

Algoritmes afwegen

Verkenning naar maatregelen ter bescherming van mensenrechten bij profilering in de uitvoering



Auteurs

Jurriën Hamer, Amber Lemmens en Linda Kool

Illustraties

Rathenau Instituut

Foto omslag

Foto: David Rozing / ANP

Bij voorkeur citeren als:

Rathenau Instituut (2022). *Algoritmes afwegen – Verkenning naar maatregelen ter bescherming van mensenrechten bij profilering in de uitvoering*. Den Haag. Auteurs: Hamer J., A. Lemmens en L. Kool

Voorwoord

Uitvoeringsorganisaties vormen een essentieel onderdeel van de digitale overheid. In toenemende mate gebruiken organisaties als de Belastingdienst, uitvoerende onderdelen van gemeentes, het UWV, de SVB en het CJIB digitale hulpmiddelen om wet- en regelgeving te vertalen naar de praktijk. Voorzieningen als uitkeringen en toeslagen worden steeds meer per algoritme aan de burger geleverd.

De afgelopen jaren lieten zien hoezeer de rechten van burgers daarbij op het spel staan. Digitale innovaties kunnen uitvoeringsorganisaties helpen efficiënter en gericht diensten aan burgers te leveren en fraude te bestrijden. Maar ze kunnen ook discriminerende effecten hebben, de privacy schenden en burgers confronteren met dienstverlening waarvan ze niets meer begrijpen. Voor zover deze risico's niet bekend waren, werden ze pijnlijk zichtbaar tijdens de kinderopvangtoeslagenaffaire.

De Rijksoverheid, denktanks, wetenschappers en uitvoeringsorganisaties werken al jaren aan het versterken van de mensenrechtelijke waarborgen. Er zijn hulpmiddelen ontwikkeld als de Impact Assessment Mensenrechten en Algoritmes (IAMA) en de Handreiking Non-Discriminatie. Uitvoeringsorganisaties hebben maatregelen genomen door bijvoorbeeld te investeren in ethische expertise.

In opdracht van het ministerie van Binnenlandse Zaken en Koninkrijksrelaties inventariseerden we de genomen maatregelen. Daarbij is vanwege de mensenrechtelijke aspecten gekozen voor een focus op profilering. We spraken met medewerkers van drie uitvoeringsorganisaties: het UWV, het CJIB en de SVB. We formuleren zeven aanbevelingen om de bescherming van mensenrechten in de uitvoering verder te versterken.

Het Rathenau Instituut wil met deze studie bijdragen aan de ethische versterking en doorontwikkeling van de verantwoorde inzet van technologie door uitvoeringsorganisaties. Voor het functioneren van onze democratie en voor het dagelijks leven van burgers, is het belangrijk dat de overheid de rechten van burgers helder in het vizier heeft, en zorgvuldig nadenkt over de gevolgen van algoritmische diensten in de praktijk, die vaak weerbarstig is. Op die manier komt een betrouwbare digitale samenleving die burgers ondersteunt dichterbij.

Drs. Henk de Jong

Directeur Rathenau Instituut

Samenvatting

Introductie

Het tijdperk van de algoritmische uitvoeringsorganisatie is al enkele decennia gaande. Organisaties gebruiken inmiddels op grote schaal algoritmes om diensten uit te voeren. Daarbij draaide het lange tijd hoofdzakelijk om het toepassen van *rule-based* algoritmes. Het structureel in gaan zetten van algoritmische profilering vormt in deze ontwikkeling de volgende stap: door het gedrag van burgers met lerende algoritmes te voorspellen zouden uitvoeringsorganisaties de regels beter kunnen handhaven, en de burger sneller en met meer maatwerk kunnen bedienen.

Deze nieuwe stap die uitvoeringsorganisaties zetten, verdiept de mensenrechtelijke uitdagingen die gepaard gaan met algoritmische systemen. Door de inzet van lerende systemen wordt het bijvoorbeeld lastiger om discriminatie op te sporen en de keuzes van systemen goed uit te leggen of zelfs volledig te kennen. In het recente verleden, en in het bijzonder tijdens de kinderopvangtoeslagenaffaire, bleek hoe de bescherming van mensenrechten tekort kan schieten, en dat hierin zowel de algoritmes zelf, als de context waarbinnen ze ontwikkeld en gebruikt worden, een rol spelen. Dat wil zeggen, ook de wetten die gemaakt worden, de aanwezige expertise en de werkcultuur zijn van invloed op de ontwikkeling en toepassing van specifieke algoritmes.

De studie

In dit rapport verkennen we op verzoek van het ministerie van Binnenlandse Zaken en Koninkrijksrelaties welke maatregelen uitvoeringsorganisaties nemen om mensenrechtelijke risico's bij de inzet van profilerende systemen te mitigeren. Ook kijken we hoe deze maatregelen verder versterkt kunnen worden. We hebben deskresearch uitgevoerd om de relevante mensenrechten te selecteren en voerden gesprekken met verschillende medewerkers van het CJIB, UWV en de SVB. De voorlopige bevindingen van deze gesprekken zijn besproken tijdens twee externe consultaties. Het onderzoek is uitgevoerd van juli 2021 tot maart 2022.

Van pilots naar productie

Hoewel de soorten profilerende systemen en de maatregelen die uitvoeringsorganisaties nemen deels verschillen per organisatie, maakt ons onderzoek een aantal grote lijnen zichtbaar. Het gaat om pilots die vanaf eind 2021/begin 2022 in de reguliere werkwijze van organisaties worden ingezet, en een ontwikkeltijd van twee tot drie jaar kennen. De doelen van de profilering variëren. Het UWV werkt bijvoorbeeld aan een systeem dat gericht hulp biedt bij

het vinden van werk, en aan systemen om gericht foutief gebruik van regelingen te detecteren. Ook het SVB ontwikkelt een pilot voor dit laatste doel. De pilot van het CJIB is gericht op het innen van openstaande boetes en het voorkomen van schuldenoploop. Bij de meeste besproken pilots ligt de nadruk vooralsnog op het versterken van de handhaving. In de toekomst willen de onderzochte uitvoeringsorganisaties profilering meer gaan inzetten om burgers te ondersteunen. Tegelijkertijd kan het lastig zijn voor hen om dit financieel rond te krijgen.

Omgang met mensenrechtelijke risico's

In de gevoerde gesprekken viel het op dat risico's vooral werden benoemd in termen van ethische principes en specifieke wettelijke eisen, zoals opgenomen in de Algemene verordening gegevensbescherming. Er werd niet of nauwelijks verwezen naar mensenrechtenverdragen. Dat wil niet zeggen dat de onderzochte uitvoeringsorganisaties zich bij de ontwikkeling van hun profilerende algoritmes niet bewust zijn van mogelijke risico's. In de gesprekken worden zaken als privacy, transparantie, uitlegbaarheid en menselijke tussenkomst genoemd. Grofweg komen de genoemde kwesties overeen met de relevante mensenrechten waar we ons in dit rapport op richten: privacyrechten, gelijkheidsrechten, vrijheidsrechten en procedurele rechten. De nadruk lijkt vooral te liggen op privacyrechten en het recht op non-discriminatie. De onderzochte uitvoeringsorganisaties beoordelen de inzet van profilerende systemen als proportioneel. Ze zien risico's, maar concluderen dat die voldoende te mitigeren zijn en dat er voldoende baten tegenover staan.

Ethische en wettelijk eisen

De maatregelen die de onderzochte organisaties nemen, dienen twee doelen. Enerzijds zijn ze gericht op het zo goed mogelijk toepassen van bestaande juridische kaders. De organisaties ervaren soms onduidelijkheid over hoe die gehanteerd dienen te worden en proberen met de maatregelen de vertaalslag te maken van wet naar praktijk. Anderzijds investeren de organisaties in de ontwikkeling van ethische expertise, vanuit een verlangen om niet alleen binnen de grenzen van de wet te handelen, maar ook te voldoen aan wat maatschappelijk wenselijk is. De maatregelen dienen de organisatie dus ook te stimuleren verder te kijken dan de wettelijke vereisten.

Breed scala van maatregelen

De onderzochte organisaties werken aan een scala van maatregelen om in lijn te werken met mensenrechten. Voorbeelden zijn het aanstellen van een ethisch adviseur en het oprichten van een ethische commissie. Ook worden normenkaders ontwikkeld, die ethische principes en specifieke wettelijke eisen koppelen aan verantwoordelijkheden van specifieke functies en medewerkers. Zo

worden algemene uitgangspunten concreet toepasbaar in de praktijk. Verder investeren de organisaties in ethische en juridische expertise, onder meer bij de ontwikkelteams en is bias-toetsing een standaard onderdeel geworden van de systeemontwikkeling. Opvallend is dat deze toetsing decentraal is georganiseerd. De onderzochte uitvoeringsorganisaties starten hun eigen samenwerkingen met externe partijen om bias te toetsen. De kennisontwikkeling vindt ook per organisatie plaats.

Ten aanzien van het recht op informatie zoeken de onderzochte uitvoeringsorganisaties vooral naar de balans tussen actieve en passieve informatievoorziening. Daarbij verschillen ze wat betreft de maatregelen die ze hebben genomen. Op het gebied van procedurele rechten nemen ze maatregelen om te zorgen voor menselijke tussenkomst en uitlegbaarheid. De organisaties verschillen van mening over de noodzaak de bestaande klachtenprocedures aan te passen.

