Typering	Beschikbaarheidsomschrijving	"Ernstige fysieke en psychologische problemen", "In therapie voor chronische"
	Begindatum, einddatum	
	Karakteristiek	"Sector transport en logistiek", "Vluchteling", "2e profiel: snel, precies en fysiek fit"
Contact	Documentatiedatum	
	Contacttype	"Email (uitgaand)", "Gesprek", "Document (uitgaand)"
	Onderwerp	"Inspanningstoets", "Traject", "(Arbeids) motivatie"
Competentie	Begindatum, einddatum	
	Competentie (titel)	"Inlevingsvermogen", "Beslissen en activiteiten initiëren", "Stressbestendig"

Categorie	Variabele	Voorbeeld datapunt
	Competentie (omschrijving)	"behandeling psycholoog ivm psych klachten", "ADHD, maakte rare taalfouten, komt niet geloofwaardig over", "analfabeet, niet aanspreekbaar in Ned. taal"
Ontheffing	Begindatum, Einddatum	
	Vrijstellingstype	"Sociale gronden", "Tijdelijke ontheffing arbeidsplicht en tegenprestatie", "Me-
Belemmering	Begindatum, einddatum	
	Belemmeringscategorie	"Psychische problemen", "Niet computervaardig", "Volledig ontbrekende verdien capaciteit"
	Belemmeringsomschrijving	"PTSS. In behandeling bij psycholoog van Pionier te Amsterdam.", "Dit gaat tijdens het gesprek op 9-2-2015 beter. Ze geeft aan dat haar dochter haar helpt.", "me- neer kampt met schulden, klant heeft een aantal regeling lopen."
Instrument	Begindatum, einddatum	
	Reden voor beëindiging	"Matchbaar", "Doelstelling bereikt, uit- stroom naar regulier werk", "Overdracht naar Prematching", "Werk/Re- integratie", "Ondersteunende instrumenten", "Activering"
Deelname activiteiten	Begindatum, einddatum	
	Herintegradieladder ('traject tot werk' programma)	"Activering", "Ondersteunende instru- menten", "Werk/Re- integratie"
	Vroegtijdige beëindiging	"Betaald werk gevonden", "Disfunction- eren", "Verhuizing buiten gemeente"
Persoon	Geslacht	
	Geboortemaand	
Persoonlijke eigen- schappen	Houding	
	Autonomie	
	Assertiviteit	
	Communicatie	
	Discipline	
	Leergierig	
	Flexibiliteit	
	Hobby's/Sport	

Categorie	Variabele	Voorbeeld datapunt
	Sollicitatiegedrag	
	Taal – Nederlandse leesvaardigheid	
	Taal – Nederlandse spreekvaardigheid	
	Taal – Nederlands taalbegrip	
	Taal – Nederlandse schrijfvaardigheid	
	Taalvereisten vervuld (spreken, luisteren, lezen, schrijven)	
	Motivatie	
	Presenteervaardigheid	
	Professioneel voorkomen	
	Overige opmerkingen	
	Arrestatie- en aanhoudingsportfolio	
	Traject tot werk activatie	
	Werk tijdens kantooruren	
	Werk buiten kantooruren	
Relatie (Kind/partner/overig)	Begindatum, einddatum	
	Type	"ouder -> kind", "verhuurder -> huurder", "partner - partner (gehuwd)"
		"kindrelatie", "medebewoner", "partnerrelatie"
Steunplan	Begindatum, voorgestelde datum, datum van ondertekenen	
	Planbeschrijving	"Bemiddeling", "Diagnose", "Sociale activering"
	Reden voor beëindiging	"Doelstelling bereikt, overdracht", "Klant ziet af van aanbod", "Uitstroom als zelfstandige"
	Beschrijving van doelen	"Uitstroom naar regulier werk door middel van korte trajecten en/of regelmatige afspraken.", "Maatschappelijke participatie door middel van activiteiten die gericht zijn op het actief positioneren in de samenleving."

Tabel 1 – Verzamelde datapunten over bijstandsontvangers in de gemeente Rotterdam³

Appendix C – Variabelenselectie methoden

Variabelenselectie door experts

Bij de gemeente Rotterdam verandert de samenstelling van expert-gestuurde risicoprofielen jaarlijks. Profielen worden gestuurd door bepaalde meerjarige profieldomeinen, zoals risico per "type huishouden", "leeftijd", en/of "wijk". Voorbeelden van type huishouden zijn "alleenstaande mannen" en "alleenstaande vrouwelijke huurders"¹⁴. In dit proces worden de betrokken experts ondersteund door de Privacy Office van de gemeente Rotterdam, die verantwoordelijkheid draagt voor rechtmatigheid van de gekozen profielen.

Variabelenselectie door algoritmes

De top-20 variabelen met de meest voorspellende waarde om onrechtmatige bijstandsuitkeringen te voorspellen (zoals weergegeven in [Figuur 3](#)) worden vermeld in [Tabel 2](#). Het gradient boosting model (gbm) wordt beschouwd als het best functionerende algoritme onder de kandidaatsmethoden voor variabelenselectie op een trainingsdataset: gbm, glmnet, random forest, rpart, en xgbtree.

Beschrijving van de top-20 meest voorspellende variabelen

Naam van variabele	Omschrijving
Leeftijd	Leeftijd van de ontvanger van bijstand
Kostendeling	Meerdere volwassenen delen een woning, maar zijn niet samenlevende. De hoeveelheid bijstand is gereduceerd
No show	Zonder aankondiging is de bijstandsontvanger niet komen opdagen bij een afspraak met de gemeente
Vakdeskundigheid	Competentie zoals beoordeeld door de Uitvoeringsinstituut Werknemersverzekeringen (UWV). Deze informatie is uit de UWV database verkregen
Dagen op adres	Aantal dagen verblijvend op een adres
Kostendeling (op hetzelfde adres)	Meerdere volwassenen leven op hetzelfde adres. De hoeveelheid bijstand is gereduceerd
Overleg met Afdeling Inkomen	Werknemers van de Afdeling Inkomen hebben interactie met de ontvanger van bijstand gehad
Dagen na ingang taaleis	Het aantal dagen sinds het ingaan van de taalvereisten
Wijze van activering	Huidige methode van "activatie" die wordt gebruikt om burgers naar werk te begeleiden. Zie ook variabele Instrument in Appendix B – Dataverzameling
Historie van ontwikkeling-splannen	Som van verleden activiteiten om bijstandsontvangers te ondersteunen. Zie ook variabele Steuplan in Appendix B – Dataverzameling.

¹⁴ Sectie 4-3 van Gekleurde Technologie, Rotterdamse Rekenkamer (2021) <https://rekenkamer.rotterdam.nl/onderzoeken/algoritmes/>

Naam van variabele	Omschrijving
Dagen sinds opvoer	Aantal dagen sinds de aanvraag voor bijstandsuitkering
Aantal uitgaande contactmomenten	Aantal brieven met beschrijving "verzonden document" aan de bijstandsontvanger in het afgelopen jaar
Optelling van resultaten	Som van vermelde interacties
Aantal keren geactiveerd	Som van 'activaties' opgeslagen in de kolom 'Heringtegratieladder'. Zie ook Appendix B – Dataverzameling
Vrijstelling op medische gronden	Aantal dagen dat een bijstandsontvanger in het verleden is vrijgesteld van het 'traject tot werk' vanwege medische omstandigheden
Partner (aantal dagen)	Aantal dagen dat een partner is geregistreerd
Aantal kinderen	Aantal geregistreerde (adoptie) kinderen
Aantal woorden	Totaal aantal woorden dat in het verleden is gebruikt door een medewerker van de gemeentes in het tekstveld
Uitgaande documenten	Aantal keer dat er in het verleden met de bijstandsontvanger gecorrepondeerd met omschrijving 'uitgaand document'
Aantal woorden (afgelopen jaar)	Totaal aantal woorden dat afgelopen jaar is gebruikt door een medewerker van een gemeente in het tekstveld

Tabel 2 – Beschrijving van top-20 variabelen met meeste voorspellende vermogen volgens het gbmodel

Appendix D – Resultaten per selectiemethode

Kwantitatieve aspecten

Kwantitatieve maten geven inzicht in de effectiviteit van profileringsmethoden. Deze maten zijn belangrijk voor het nemen van bestuurlijke besluiten, maar zouden ook voor ethische evaluatie gebruikt moeten worden. Bijvoorbeeld: als algoritmische profilering een grote winst in nauwkeurigheid oplevert, worden de ethische risico's anders gewogen dan wanneer de nauwkeurigheidswinst slechts minimaal is. Zowel kwantitatieve als kwalitatieve aspecten zouden bij de ethische beoordeling een rol moeten spelen.