Algoritmes leren afwegen

De organisaties zien dat er risico's bij de inzet van profilerende systemen kunnen ontstaan en nemen maatregelen om die zoveel mogelijk te verkleinen. Uitvoeringsorganisaties voelen zich ondanks deze risico's genoodzaakt om profilerende systemen te gebruiken, vanwege de maatschappelijke en politieke wensen om enerzijds effectief fraude op te sporen en anderzijds te zorgen voor een soepele dienstverlening.

De onderzochte organisaties zijn bij het nemen van maatregelen op zoek naar houvast: welke maatregelen zorgen ervoor dat zij zowel in wettelijk als in ethisch opzicht voldoen aan de gerechtvaardigde verwachtingen van politiek en samenleving? Steeds meer sijpelt bij de onderzochte uitvoeringsorganisaties het besef door dat ze niet alleen moeten handelen binnen en volgens de bedoeling van de wet, maar ook maatschappelijk onwenselijke toepassingen moeten vermijden, die wettelijk wel zijn toegestaan.

De waarde van de genomen maatregelen zal zich in de komende tijd moeten bewijzen. De opdracht is om nut en noodzaak van profilerende algoritmen adequaat te leren afwegen. Uiteindelijk is dit een gedeelde taak van uitvoeringsorganisaties, politici en andere beleidsmakers. Daarbij zal het steeds moeten gaan over hoe maatschappelijke doelen het beste kunnen worden bereikt en de rechten van burgers het beste worden beschermd.

Aanbevelingen

Op basis van het onderzoek formuleert het Rathenau Instituut zeven aanbevelingen. De aanbevelingen vloeien voort uit de inventarisatie van

maatregelen bij de drie door ons gesproken uitvoeringsorganisaties, maar dienen als uitgangspunt voor elke uitvoeringsorganisatie die profilerende algoritmen ontwikkelt of inzet.

1. Zorg ervoor dat juridische vereisten en ethische beginselen beter hanteerbaar worden

- Het is zaak dat iedere uitvoeringsorganisatie een normenkader opstelt, dat specificeert welke mensenrechtelijke aspecten bij welke functies zijn belegd.
- Controleer als uitvoeringsorganisatie of het gebruikte normenkader volledig is en houd daarbij oog voor brede afwegingen die relevant zijn voor de beoordeling van mensenrechtelijke risico's. Maak gebruik van algemene hulpmiddelen zoals IAMA, CODIO, het toetsingskader voor risicoprofielen van het College voor de Rechten van de Mens en de handreiking non-discriminatie. Voer deze controle periodiek uit.
- Stel als uitvoeringsorganisaties een overleg in over de ontwikkelde normenkaders om te komen tot *best practices*. Het ministerie van Binnenlandse Zaken en Koninkrijksrelaties kan hierin een coördinerende en stimulerende rol spelen.

2. Stimuleer ethische expertise in de gehele organisatie

- Zorg als uitvoeringsorganisatie dat ontwikkelteams beschikken over de relevante juridische en ethische vaardigheden. Investeer in interne opleidingen en train alle medewerkers in de organisatie.
- Stel een ethisch adviseur in om de interne expertise-ontwikkeling te versterken en een doordacht gebruik van normenkaders te ondersteunen.
- Zorg voor een gezamenlijk overleg tussen uitvoeringsorganisaties over hoe zij hun ethische commissie willen invullen en waar ze kunnen leren van elkaars ervaringen. Iedere invulling kent namelijk voor- en nadelen.
- Richt als uitvoeringsorganisatie een burgerpanel op waarin de gevolgen van algoritmische systemen worden bekeken. Dat ondersteunt een meer voldragen risicobeoordeling.

3. Blijf investeren in privacytoetsing

- Zorg als uitvoeringsorganisaties dat de functionaris gegevensbescherming zo gepositioneerd is dat hij doortastend toezicht kan houden op profilerende systemen.
- Blijf het gevoerde privacybeleid kritisch evalueren en waar nodig aanscherpen.

4. **Versterk de manier waarop biastoetsing plaatsvindt**
 - Geef als uitvoeringsorganisatie meer inzicht in hoe biastoetsing plaatsvindt.
 - Zet een nationaal kennisplatform voor biastoetsing op waar expertise kan worden ontwikkeld en gedeeld. Bepaal welke mate van standaardisatie gewenst is en of nadere wettelijke eisen nodig zijn.
5. **Investeer in een actieve informatievoorziening**
 - Versnel de ontwikkeling en het gebruik van een algoritmeregister.
 - Ontwikkel de informatievoorziening in samenspraak met representatieve burgerpanels.
 - Ontwikkel als overheid helder en consistent taalgebruik over algoritmes.
6. **Versterk de waarborgen voor procedurele rechten**
 - Zorg als uitvoeringsorganisatie ervoor dat de klachtenprocedure bij een nieuw profilerend systeem opnieuw wordt bekeken en zo nodig verbeterd.
 - Verhelder als wetgever de eisen aan de uitlegbaarheid van beleidsondersteunende lerende algoritmes.
 - Overweeg als wetgever om nadere wettelijke eisen te stellen aan het toepassen van lerende systemen. Kijk in het bijzonder naar de toelaatbaarheid van *unsupervised learning*.
7. **Voorkom verkoking bij de ontwikkeling en beoordeling van algoritmische toepassingen**
 - Zorg dat binnen en buiten uitvoeringsorganisaties ook de zogenoemde nuloptie op tafel ligt bij het toetsen van de proportionaliteit van algoritmische profilering.
 - Zorg dat de effectiviteit van de profilerende algoritmes op een vergelijkbare manier wordt geëvalueerd en centraal bekend is.
 - Verken als uitvoeringsorganisatie profilerende toepassingen die burgers meer ondersteuning bieden. Het ministerie van Binnenlandse Zaken en Koninkrijksrelaties kan deze ontwikkeling stimuleren.

Inhoud

Voorwoord	3
Samenvatting	4
Inleiding	13
1.1 De algoritmische uitvoeringspraktijk	13
1.1 Doel, onderzoekopzet en centrale begrippen	15
1.2 Aanpak	17
2 Mensenrechtelijke aspecten van het gebruik van profilering door uitvoeringsorganisaties	20
2.1 Privacy	20
2.2 Gelijkheidsrechten	23
2.3 Vrijheidsrechten	25
2.4 Procedurele rechten	26
2.5 AI-verordening	28
3 Maatregelen bij UWV, SVB, CJIB	29
3.1 De inzet van profilerende algoritmes	29
3.2 Gesignaleerde risico's	33
3.2.1 Een inventarisatie van de belangrijkste maatregelen	33
3.2.2 Functies	33
3.2.3 Expertise	37
3.2.4 Hulpmiddelen	42
3.2.5 Werkprocessen	44
3.2.6 Technische maatregelen	48
3.3 Blik op de toekomst	49
3.4 Samenvattend overzicht	52
4 Analyse	55
4.1 Inleiding	55
4.2 Bevindingen	56
5 Conclusies en aanbevelingen	66
5.1 Inleiding	66
5.2 Maatregelen voor profilerende systemen	67

5.3 Algoritmes leren afwegen	70
Literatuurlijst	82
Bijlage 1: interviews.....	86
Bijlage 2: interviewprotocol	88
Bijlage 3: externe consultaties	90

Begrippenlijst

Algoritme / algoritmisch systeem. Een algoritme of een algoritmisch systeem verwijst naar een set instructies die worden ingezet om bepaalde problemen op te lossen. Algoritmes kunnen in computercode worden geprogrammeerd. Er bestaan vele verschillende soorten algoritmes.

Rule-based algoritme. Een algoritme dat volledig door mensen wordt geschreven. Het bevat een reeks 'als-dan instructies', en wordt daarom ook wel een 'als-dan algoritme' genoemd. Dit soort algoritmes zijn in theorie inzichtelijk voor mensen, maar in de praktijk vaak complex en omvangrijk. Dit type systemen wordt door de overheid veel gebruikt om bepaalde wettelijke bepalingen geautomatiseerd uit te voeren.

Lerend algoritme. Een algoritme dat leert om bepaalde problemen op te lossen. Op basis van grote hoeveelheden data leert het algoritme patronen of verbanden te ontdekken. Dit wordt ook *machine learning* of een *case based* algoritme genoemd. Er bestaan verschillende soorten lerende algoritmes. Lerende algoritmes worden meestal niet gebruikt voor geautomatiseerde besluiten, maar ter voorbereiding of ondersteuning van een besluit.

Supervised learning. Verwijst naar lerende algoritmes die onder begeleiding van mensen nieuwe regels afleiden. Deze systemen worden met name getraind door data te labelen. Zo leren programmeurs hoe het algoritme met aangereikte voorbeelden moet omgaan.

Unsupervised learning. Verwijst naar lerende algoritmes die zonder begeleiding van mensen nieuwe regels afleiden. Deze systemen leren hoe ze onderscheid kunnen maken in de aangeleverde data. De data worden niet gelabeld. Deze algoritmes zijn het meest complex en minst inzichtelijk.

Profilerende systemen. Lerende systemen die profielen opstellen om gedrag van individuen te voorspellen, doorgaans door het koppelen van verschillende (persoons-)gegevens uit verschillende databronnen. De systemen dienen verschillende doelen, bijvoorbeeld om gericht ondersteuning te bieden aan burgers, of om gericht fraude te detecteren. De systemen voeren niet geautomatiseerd besluiten door.

Afkortingen

AI: Artificial Intelligence

AVG: Algemene verordening gegevensbescherming

DPIA: Data Protection Impact Assessment

IAMA: Impact Assessment Mensenrechten en Algoritmes

FG: Functionaris gegevensbescherming

UWV: Uitvoeringsinstituut Werknemersverzekeringen

SVB: Sociale Verzekeringsbank

CJIB: Centraal Justitieel Incassobureau

Inleiding

1.1 De algoritmische uitvoeringspraktijk

Uitvoeringsorganisaties zijn de afgelopen dertig jaar in vergaande mate gedigitaliseerd. Verkeersboetes worden automatisch geïnd, de kinderbijslag wordt automatisch gestort en belastingformulieren worden automatisch ingevuld. De gedachte achter deze grootschalige digitalisering is helder: met behulp van algoritmes zou de slagkracht en efficiëntie van uitvoeringsorganisaties sterk vergroot kunnen worden. Zo kan de politie met automatisch cameratoezicht veel meer auto's flitsen en boetes uitschrijven. In 2020 werd elke 11 seconden een hardrijder geflitst en waren er 6,4 miljoen boetes voor hardrijden (RTL Nieuws 2021). Het gaat in deze algoritmische uitvoeringspraktijk grotendeels om systemen die wettelijke voorschriften geautomatiseerd kunnen uitvoeren.

De kinderopvangtoeslagaffaire en andere misstanden maakten duidelijk dat het streven naar digitale efficiëntie een donkere kant kan hebben: de overheid heeft de rechten van burgers ernstig geschonden (Tweede Kamer 2020).¹ Het gebruik van algoritmes door de overheid ligt daarom zowel in de maatschappij als in de politiek onder een vergrootglas. Daarbij staan de algoritmes niet op zichzelf. Ze zijn ingebed in een organisatie en vormen een *algoritmische praktijk*, met onder andere een bepaalde technologie, expertise, organisatiestructuur en informatiebeleid (Grimmelikhuijsen, S. & Meijer, A., 2020). Als die praktijk als geheel onvoldoende waarborgen kent, en belangenafwegingen onzorgvuldig gemaakt worden, brengt hij grote mensenrechtelijke risico's met zich mee.

Een lange reeks onderzoeken in het afgelopen decennium wijst op structurele problemen bij de ontwikkeling en toepassing van ICT-systemen bij de overheid. Zo wordt een digitaal systeem vaak primair vanuit efficiëntie voor de eigen bedrijfsvoering ontworpen (Widlak 2021; Rathenau Instituut 2011). Het blijft lastig om de belangen van burgers minstens zo zwaar mee te laten wegen als de belangen van de eigen organisatie (Rathenau Instituut 2010; Rathenau Instituut 2012). Dat leidt bijvoorbeeld tot gebruiksonvriendelijke systemen, een omkering van de bewijslast en onduidelijke verantwoordelijkheden (Widlak 2021; WRR 2011; TCU, 2021; POK, 2020; Ombudsman 2012 & 2016; Widlak & Peeters, 2018).

¹Zie ook de uitspraak over het Systeem Risico Indicatie (SyRI): Rechtbank Den Haag 5 februari 2020.

De opkomst van lerende systemen verdiept deze zorgen. Want hoe leren de systemen precies, en valt dat voldoende uit te leggen aan burgers en de rechter? Hoe valt te controleren dat er geen bias in de systemen ontstaat, en te waarborgen dat groepen burgers niet worden gediscrimineerd? Weet een burger wanneer lerende systemen worden gebruikt, wat daarvan het effect is, en of hij daartegen in verweer kan komen?

In de zoektocht naar verdere efficiëntie en verbetering van de dienstverlening zijn overheden aan het experimenteren met lerende systemen. Het vorige kabinet heeft diverse onderzoeken laten uitvoeren om meer zicht te krijgen op mensenrechtelijke risico's van dit soort systemen (Vetzo et al. 2018; Rathenau Instituut, 2017; Kamerstukken II 2019-2020. 26643, nr. 642; Kamerstukken II 2020-2021. 26643, nr. 765; Kulk & van Deursen, 2020). Er zijn risico's geconstateerd met betrekking tot privacy, gelijke behandeling, eerlijkheid en maatwerk, procedurele rechten en vrijheidsrechten. Bovendien hebben Hoge Colleges van Staat in 2021 gewezen op bestuurlijke en staatsrechtelijke risico's bij digitale overheidsbesluitvorming (RvS 2021, ARK 2021). Ook de Raad voor het Open Bestuur deed dit (ROB 2021). Ze publiceerden toetsingskaders om de wet beter toepasbaar te maken en mensenrechtelijke risico's in een vroeg stadium te identificeren en te verkleinen.

Ook andere partijen, nationaal en internationaal, ontwikkelden richtlijnen om organisaties te helpen bij het beschermen van mensenrechten (bijvoorbeeld de *ethische richtlijnen voor betrouwbare AI*, Deskundigengroep op hoog niveau inzake kunstmatige intelligentie, 2019). Deze bieden een leidraad voor hoe organisaties al tijdens het ontwerp van algoritmische systemen rekening kunnen houden met mensenrechten. In Nederland verscheen in 2021 onder meer het Impact Assessment Mensenrechten en Algoritmes en de handreiking non-discriminatie (IAMA, Gerards et al. 2021, van der Sloot et al. 2021). Daarnaast wordt in publicaties als de *Code goed digitaal openbaar bestuur* de verbinding gelegd tussen digitalisering, mensenrechten en principes van behoorlijk bestuur (Meijer en Ruijter, 2021).

Er is echter weinig zicht op hoe uitvoeringsorganisaties momenteel hun organisatie inrichten om bij de inzet van algoritmische systemen mensenrechtelijke risico's te identificeren, en zoveel mogelijk te verkleinen. Welke werkprocessen hebben organisaties hiervoor opgetuigd? Bij wie zijn verantwoordelijkheden belegd? Welke expertise zijn zij aan het ontwikkelen? Gebruiken uitvoeringsorganisaties de bovengenoemde hulpmiddelen of ontwikkelen ze hun eigen instrumenten? Kortom, welke waarborgen worden in de uitvoeringspraktijk bij profilerende systemen aangebracht?

Om deze vragen te beantwoorden, heeft het ministerie van Binnenlandse Zaken en Koninkrijksrelaties het Rathenau Instituut opdracht gegeven deze studie uit te voeren. Omdat profilerende systemen in opkomst zijn en mensenrechtelijke risico's vergroten, richt dit onderzoek zich op deze technologie.

1.1 Doel, onderzoeksopzet en centrale begrippen

Het doel van dit onderzoek is om te verkennen hoe uitvoeringsorganisaties hun organisatie inrichten om mensenrechtelijke risico's van profilerende systemen te voorkomen.

De onderzoeksvraag luidt daarom als volgt:

Welke maatregelen nemen uitvoeringsorganisaties om bij de ontwikkeling en het gebruik van profilerende systemen in lijn te werken met mensenrechten, en hoe kunnen ze dit versterken?

De drie deelvragen zijn:

1. Voor welke taken zetten uitvoeringsorganisaties profilerende systemen in? En welke risico's op mensenrechtenschending zien zij hierbij ontstaan?
2. Welke maatregelen nemen uitvoeringsorganisaties momenteel om risico's te vermijden? Welke versterking achten zij nodig in de toekomst?
3. Waar liggen kansen voor verbetering? Wat kunnen uitvoeringsorganisaties van elkaar leren?

De deelvragen bestrijken twee onderzoeksfases. Voor het beantwoorden van vraag 1 en 2 inventariseren we bij uitvoeringsorganisaties welke mensenrechtelijke vraagstukken leven, welke maatregelen ze nu nemen en welke ze in de toekomst nodig achten. Vervolgens komen we, met behulp van de inzichten van externe experts, tot een eigen analyse en antwoord op de onderzoeksvraag (zie verder paragraaf 1.3).

De hoofdvraag bevat twee centrale begrippen: maatregelen en profilerende systemen. We lichten deze begrippen en hoe we ze zullen gebruiken in dit onderzoek nader toe.

Met maatregelen bedoelen we de verschillende manieren waarop uitvoeringsinstanties hun organisaties inrichten om in lijn te werken met mensenrechten. We kijken in dit onderzoek naar vier soorten maatregelen: welke

werkprocessen, functies, expertise en hulpmiddelen gebruikt of ontwikkelt de organisatie?

Met werkprocessen doelen we op de stappen die het algoritmisch systeem in de organisatie doorloopt. Legt een uitvoeringsorganisatie systemen altijd voor aan burgerpanels voordat ze die structureel in gebruik neemt? Zorgt ze structureel voor diversiteit in ontwikkelteams? Is er een governancestructuur afgesproken waarin verantwoordelijkheden en toezicht zijn belegd?

Functies verwijzen naar de posities en organen die zijn gecreëerd om mensenrechten te beschermen, zoals een functionaris gegevensbescherming, een ethisch adviseur of een ethische commissie.

Onder expertise verstaan we de kennis en vaardigheden van mensen bij uitvoeringsorganisaties in relatie tot de mensenrechtelijke risico's van algoritmische systemen. Weten medewerkers genoeg over privacy, biasdetectie of *fairness*, en hoe ondersteunt de organisatie het vermogen van medewerkers om zorgvuldige afwegingen te maken?

Met de term hulpmiddelen verwijzen we naar de beschikbare richtlijnen, handreikingen, codes en andere tools die medewerkers van uitvoeringsorganisaties helpen om mensenrechtenkwesties in beeld te krijgen en recht te doen.

Deze soorten maatregelen kunnen onderling overlappen. Ze zijn gekozen om zo breed mogelijk te inventariseren wat uitvoeringsorganisaties doen en kunnen doen aan mensenrechtenbescherming.