Resultaten per selectiemethode in Rotterdam

De gemeente Rotterdam heeft de uitkomsten van heronderzoeken per selectiemethode gedeeld voor de periode 2017-2021. Er zijn vier verschillende uitkomsten van een heronderzoek: geen actie, beëindiging, terugvordering en administratieve aanpassingen. Alleen de laatste drie uitkomsten zijn weergegeven in [Tabel 3](#) - [Tabel 7](#). Het totaal aantal ontvangers van bijstand in de gemeente Rotterdam in 2017-2021 was: 44.600 (juni 2017), 42.110 (juni 2018), 39.840 (juni 2019), 48.830 (juni 2020) en 46.090 (juni 2021)¹⁵.

2017	Afgeronde heronderzoeken	Beëindiging		Terugvordering		Administratieve aanpassingen	
		Aantal	Procent	Aantal	Procent	Aantal	Procent
Algoritmische profilering	255	44	17,3%	65	25,5%	-	0,0%
Expert profilering	1.817	317	17,4%	269	14,8%	9	0,5%
Signaal steekproef	166	73	44,0%	32	19,3%	-	0,0%
Aselecte steekproef	2	-	0,0%	-	0,0%	-	0,0%
Totaal	2.240	434	19,4%	366	16,3%	9	0,4%

Tabel 3 – Nauwkeurigheid van profilerings- en steekproefmethoden in de gemeente Rotterdam in 2017

2018	Afgeronde heronderzoeken	Beëindiging		Terugvordering		Administratieve aanpassingen	
		Aantal	Procent	Aantal	Procent	Aantal	Procent
Algoritmische profilering	477	86	18,0%	100	21,0%	6	1,3%
Expert profilering	2.013	264	13,1%	395	19,6%	22	1,1%
Signaal steekproef	305	116	38,0%	54	17,7%	1	0,3%
Aselecte steekproef	2.961	190	6,4%	515	17,4%	35	1,2%
Totaal	5.756	656	11,4%	1.064	18,5%	64	1,1%

Tabel 4 – Nauwkeurigheid van profilerings- en steekproefmethoden in de gemeente Rotterdam in 2017

¹⁵ Centraal Bureau van de Statistiek <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/80794ned/table?dl=34613>

2019	Afgeronde heronderzoeken	Beëindiging		Terugvordering		Administratieve aanpassingen	
		Aantal	Procent	Aantal	Procent	Aantal	Procent
Algoritmische profilering	1.428	188	13,2%	381	26,7%	18	1,3%
Expert profilering	2.946	210	7,1%	612	20,8%	36	1,2%
Signaal steekproef	824	251	30,5%	199	24,2%	14	1,7%
Aselecte steekproef	1.169	64	5,5%	278	23,8%	20	1,7%
Totaal	6.367	713	11,2%	1.470	23,1%	88	1,4%

Tabel 5 – Nauwkeurigheid van profilerings- en steekproefmethoden in de gemeente Rotterdam in 2020

2020	Afgeronde heronderzoeken	Beëindiging		Terugvordering		Administratieve aanpassingen	
		Aantal	Procent	Aantal	Procent	Aantal	Procent
Algoritmische profilering	1.207	102	8,5%	431	35,7%	317	26,3%
Expert profilering	1.850	67	3,6%	418	22,6%	384	20,8%
Signaal steekproef	265	73	27,5%	101	38,1%	45	17,0%
Aselecte steekproef	24	1	4,2%	13	54,2%	8	33,3%
Totaal	3.346	243	7,3%	963	28,8%	754	22,5%

Tabel 6 – Nauwkeurigheid van profilerings- en steekproefmethoden in de gemeente Rotterdam in 2021

2021	Afgeronde heronderzoeken	Beëindiging		Terugvordering		Administratieve aanpassingen	
		Aantal	Procent	Aantal	Procent	Aantal	Procent
Algoritmische profilering	253	12	4,7%	105	41,5%	98	38,7%
Expert profilering	3.841	77	2,0%	359	9,3%	388	10,1%
Signaal steekproef	134	48	35,8%	53	39,6%	26	19,4%
Aselecte steekproef	242	4	1,7%	8	3,3%	15	6,2%
Totaal	4.470	141	3,2%	525	11,7%	527	11,8%