Met profilerende systemen doelen we op lerende systemen die op basis van data-analyse kenmerken van burgers in kaart brengen (het profiel). Daarmee kunnen uitvoerende functionarissen mensen selecteren die in aanmerking komen voor nadere dienstverlening of onderzoek (ARK 2021, TNO 2019; TNO 2021). Daarbij gaat het dan bijvoorbeeld om: 'kansrijk om het preventie-instrument telefonisch innen in te zetten', 'kansrijk om nader hulp te geven bij het solliciteren' of 'kansrijk om nader onderzoek te doen of er sprake is van fraude'. De profielen dienen vaak niet om automatisch besluiten te nemen, maar ondersteunen de uitvoerende functionarissen bij het maken van analyses en het nemen van besluiten (ARK 2021).

Deze profilerende systemen verdiepen diverse mensenrechtelijke risico's. Zo is het moeilijk uit te leggen hoe de lerende algoritmes precies werken, is bias lastig

op te sporen, en is slecht voorspelbaar wat het algoritme precies leert (ARK 2021; Considerati 2020). Hoofdstuk 2 gaat op deze punten uitgebreider in.

Omdat het onderzoek zich toespitst op de verschillende maatregelen, verwachten we dat de bevindingen ook relevant zijn voor de ontwikkeling en het gebruik van andere typen algoritmische systemen.

1.2 Aanpak

Voor deze studie maakten we gebruik van een combinatie van deskresearch, interviews en externe consultaties.

Op basis van deskresearch maakten we een selectie van de mensenrechtelijke risico's waar we in deze studie naar hebben gekeken. We keken daarvoor naar recente wetenschappelijke literatuur, met name over de relatie tussen mensenrechtenbescherming en de toepassing van algoritmes door de overheid. Ook zijn beleidsdocumenten, rapporten van toezichthouders en journalistieke publicaties in de studie meegenomen.

Vervolgens voerden we gesprekken bij drie uitvoeringsorganisaties om een beeld te krijgen van het type maatregelen dat zij in relatie tot profilerende systemen en mensenrechten hebben genomen, en willen nemen. We vroegen hen naar documentatie over de genomen maatregelen, zoals vastgelegde werkprocessen en verantwoordelijkheden, of gebruikte hulpmiddelen.

Om de opbrengst van de interviews met elkaar te kunnen vergelijken, kozen we ervoor de gesprekken te houden bij drie landelijke uitvoeringsorganisaties, die:

1. profilerende systemen inzetten waarbij mensenrechtelijke risico's spelen;
2. aangeven actief maatregelen te nemen om deze risico's te mitigeren; en
3. mee willen werken aan het onderzoek.

De gekozen organisaties zijn: het Centraal Justitieel Incassobureau (CJIB), de Sociale Verzekeringsbank (SVB) en het Uitvoeringsinstituut Werknemersverzekeringen (UWV).

Bij deze organisaties hebben we drie of vier semigestructureerde interviews afgenomen bij medewerkers van verschillende afdelingen. Naderhand hebben we aan een deel van hen vervolgvragen gesteld. Zo kregen we maatregelen die bij verschillende functies belegd zijn in beeld, alsook mogelijke perspectieven op die

maatregelen. Omdat ze van groot belang zijn voor de bescherming van mensenrechten tijdens de ontwikkeling en het gebruik van profilerende systemen, hebben we ernaar gestreefd om bij elke organisatie te spreken met iemand van het ontwikkelteam, de functionaris gegevensbescherming en een directielid of bestuurder. Helaas wilden twee personen die we graag wilden interviewen geen medewerking verlenen aan het onderzoek. Bij het CJIB verleende de functionaris gegevensbescherming geen medewerking, daar is met de *privacy officer* gesproken. Bij de SVB verleende het bestuur geen medewerking. In totaal zijn 13 personen geïnterviewd (zie bijlage 1).

De interviews zijn uitgeschreven en omgezet in gespreksverslagen die ter goedkeuring aan de geïnterviewden zijn voorgelegd. Ook de opgenomen citaten zijn ter goedkeuring aan hen voorgelegd. In bijlage 2 is het interviewprotocol opgenomen. De interviews, de overlegde documentatie en een aantal vervolgvragen vormen de basis van hoofdstuk 3.

Vervolgens hebben we twee externe consultaties georganiseerd om de opgehaalde informatie te bespreken met externe experts. Daarbij presenteerden we de voorlopige bevindingen, alsook suggesties voor mogelijke versterkingen van de mensenrechtenbescherming. De externe experts hebben we gevraagd in hoeverre zij het gepresenteerde beeld deelden, en welke aanvullende zaken en knelpunten aandacht vereisen. Voor de consultatie hebben we vertegenwoordigers van maatschappelijke organisaties, wetenschappers, uitvoeringsorganisaties en toezichthouders uitgenodigd. Ook vertegenwoordigers van de drie geïnterviewde organisaties waren aanwezig. Daardoor konden de aanwezigen zowel elkaars uitspraken als de voorlopige bevindingen bevragen en weerleggen. De opdrachtgever was als toehoorder aanwezig. Zie bijlage 3 voor de aanwezigen en de opzet van de consultatie. De analyse van de interviews en consultaties vormen de basis van hoofdstuk 4 en 5.

Het integrale rapport is niet ter goedkeuring voorgelegd aan de gesproken organisaties, om de onafhankelijkheid van het onderzoek en de integriteit van de gevoerde gesprekken te waarborgen.

Afbakening

Dit onderzoek inventariseerde niet welke algoritmes uitvoeringsorganisaties precies gebruiken. We keken niet in hun digitale systemen. Dit onderzoek richtte zich op de maatregelen die organisaties nemen om in lijn te werken met mensenrechten. Ook beperkt dit onderzoek zich tot specifieke posities die van bijzonder belang zijn voor het beschermen van mensenrechten bij de ontwikkeling van profilerende systemen, zoals een functionaris gegevensbescherming of

ethisch adviseur. Hun observaties vormden de basis voor de bevindingen die tijdens de externe consultaties zijn gepresenteerd.

1.3 Leeswijzer

Naast de inleiding telt dit rapport vier hoofdstukken. In hoofdstuk 2 worden de mensenrechten die op het spel staan bij de inzet van profilerende algoritmes door uitvoeringsorganisaties geïntroduceerd en toegelicht. Hierbij worden ook de verbanden gelegd tussen mensenrechten en andere principes en normen, zoals de algemene beginselen van behoorlijk bestuur. In hoofdstuk 3 bespreken we de bij de drie uitvoeringsorganisaties gevoerde gesprekken: hoe brengen het UWV, de SVB en het CJIB algoritmische profilering in de praktijk, en welke maatregelen achten ze nodig om daarbij de mensenrechten te beschermen? In hoofdstuk 4 analyseren we deze resultaten, mede aan de hand van de bevindingen uit de externe consultaties. In hoofdstuk 5 presenteren we onze conclusies en aanbevelingen.

2 Mensenrechtelijke aspecten van het gebruik van profilering door uitvoeringsorganisaties

Wanneer overheden burgers profileren, kunnen mensenrechten in het geding komen. Maar welke precies? En hoe raakt dat burgers? Dat lichten we in dit hoofdstuk nader toe aan de hand van vier clusters van mensenrechten: privacyrechten, gelijkheidsrechten, vrijheidsrechten en procedurele rechten. Deze indeling nemen we over van de studie *Algoritmes en grondrechten*, en het daarna ontwikkelde Impact Assessment Mensenrechten en Algoritmes (Gerards et al. 2018, Min. BZK 2021). Per cluster bespreken we op hoofdlijnen welke mensenrechten voor uitvoeringsorganisaties het meest relevant zijn.

2.1 Privacy

Doordat de overheid steeds meer persoonlijke gegevens over burgers verzamelt, is het in acht nemen van privacy steeds belangrijker geworden bij de ontwikkeling van profilerende algoritmische systemen. Het fundament van het recht op privacy is gelegd door artikel 8 van het *Europees Verdrag tot bescherming van de rechten van de mens en de fundamentele vrijheden* (EVRM). Hierin staat: 'een ieder heeft het recht op respect voor zijn privéleven, zijn familie- en gezinsleven, zijn woning en zijn correspondentie'. Deze verdragsbepaling is de afgelopen decennia in rechtszaken en wetten ingevuld en verder geconcretiseerd. Zo is in het *Handvest van de grondrechten van de Europese Unie* ook een recht op dataprotectie vastgelegd, dat deels overlapt met het recht op privacy. Dit recht bepaalt dat de verwerking van persoonsgegevens alleen mag plaatsvinden op basis van een aantal principes. Dit recht op gegevensbescherming speelt in dit rapport een centrale rol.

De vereisten voor deugdelijke gegevensbescherming zijn uitgewerkt in de *Algemene verordening gegevensbescherming* (AVG). Zo mogen persoonsgegevens alleen worden verzameld met een gerechtvaardigd doel en moet de verzamelende organisatie transparant zijn over hoe en waarom ze persoonsgegevens verwerkt. Ook moet de gegevensverwerking op een passende manier worden beveiligd. Daarnaast gelden extra strenge regels voor gegevens over bijvoorbeeld gezondheid of geloofsovertuiging.

We bespreken de centrale beginselen van de AVG, en geven aan waarom juist deze relevant zijn voor uitvoeringsorganisaties. We noemen achtereenvolgens: doelbinding, minimale gegevensverwerking, beperkte opslag, juistheid, integriteit, en vertrouwelijkheid. Bij de toetsing van deze beginselen is het niet alleen zaak om per systeem te kijken of aan de AVG wordt voldaan, maar ook om te bezien wat het totale effect van een reeks systemen is op de privacy van burgers, en of die impact gewenst is. In hoofdstuk 4 komen we op dit punt terug.

Doelbinding

Overheidsinstanties moeten zich bij het verzamelen van gegevens houden aan doelbinding: persoonsgegevens mogen alleen verwerkt worden indien er sprake is van: 'welbepaalde, uitdrukkelijk omschreven doeleinden'. Hierbij kun je denken aan het verzamelen van een e-mailadres om informatie naartoe te sturen die de burger nodig heeft. Op zijn beurt moet dit doel weer gekoppeld zijn aan een wettelijke grondslag: een uitvoeringsorganisatie voert een wettelijke taak uit, en daar horen bepaalde doelen bij.

Minimale gegevensverwerking

Dit beginsel schrijft voor dat, gegeven een bepaald doel, er niet meer persoonsgegevens verzameld of verwerkt mogen worden dan noodzakelijk zijn om het geformuleerde doel te bereiken. De persoonsgegevens moeten: 'toereikend zijn, ter zake dienend en beperkt tot wat noodzakelijk is voor de doeleinden waarvoor zij worden verwerkt'.

Beperkte opslag

De uitvoeringsorganisatie moet zich ook houden aan opslagbeperking. De persoonsgegevens moeten: 'worden bewaard in een vorm die het mogelijk maakt de betrokkenen niet langer te identificeren dan voor de doeleinden waarvoor de persoonsgegevens worden verwerkt noodzakelijk is'. Als de uitvoeringsorganisatie de gegevens wel langer wil bewaren, voor bijvoorbeeld wetenschappelijk of historisch onderzoek, dan moeten er: 'passende technische en organisatorische maatregelen worden getroffen om de rechten en vrijheden van de betrokkene te beschermen'.

Juistheid

De persoonsgegevens die overheidsinstanties verzamelen, moeten juist en actueel zijn. 'Alle redelijke maatregelen moeten worden genomen om de persoonsgegevens die, gelet op de doeleinden waarvoor zij worden verwerkt, onjuist zijn, onverwijld te wissen of te rectificeren.'

Integriteit en vertrouwelijkheid

Daarnaast vraagt de AVG om passende technische en organisatorische maatregelen die ervoor zorgen dat persoonsgegevens op een veilige manier verwerkt worden. Dit betekent dat persoonsgegevens: 'onder meer beschermd zijn tegen ongeoorloofde of onrechtmatige verwerking en tegen onopzettelijk verlies, vernietiging of beschadiging'.

De AVG introduceert ook het recht van iemand om: 'niet te worden onderworpen aan een uitsluitend op geautomatiseerde verwerking, waaronder profilering, gebaseerd besluit waaraan voor hem rechtsgevolgen zijn verbonden of dat hem anderszins in aanmerkelijke mate treft' (art. 22 lid 1). De AVG omschrijft profilering als: '[...] elke vorm van geautomatiseerde verwerking van persoonsgegevens waarbij aan de hand van persoonsgegevens bepaalde persoonlijke aspecten van een natuurlijke persoon worden geëvalueerd, met name met de bedoeling zijn beroepsprestaties, economische situatie, gezondheid, persoonlijke voorkeuren, interesses, betrouwbaarheid, gedrag, locatie of verplaatsingen te analyseren of te voorspellen' (AVG art. 4).

In de uitvoeringspraktijk is er wat profilering betreft meestal geen sprake van besluiten die uitsluitend gebaseerd zijn op een geautomatiseerde verwerking, omdat op een bepaald moment een mens de uitkomsten van de verwerking analyseert en een besluit neemt. Profilerende systemen worden dus vaak ingezet als tussenstap. Ook is belangrijk om te vermelden dat artikel 22 uitzonderingen biedt, waaronder voor geautomatiseerde verwerkingen die bij wet worden toegestaan – mits is voorzien in passende beschermende maatregelen (art. 22 lid 2 sub b). Het kabinet overwoog in 2019 om de wettelijke waarborgen voor profilering te versterken, maar wacht inmiddels de komende Europese AI-verordening af (Kamerstukken 2019-2020, 26643, nr. 641). We bespreken de AI-verordening kort in paragraaf 2.5.

Vanwege de AVG hebben organisaties de afgelopen jaren een governance-model opgetuigd dat toeziet op het respecteren van gegevensbescherming. Zo moet er een *data protection impact assessment* (DPIA) worden afgenomen, dat getoetst moet worden door een functionaris gegevensbescherming (FG). Veelal zijn er *privacy officers* die de ontwikkelteams ondersteunen bij het rekening houden met de AVG-bepalingen.

Toch blijkt in de praktijk dat uitvoeringsorganisaties er niet altijd in slagen om deze beginselen te respecteren. Zo zette de politie tijdens de lockdowns camera's in om gedrag te monitoren, terwijl daar onvoldoende grondslag voor was, en die camera's daar niet voor bedoeld waren (De Groene Amsterdammer, 9 juni 2021). Een ander voorbeeld betrof de camera's die langs autowegen automatisch foto's

nemen van langsrijdende voertuigen om kentekens te controleren. Een deel van deze camera's bezat geavanceerde functies, waardoor ze ook de inzittenden konden fotograferen. Hoewel er geen juridische grondslag was om automobilisten en bestuurders herkenbaar in beeld te brengen, is dat de afgelopen vijf jaar wel gebeurd. De foto's zijn onder meer gebruikt voor opsporingsdoeleinden en strafrechtelijke onderzoeken (NRC, 3 augustus 2021).

Een ander voorbeeld is SyRI, het Systeem Risico Indicatie. Dit is een systeem dat vanaf 2014 werd ingezet om fraude en misbruik met uitkeringen, toeslagen en belastingen op te sporen. Het systeem maakte een risicobeoordeling op basis van verschillende databronnen die gekoppeld werden, waaronder die van onder andere gemeenten, de Belastingdienst en uitvoeringsorganisaties. De risicomeldingen werden vervolgens doorgegeven aan overheidsinstanties die om een analyse hadden gevraagd. De Rechtbank Den Haag oordeelde dat het systeem niet voldeed aan artikel 8 van het Europees Verdrag voor de Rechten van de Mens (EVRM).

Deze misstanden laten zien dat ook bij de inzet van profilering het belangrijk is om te waken voor privacyschendingen.

2.2 Gelijkheidsrechten

Ook gelijkheidsrechten kunnen bij profilering op het spel staan. Gelijkheidsrechten vloeien voort uit het principe dat elke burger gelijkwaardig is, en daarom voor de wet gelijk is en niet gediscrimineerd mag worden. Een burger mag dus niet uitgesloten worden van een regeling, of slechter behandeld worden, vanwege zijn huidskleur, geslacht of seksuele oriëntatie. Dit betekent niet dat mensen altijd volstrekt gelijk moeten worden behandeld. Het is vaak logisch om ongelijke gevallen ook ongelijk te behandelen, als er op die manier maatwerk geleverd wordt. Zo kan iemand die laaggeletterd is recht hebben op overheidshulp bij het begrijpen van overheidsinformatie, terwijl deze hulp niet, of niet in dezelfde mate beschikbaar is voor mensen die goed kunnen lezen en schrijven. Het gaat er uiteindelijk om dat mensen in gelijke gevallen gelijk behandeld worden (Grondwet art. 1).

De volgende gelijkheidsbeginselen en -normen zijn in het bijzonder relevant voor uitvoeringsorganisaties: gelijkheid voor de wet, verbod van direct onderscheid op bepaalde gronden, verbod van indirect onderscheid op bepaalde gronden, materieel onderscheid/maatwerk en het verbod van profilering. We bespreken ze ieder kort.

Gelijkheid voor de wet

Burgers hebben recht op een gelijke toepassing van algemene wetgeving, waarbij geen willekeur plaatsvindt en er voldaan wordt aan het consistentievereiste en het rechtszekerheidsbeginsel.

Verbod van direct onderscheid op bepaalde gronden

Beslissingen of regels mogen niet in doorslaggevende mate zijn ingegeven door of gebaseerd op beschermde persoonskenmerken, zoals geslacht of seksuele oriëntatie.

Verbod van indirect onderscheid op bepaalde gronden

Beslissingen of regels mogen niet disproportioneel benadelend uitpakken voor personen die behoren tot groepen met beschermde persoonskenmerken. Deze norm geldt voor situaties waarin niet op verboden gronden wordt geselecteerd, maar een bepaalde groep toch nadelige gevolgen ondervindt, bijvoorbeeld omdat er gekeken wordt naar een bepaald postcodegebied waar een etnische groep sterk vertegenwoordigd is.

Recht op materieel onderscheid/maatwerk

Er is een plicht om rekening te houden met verschillen tussen mensen en groepen. Dit beginsel kan op gespannen voet staan met de rechtszekerheid, die noopt verschillende burgers zo consistent mogelijk te behandelen. Er zal dus telkens een afweging nodig zijn tussen maatwerk en consistentie.

Ook gelijkheidsrechten zijn in de praktijk onder druk komen te staan. Zo maakte de Belastingdienst tussen 2012 en 2020 gebruik van de Fraude Signalerings Voorziening (FSV), een risicosignaleringssysteem waarin werd geregistreerd of er bij een persoon een vermoeden van fraude bestond. Dit gebeurde bijvoorbeeld wanneer iemand die dicht bij zijn werk woonde hoge reiskosten opgaf. Maar uit onderzoek van PwC bleek ook dat in tientallen gevallen een frauderegistratie was gebaseerd op nationaliteit of uiterlijk – wat strijdig is met het non-discriminatiebeginsel (PwC, 2021). De consequenties van een registratie in de FSV konden ernstig zijn. Zo konden mensen vanwege de registratie niet in aanmerking komen voor een persoonlijke betalingsregeling.