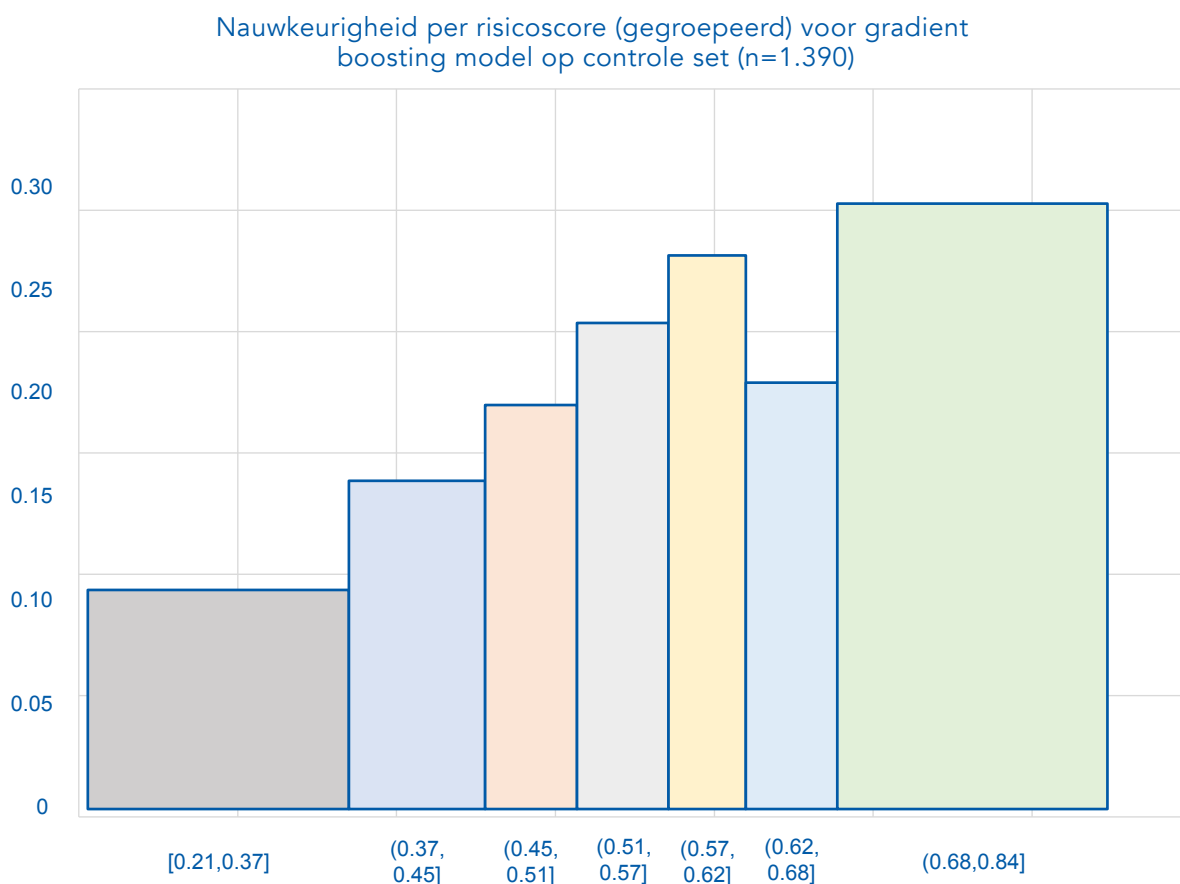
Tabel 7 – Nauwkeurigheid van profilerings- en steekproefmethoden in de gemeente Rotterdam in 2021

Nauwkeurigheid van de gradient boosting model op nieuwe data

De prestatie van het gradient boosting model op controle dataset, d.w.z. nieuw verkregen data die geen onderdeel heeft uitgemaakt van de trainings-, validatie- en test dataset waarop het model getraind is. De dataset bestaat uit ontvangers van bijstand die zijn geselecteerd door expert, signaal of aselechte steekproeven en de uitkomst van het heronderzoek dat heeft plaatsgevonden.

Heronderzoekstatistieken uit Amsterdam, Den Haag en Rotterdam

Drie van de vier grootste Nederlandse gemeenten (Amsterdam, Den Haag en Utrecht) hebben statistieken gedeeld over het aantal heronderzoek die door de gemeente zijn uitgevoerd en de uitkomsten hiervan (rechtmatig of onrechtmatig toegekend). Alleen in de gemeente Rotterdam is naar ons weten in deze periode gebruik gemaakt van algoritmische steekproeven. Bijstandsontvangers in andere gemeenten zijn geselecteerd voor heronderzoek middels expert, signaal en aselechte steekproeven. In 2021, woonde circa 30% van alle ontvangers van bijstand in Nederland (± 365.000) in een van de vier grote steden¹⁶.



Figuur 4 – Nauwkeurigheid van gbmodel op nieuwe data set (n=1.390)

¹⁶ Dashboard over bijstandsuitkeringen, Centraal Bureau van de Statistiek (CBS) <https://dashboards.cbs.nl/v2/dashboardSOZ/>

Amsterdam

De gemeente Amsterdam leverde cijfers van het aantal afgeronde heronderzoeken en de uitkomsten daarvan, inclusief de totale hoeveelheid teruggevorderde bedragen.