2.3 Vrijheidsrechten

Bij profilering moet ook rekening gehouden worden met vrijheidsrechten. Deze rechten dienen burgers de ruimte en gelegenheid te geven om onbelemmerd uitdrukking te geven aan hun identiteit en voorkeuren. Burgers moeten op autonome wijze keuzes kunnen maken, of het nu gaat om politieke standpunten of iemands seksuele oriëntatie, en deze keuzes in vrijheid kunnen delen met anderen. Hiervoor zijn tal van vrijheidsrechten nodig, waaronder het recht op demonstratie en het kiesrecht. In de context van algoritmische overheidsbesluitvorming en uitvoeringsorganisaties draait het vooral om het recht op informatie. Er is immers ook informatie nodig om tot een onderbouwde keuze te kunnen komen, bijvoorbeeld over het wel of niet aanvechten van een bepaald overheidsbesluit. Het recht op informatie draait om het passieve recht om toegang tot informatie te verkrijgen en het actieve recht om informatie te ontvangen.

Passief recht om toegang tot informatie te verkrijgen

Het passieve informatierecht houdt in dat mensen het recht hebben op toegang tot bepaalde informatie, bijvoorbeeld over de totstandkoming van een overheidsbesluit. Ze moeten de informatie dan wel zelf opzoeken, bijvoorbeeld door de website van een uitvoeringsorganisatie te raadplegen.

Actief recht om informatie te ontvangen

Het actieve recht om informatie te ontvangen, verlangt meer van de partij die over bepaalde informatie beschikt. Deze moet bepaalde informatie actief doorgeven aan de burger, bijvoorbeeld als het gaat om de motivatie om een bepaald besluit te nemen. De vraag is natuurlijk hoe je dit zo goed mogelijk vormgeeft, en, bijvoorbeeld, zo gecommuniceerd wordt dat burgers ook begrijpen welke informatie openbaar wordt gemaakt.

Deze rechten zijn uitgewerkt in verschillende wetten, zoals de Wet open overheid (WOO), en zijn ook verbonden met de vrijheid van meningsuiting (art. 10 EVRM). De WOO moet er onder andere voor zorgen dat overheidsinformatie beter vindbaar is, en legt overheden inspanningsverplichtingen op om bepaalde informatie actief openbaar te maken. Daar zijn verschillende redenen voor: informatie draagt bijvoorbeeld bij aan democratische verantwoording, zodat bijvoorbeeld burgers, journalisten en maatschappelijke organisaties kunnen zien en controleren wat de overheid doet.² In de volgende paragraaf gaan we nader in op dit controlerende mechanisme.

² Handreiking: Actief openbaar maken doe je zo! | Instrument | Rijksprogramma voor Duurzaam Digitale Informatiehuishouding

In de praktijk zullen uitvoeringsorganisaties dus zowel actief als passief informatie beschikbaar moeten stellen. Een voorbeeld waarin het informatierecht onder druk is komen te staan, is de toeslagenaffaire waarover in december 2020 de parlementaire ondervragingscommissie Kinderopvangtoeslag het verslag *Ongekend onrecht* publiceerde. De commissie concludeert dat de informatie die ouders kregen vaak niet voldoende was (POK, 2020, p.96). De Belastingdienst deed jarenlang onderzoek naar de aanvragen van ouders en liet in veel gevallen bij kleine fouten de kinderopvangtoeslag vervallen. Dit werd ook wel een zero-tolerance-aanpak genoemd. Tijdens deze handhavingsprocessen kregen ouders onvoldoende of zelfs geen informatie van de fiscus.

2.4 Procedurele rechten

Het is ook zaak bij profilering te letten op procedurele rechten. Deze rechten zorgen ervoor dat geschillen op een effectieve en objectieve manier kunnen worden opgelost en dat burgers bij de aantasting van hun individuele belangen bezwaar kunnen maken. Ze zijn noodzakelijk voor de controle op de uitoefening van de wetgevende en uitvoerende macht – burgers moeten zich tegen deze macht kunnen weren. Toegang tot een effectief rechtsmiddel is hierbij essentieel. Dit kan in de vorm van een rechter aan wie een zaak voorgelegd wordt, maar ook als klacht bij een toezichthouder.

Procedurele rechten zijn niet alleen van toepassing wanneer een besluit genomen is en de burger daarover wil klagen, maar gelden ook al in de besluitvormingsfase. Bepaalde algemene beginselen van behoorlijk bestuur, zoals transparantie, motivering en zorgvuldigheid van besluitvorming, worden ook beschouwd als procedurele rechten.

We bespreken de algemene beginselen van behoorlijk bestuur, waaronder ook de motiveringsplicht valt, het recht op *equality of arms*, het recht op hoor en wederhoor en de onschuldspresumptie.

De algemene beginselen van behoorlijk bestuur

De algemene beginselen van behoorlijk bestuur zijn in de loop der jaren in de rechtspraak ontstaan en deels opgenomen in de *Algemene wet bestuursrecht* (Awb). Zo behoort de overheid transparant te zijn en waar nodig informatie te geven, moet de besluitvorming zorgvuldig plaatsvinden, mag er geen sprake zijn van willekeur en moeten de bevoegdheden van overheden aan specifieke doelen zijn verbonden.

De motiveringsplicht

Een beginsel van behoorlijk bestuur dat in het bijzonder van belang is bij profilerende algoritmische systemen, is de motiveringsplicht: overheden dienen hun besluiten deugdelijk te onderbouwen. Dan weet de burgers immers waartegen hij bezwaar kan maken, en kan op politiek niveau een discussie gevoerd worden over de werking van bepaalde wetten. Deze motiveringsplicht zit ook deels verankerd in de AVG, die overheden verplicht in sommige gevallen algoritmische besluiten aan burgers uit te leggen.

Het recht op *equality of arms*

Dit recht verwijst naar het vermogen van burgers om vanuit een gelijkwaardige positie beslissingen van de overheid aan te vechten. Zo moeten burgers toegang krijgen tot alle relevante documenten en genoeg tijd hebben om zich voor te bereiden. De doorgaans sterke positie van de overheid moet gebalanceerd worden door de burger adequate faciliteiten te bieden.

Het recht op hoor en wederhoor

Als overheden besluiten nemen over burgers, hebben burgers het recht hierover geraadpleegd te worden. Wat vinden zij van het besluit dat hen treft? Dit recht borgt dat het perspectief van burgers bij besluiten die hen raken niet genegeerd wordt.

Dat algoritmische overheidssystemen in de praktijk op gespannen voet kunnen staan met de motiveringsplicht, blijkt uit de uitspraken van de Raad van State over het Aerius-systeem. Dit geautomatiseerde rekenprogramma berekent onder meer de uitstoot van stikstof, zodat op basis daarvan vergunningen verleend kunnen worden. In 2017 oordeelde de Raad dat de totstandkoming van de besluitvorming met het systeem ondoorzichtig was, waardoor een ongelijkwaardige procespositie kon ontstaan (Raad van State 2017). De Raad gaf aan dat de gemaakte keuzes, gebruikte gegevens en aannames van het systeem ‘volledig, tijdig en op passende wijze’ beschikbaar moeten worden gesteld (Raad van State 2017).³ De uitspraak en latere verfijning van de Raad van State passen in de huidige discussie of nadere wettelijke normering van algoritmische besluitvorming nodig is (zie bijvoorbeeld Wolswinkel 2021).

³ In 2018 verfiende de Raad van State deze uitspraak. Het Aerius-systeem maakt gebruik van standaardinvoergegevens en maatwerkinvoergegevens, die de gebruiker zelf invoert. De Raad van State geeft aan dat de overheid inzichtelijk moet maken op welke maatwerkgegevens een besluit is gebaseerd, op basis van de Algemene wet bestuursrecht. De standaardgegevens hoeven niet zonder meer uit eigen beweging te worden overlegd.

Voor op regels gebaseerde algoritmische systemen geldt dat de algoritmes en de gemaakte keuzen door mensen zijn geprogrammeerd, en daarmee in principe voor mensen bekend en te begrijpen zijn. Dat vergt wel een zorgvuldig gedocumenteerd ontwerpproces, waarin gemaakte keuzes, gebruikte data en aannames te traceren zijn. Wetenschappers en beleidsmakers experimenteren met innovatieve vormen om gemaakte keuzes, en de vertaling van wet naar systeem te verbeteren (zoals wendbaar wetgeven, of Calculemus-Flint).

Bij lerende systemen, waaronder vaak profilerende systemen, zijn de gemaakte keuzes op een fundamenteel niveau moeilijker navolgbaar. Dit betekent niet dat er geen mensen betrokken zijn bij de besluitvorming door systemen, of de systeemontwikkeling. Softwareontwikkelaars leren het systeem patronen te herkennen met behulp van trainingsdata, kunnen 'onder de motorkap' kijken en het algoritme aanpassen. Toch geeft dat beperkt inzicht in de manier waarop het systeem criteria hanteert. Wetenschappers werken aan methodes om dit te verbeteren, onder de vlag *explainable AI*.

2.5 AI-verordening

Tot slot merken we op dat er in Europa gewerkt wordt aan specifieke wetgeving over algoritmische systemen, waaronder profilerende systemen. De Europese Commissie presenteerde in maart 2021 het voorstel voor de *Wet op de Artificiële Intelligentie*. Hierin staan voor ontwikkelaars en gebruikers van systemen nadere plichten omschreven, omdat algoritmische systemen mensenrechten kunnen schaden. Voor zogenoemde hoog-risico-systemen bevat het wetsvoorstel onder andere verplichtingen om risico's van algoritmische systemen te identificeren, te beoordelen en maatregelen te nemen om ze zoveel mogelijk uit te sluiten of te beperken. Daarbij gaat het om de hele levenscyclus van een algoritmisch systeem. Het wetsvoorstel stelt eisen aan data en databaseer, onder meer ten aanzien van representativiteit van de dataset en de beoordeling van mogelijke vertekeningen (bias). Ook staan er in het wetsvoorstel verplichtingen ten aanzien van transparantie en informatieverstrekking, zowel voor gebruikers als voor toezichthouders. De voorgestelde AI-verordening speelt zo in op verschillende van de hiervoor genoemde mensenrechten die relevant zijn bij het gebruik van algoritmische systemen.

Tegelijkertijd zijn er nog veel vragen: wat houdt de eis dat de werking voldoende transparant moet zijn precies in? Hoe moet zo'n algoritmisch systeem voor risicobeoordeling eruitzien? Hoe spoor je vertekeningen op de juiste wijze op en welke verantwoording leg je daarover af? Het Europese wetgevingsproces zal uitwijzen in hoeverre het wetsvoorstel op deze punten wordt geconcretiseerd.

3 Maatregelen bij UWV, SVB, CJIB

Om te inventariseren wat voor soort maatregelen uitvoeringsorganisaties nemen om mensenrechten te beschermen bij de inzet van profilerende algoritmes, hebben we dertien gesprekken gevoerd met medewerkers van het Uitvoeringsinstituut Werknemersverzekeringen (UWV), de Sociale Verzekeringsbank (SVB) en het Centraal Justitieel Incassobureau (CJIB). We spraken met medewerkers van verschillende afdelingen, zoals het ontwikkelteam, *privacy officers*, de functionaris gegevensbescherming en leden van de directie of het bestuur. Zie bijlage 1 voor een overzicht van de gesproken medewerkers en bijlage 2 voor het interviewprotocol.

In dit hoofdstuk beschrijven we eerst waarvoor de organisaties profilerende algoritmes momenteel inzetten, welke mensenrechtelijke risico's ze signaleren en wat ze doen om deze risico's te mitigeren, in termen van werkprocessen, functies, expertise en hulpmiddelen. We sluiten af met een overzicht van de bevindingen.

3.1 De inzet van profilerende algoritmes

De onderzochte uitvoeringsorganisaties willen gebruik gaan maken van profilerende algoritmes. De doelen daarvoor lopen uiteen. De profilerende algoritmes van het UWV en de SVB worden ingezet om fraude en foutief gebruik van regelingen op te sporen. Het profilerende algoritme van het CJIB is gericht op het innen van openstaande boetes en het voorkomen van schuldenoploop. In de kaders hieronder geven we weer hoe de organisaties de inzet van profilerende algoritmes beschrijven. In het volgende hoofdstuk zullen we hierop reflecteren. Samengenomen geven de voorbeelden een beeld van organisaties die bijna klaar zijn met het ontwikkelen van profilerende systemen. Alle voorgestelde profilerende algoritmische systemen zouden in eind 2021/begin 2022 in productie moeten gaan.

Kader 1 **Profilerende algoritmische systemen bij het UWV**

Het UWV maakt binnen verschillende afdelingen gebruik van verschillende typen algoritmes – niet alleen in de toetsings- en handhavingspraktijk, maar ook als dienstverlening bij het matchen van vacatures, cv-bestanden en werkgeversvragen. Binnen de afdeling Handhaving worden verschillende risicoscans ontwikkeld die specifieke regelovertredingen helpen te detecteren. De risicoscans worden gebruikt om inzicht te krijgen in het gedrag dat burgers laten zien, zoals het online invullen van een formulier of aanvraag, of de telefoongesprekken met het callcenter van het UWV. Het UWV kijkt naar gedrag dat direct samenhangt met de kans op die regelovertredingen. Het UWV geeft aan niet naar kenmerken als leeftijd of gender te kijken en geen data van sociale media of openbare websites te halen.

Er zijn drie verschillende risicoscans in ontwikkeling die naar verwachting in 2022 live zullen gaan: *verwijtbare werkloosheid*, *onrechtmatig verblijf buiten Nederland* en *sollicitatiescan*. De eerste twee risicoscans vormen een aanvulling op andere manieren van detectie, zoals interne en externe meldingen van regelovertredingen en steekproeven. De sollicitatiescan wordt ingezet om te ontdekken welke burgers aanvullende dienstverlening nodig hebben bij het solliciteren.

Afgezet tegen een willekeurige steekproef blijkt volgens het UWV dat de risicoscans 2,5 tot 8 keer effectiever zijn. Het gebruik van risicoscans is bovendien volgens het UWV voordelig voor de burger. Burgers die zich aan regelingen houden, worden met behulp van de risicoscans minder aan onderzoek van het UWV onderworpen. Ook dient het de tijd die ligt tussen de WW-aanvraag en detectie van een regelovertreding te verkorten, zodat het aantal terugvorderingen dat in een laat stadium plaatsvindt, wordt verminderd.

Kader 2 Profilerende algoritmische systemen bij de SVB

De SVB zet algoritmes onder andere in voor fraude-detectie, het vaststellen waar een burger recht op heeft en ook het vaststellen van niet-gebruik. De algoritmes helpen om te bepalen of mensen krijgen waar ze recht op hebben, zoals de kinderbijslag of het persoonsgebonden budget. Het gaat veelal om *rule based* algoritmes die een één-op-één vertaling zijn van de wet.

De SVB ontwikkelt ook profilerende algoritmes. Het gaat dan om zogenaamde *case based* algoritmes. Een daarvan gaat in 2022 live en heeft de naam SWAN (samenwonen alleenstaand). Burgers die een AOW-uitkering ontvangen, kunnen bewust of onbewust misbruik maken van de regeling door samen te wonen en allebei een AOW-uitkering voor alleenstaanden te ontvangen. Zo ontvangen ze meer dan waar zij recht op hebben. Door middel van het SWAN-algoritme kan er met behulp van een risicoscore een kans berekend worden dat iemand onrechtmatig een uitkering ontvangt. Het algoritme kijkt vooral naar het moment van samenwonen.

Op basis van deze sortering zouden fraude en foutief gebruik sneller opgespoord moeten worden. Dat is nuttig, geeft de SVB aan, omdat de handhavingscapaciteit beperkt is. Het algoritme heeft een signalerende functie en zal een aanvulling vormen op een *rule based* systeem dat al bestaat. De werkelijke vaststelling van fraude of niet-gebruik blijft volgens de SVB altijd mensenwerk.

Het SWAN-algoritme is volgens de SVB niet alleen belangrijk om fraude tegen te gaan, maar ook om ervoor te zorgen dat burgers niet te hoge bedragen aan terugvordering hoeven te betalen en vast komen te zitten in schulden. Het beoogde effect is dus tweeledig: het aanpakken van fraude en daarmee borgen van draagvlak voor de verzorgingsstaat, en het beschermen van onschuldige mensen.

Kader 3 Profilerende algoritmische systemen bij het CJIB

Het CJIB ontwikkelt in het Innovatielab⁰⁵⁸ profilerende algoritmes onder meer om schuldenoploop bij burgers te voorkomen. Deze oploop kan ontstaan door de stapeling van openstaande verkeersboetes. Een instrument dat deze schuldenoploop moet voorkomen, is het algoritmische systeem *Telefonisch Innen* (werktitel) dat medio 2022 waarschijnlijk live zal gaan. Dit betekent dat de pilotfase dan is afgerond en het systeem in de reguliere werkwijze wordt opgenomen. Telefonisch Innen zal worden ingezet voor debiteuren die:

1. een boete hebben gekregen voor een overtreding van de Wet administratieve handhaving verkeersvoorschriften (Wahv);
2. de tweede aanmaning hiervoor hebben ontvangen, maar niet hebben betaald; en
3. waarschijnlijk zonder tussenkomst van een deurwaarder de boete alsnog kunnen en willen betalen.

Het profilerende algoritme is bedoeld om de meest kansrijke zaken voor een telefonische interventie te selecteren. Met andere woorden: welke mensen kan het CJIB het beste bellen om een succesvolle betaalafpraak mee te maken?

Het type algoritme dat is ontwikkeld voor *Telefonisch Innen* is een beslisboom, een lerend algoritme dat op basis van data een voorspelling doet over de kans op betaling. Het gaat om zaakgegevens van de betrokkene. De personen die het algoritme selecteert, worden gebeld door medewerkers van het CJIB, waarbij de medewerker beslist of er gebeld wordt. Het doel is drieledig: te komen tot betaalafspraken, daarmee schuldenoploop tegengaan en de inzet van deurwaarders voorkomen. Met dit algoritme is er volgens het CJIB sprake van 30% meer effectiviteit ten opzichte van de situatie waarin het instrument niet wordt ingezet.

3.2 Gesignaleerde risico's

De onderzochte uitvoeringsorganisaties zijn zich bij de ontwikkeling van hun profilerende algoritmes bewust van mogelijke risico's. Op de vraag welke mensenrechtelijke risico's bij de organisaties op de radar staan, noemden de geïnterviewde medewerkers zaken als privacy, bias en transparantie, uitlegbaarheid en menselijke tussenkomst. Het viel echter op dat de medewerkers de risico's vaak niet benoemden als *mensenrechtelijke* risico's: men antwoordde in termen van specifieke wettelijke vereisten of ethische principes, en benoemde vaak al direct de genomen maatregelen. De medewerkers gaven het belang aan van investeringen in het ethische karakter van de organisatie. De maatregelen die de onderzochte organisaties nemen om risico's te mitigeren komen grofweg overeen met de indeling die we in hoofdstuk twee beschreven: privacyrechten, gelijkheidsrechten, vrijheidsrechten en procedurele rechten. In de gesprekken lag de nadruk op privacyrechten en het recht op non-discriminatie.