Jaar	Totaal aantal bijstandsgerechtigden ¹⁶	Aantal afgeronde heronderzoeken ¹⁷	Terugvorderingen
2021	39.378	1.728	€4.153.540
2020	40.298	2.715	€6.797.687
2019	39.633	2.277	€4.873.928

Tabel 8 – Statistieken over uitgevoerde heronderzoeken in de gemeente Amsterdam

Den Haag

De gemeente Den Haag leverde cijfers van het aantal afgeronde heronderzoeken en de uitkomsten daarvan (rechtmatig of onrechtmatig toegekend) en de totale hoeveelheid teruggevorderde bedragen.

Jaar	Totaal aantal bijstandsgerechtigden ¹⁶	Aantal afgeronde heronderzoeken ¹⁸	Uitkomst	Aantal	Terugvorderingen
2021	24.020	1.698	Rechtmatig toegekend	909	€4.153.540
			Onrechtmatig toegekend	789	
2020	24.683	1.292	Rechtmatig toegekend	685	€6.797.687
			Onrechtmatig toegekend	607	
2019	24.668	1.937	Rechtmatig toegekend	1.289	€4.873.928
			Onrechtmatig toegekend	648	

Tabel 9 – Statistieken over uitgevoerde heronderzoeken in de gemeente Den Haag

¹⁷ Data verstrekt door de gemeente Amsterdam op verzoek

¹⁸ Data verstrekt door de gemeente Den Haag op verzoek

Utrecht

De gemeente Utrecht publiceert kwartaalcijfers over het aantal afgeronde heronderzoeken en de uitkomsten hiervan (rechtmatig of onrechtmatig toegekend).

Jaar	Totaal aantal bijstandsgerechtigden ¹⁶	Aantal afgeronde heronderzoeken ¹⁹	Uitkomst	Aantal
2021	10.460	523	Rechtmatig toegekend	202
			Onrechtmatig toegekend	321
2020	10.695	741	Rechtmatig toegekend	293
			Onrechtmatig toegekend	448
2019	10.518	1.629	Rechtmatig toegekend	611
			Onrechtmatig toegekend	1.018

Tabel 10 – Statistieken over uitgevoerde heronderzoeken in de gemeente Utrecht

¹⁹ Data verstrekt bij de gemeente Utrecht in brieven naar de gemeenteraad <https://utrecht.bestuurlijkeinformatie.nl/Reports/Item/779c76d3-7fcd-4661-9c99-df75ba58a3b1>

Financiering van Algorithm Audit

Stichting Algorithm Audit wordt ondersteund door onafhankelijke publieke en filantropische subsidies. Budget wordt besteed aan het opstellen van ongevraagd probleemstellingen waarin ethische kwesties, die zich voordoen bij de inzet van algoritmes, worden gepresenteerd. De experts die deelnemen in een auditcommissie worden een vergoeding aangeboden. We delen de inzichten voortkomend uit ons werk met zowel de samenleving als de internationale AI auditing community om publieke kennis op te bouwen over de verantwoorde inzet van algoritmen. Werken zonder winst oogmerk dient de doelen van de stichting het beste.



Structurele partners van Algorithm Audit

SIDNfonds

SIDN Fonds

Het SIDN Fonds staat voor een sterk internet voor iedereen. Het Fonds investeert in projecten met lef en maatschappelijke meerwaarde, met als doel het borgen van publieke waarden online en in de digitale democratie.

European
Artificial Intelligence
& Society Fund

European AI&Society Fund

Het European AI&Society Fund ondersteunt organisaties uit heel Europa die AI beleid vormgeven waarin mens en maatschappij centraal staan. Het fonds is een samenwerkingsverband van 14 Europese en Amerikaanse filantropische organisaties.



Ministerie van Binnenlandse Zaken en
Koninkrijksrelaties

Ministerie van Binnenlandse Zaken en Koninkrijksrelaties

Het ministerie van BZK maakt zich sterk voor een democratische rechtsstaat, met een slagvaardig bestuur. Ze borgt de kernwaarden van de democratie. BZK staat voor een goed en digitaalvaardig openbaar bestuur en een overheid waar burgers op kunnen vertrouwen.



 www.algorithmaudit.eu

 www.github.com/NGO-Algorithm-Audit

 info@algorithmaudit.eu

Stichting Algorithm Audit is geregistreerd bij de Kamer
van Koophandel onder nummer 83979212