3.2.1 Een inventarisatie van de belangrijkste maatregelen

De volgende paragrafen vatten de belangrijkste maatregelen samen die de onderzochte uitvoeringsorganisaties hebben genoemd om de gesignaleerde risico's te mitigeren. We zullen de organisaties in samenhang bespreken, en de maatregelen indelen in drie categorieën: functies, expertise en hulpmiddelen, en werkprocessen. Tussen deze categorieën zit enige overlap. Het gaat erom dat ze samen een rijk beeld schetsen van de bescherming van mensenrechten bij algoritmische profilering. Dit hoofdstuk is een weergave van de gesprekken die wij hebben gevoerd met medewerkers van de onderzochte uitvoeringsorganisaties. In hoofdstuk vier geven we onze analyse van deze gesprekken.

3.2.2 Functies

Onder functies beschrijven we de belangrijkste posities en bijbehorende verantwoordelijkheden die binnen de onderzochte uitvoeringsorganisatie gecreëerd zijn of doorontwikkeld worden om mensenrechten te beschermen. We bespreken de functies en besteden daarbij aandacht aan de manier waarop ze zijn ingericht.

Functionaris gegevensbescherming

De onderzochte organisaties kennen een functionaris gegevensbescherming (FG), die de interne en onafhankelijke toezichthouder is op het gebied van alle

gegevensverwerking die bij een organisatie plaatsvindt. Deze functie vloeit voort uit de AVG. De FG houdt intern toezicht op de toepassing en naleving van deze wet. De FG toetst in dit kader de uitvoering van de data protection impact assessment (DPIA), ook wel een gegevensbeschermingseffectbeoordeling (GEB) genoemd.

Bij een DPIA of GEB wordt bijvoorbeeld vastgelegd met welke gegevens risicoprofielen van burgers tot stand komen. Als de FG constateert dat een voorgesteld systeem niet voldoet aan de AVG, kan hij of zij om aanpassingen vragen of geen goedkeuring verlenen. Beide zaken komen voor, geven de FG's van het UWV en de SVB aan, ook bij de toepassing van lerende algoritmes. Er kan spanning bestaan tussen wat ontwikkelteams willen en wat wettelijk mogelijk is. Als er een hoog risico wordt gedefinieerd door de FG moeten er aanvullende maatregelen komen.

Functionaris gegevensbescherming Gerjan van Schaik van het UWV over zijn rol:

'In het verleden probeerden we het helemaal sluitend te krijgen. Dat lukt niet. Dan kom je nooit tot een eindpunt. Dus we proberen na twee, drie keer een definitief advies uit te brengen. Het is voor ons belangrijk dat de basis op orde is, dus geen onrechtmatigheid of hoog restrisico, voor het resterende risico definieer je maatregelen. Als het restrisico voor de directie acceptabel is, zetten we er gewoon een streep onder en wordt het gefiled.'

Volgens de geïnterviewde FG's is de AVG geen eenvoudige wet, waarbij aandacht voor de context belangrijk is. Verschillende afdelingen zijn betrokken bij discussies over hoe de wet geïnterpreteerd moet worden, bijvoorbeeld over zaken als grondslagen, doelbindingen en proportionaliteit. Uit de gesprekken met de FG's van het UWV en de SVB bleek dat de FG's ook kijken naar mogelijke discriminatoire effecten van voorgestelde systemen.

Functionaris gegevensbescherming Hatice Dogan van de SVB over de complexiteit van de AVG:

'De AVG is eigenlijk de meest complexe wet die je kan bedenken. Ik hoor mensen weleens zeggen dat ze de AVG helemaal begrijpen. Maar ik werk hier al jaren en snap hem nog steeds niet helemaal – je moet telkens goed kijken naar de context en een interpretatie geven van open normen.'

De FG van SVB geeft aan dat de expertise bij de ontwikkelteams nog niet dusdanig is, dat ze zich kan terugtrekken en slechts steekproefsgewijs DPIA's kan toetsen. Bij het CJIB speelt de FG een andere rol dan bij de SVB en het UWV, omdat de FG meer op afstand staat. De FG van het CJIB is ondergebracht bij het ministerie van Justitie en Veiligheid en ziet ook toe op andere organisaties die onder het ministerie vallen. De FG's van het UWV en de SVB zien alleen toe op hun eigen organisatie. Bij het CJIB speelt de privacy officer daarom een grote rol bij het toetsen van de DPIA's – deze functie bespreken we hieronder.

Privacy officer

De onderzochte organisaties kiezen naast de FG doorgaans ook voor het instellen van een *privacy officer*. In tegenstelling tot de FG is het instellen van een *privacy officer* niet verplicht. Een *privacy officer* is bedoeld om vooraf advies te geven en is vaak actiever betrokken bij de ontwikkeling van nieuwe systemen. Bij de onderzochte organisaties zijn zij het eerste aanspreekpunt bij vragen over gegevensverwerking en bij het uitvoeren van een DPIA. Zij kunnen daarbij ondersteund worden door de juridische afdeling. Als er bij het CJIB bijvoorbeeld data ter beschikking van de ontwikkelteams worden gesteld, gaat het verzoek hiervoor eerst langs de *privacy officer*. Het CJIB kent wat data betreft twee hoofdlijnen: de data die onder de AVG vallen, namelijk de data vanuit de Wet administratiefrechtelijke handhaving verkeersvoorschriften (Wahv), en de data die bestaan uit de tenuitvoerleggingsgegevens waarbij de Wet justitiële en strafvorderlijke gegevens (Wjsg) van toepassing is. Deze kaders sluiten elkaar expliciet uit en mogen dus niet gecombineerd worden.

Privacy officer Anja Tol van het CJIB is nauw betrokken bij het werkproces van het ontwikkelteam:

'Het Innovatielab komt vaak bij ons op de lijn: we willen deze data gebruiken, mag dit? Dan geven wij aan welke data wel en niet gebruikt mogen worden. (...) Alles wat zij (het Innovatielab) doen, wordt wekelijks met ons besproken. Vooraf denken we mee en daarnaast werken we vanuit privacy by design, zodat we vanaf de beginfase rekening kunnen houden met privacy.'

Adviseur Recht en Ethiek, kwartiermaker data-autoriteit CIO

De onderzochte organisaties ontwikkelen ook functies die specifiek gericht zijn op ethiek. Zo kent de SVB al een aantal jaar een adviseur Recht en Ethiek, en is er een kwartiermaker data-autoriteit CIO aangesteld die zich onder andere bezighoudt met het opstellen van het normenkader dat bij de SVB gebruikt gaat worden – we zullen in paragraaf 3.3.3 het normenkader nader toelichten. De kwartiermaker data-autoriteit werkt ten behoeve van dit normenkader – waarin

ethiek een rode draad vormt – nauw samen met de afdeling van de adviseur Recht en Ethiek. De FG van de SVB werkt ook samen met de adviseur Recht en Ethiek. Het doel van dit schakelen met een ethicus is om de interpretatie van normen (van onder andere de AVG) goed door te spreken en er doordachter invulling aan te geven.

Ethische commissie

Het CJIB is betrokken bij een werkgroep ethiek binnen het ministerie van Justitie en Veiligheid, waar sprake is van het opzetten van een externe klankbordgroep die ethiek advies geeft. Het UWV is voornemens om begin 2022 een Commissie Data Ethiek op te richten die volgens medewerkers van het UWV toe kan zien op de naleving van het Kompas Data Ethiek en het voeren van het juiste gesprek hierover. Hoewel de commissie nog in ontwikkeling is, wordt het idee geopperd om de commissie te laten kijken naar het concretiseren van normen. Ook zal bepaald worden hoe het kompas precies samenhangt met de al bestaande en wettelijk verplichte gegevensbeschermingseffectbeoordeling. Het voorlopige idee is dat de commissie eens in de zes weken bij elkaar komt. Tijdens een leerjaar zullen er ervaringen worden opgedaan over hoe de commissie optimaal kan bijdragen aan het vermogen van het UWV om verantwoord met data om te gaan.

Projectmanager betrokken bij de Commissie Data Ethiek Eveline Jacobson van het UWV over de rol van de commissie bij de interpretatie van de AVG:

‘De sleutel lijkt in de open normen van de AVG te liggen, bijvoorbeeld ten aanzien van transparantie, behoorlijkheid en menselijke tussenkomst. Die normen moet het UWV voor zichzelf invullen en intern voortdurend bediscussiëren, daar kan de Commissie Data Ethiek een bijdrage aanleveren, aan de hand van het Kompas Data Ethiek. Feitelijk loopt het UWV dus vooruit op de rechtsontwikkeling, aangezien rechters deze normen ook steeds verder preciseren. Dat is ook niet gek, aangezien de AVG nog jong is.’

Er zijn in september 2021 drie manieren geïdentificeerd waarop de Commissie Data Ethiek van het UWV zou kunnen werken:

1. op bepaalde momenten in een project kan het verplicht worden om een oordeel van de commissie te vragen;
2. een projectteam kan vrijwillig om een oordeel vragen; en
3. de commissie kan zich op eigen initiatief ergens over uitspreken, bijvoorbeeld naar aanleiding van ontwikkelingen in de media of de wetenschap.

Volgens een medewerker van het UWV is besloten dat de commissie advies geeft over specifieke bestaande en te ontwikkelen datatoepassingen. Daarnaast kan de commissie op verzoek over ethische dilemma's adviseren, bijvoorbeeld naar aanleiding van de actualiteit. Ook kan de commissie ongevraagd advies geven, bijvoorbeeld naar aanleiding van maatschappelijke ontwikkelingen.

Net als bij het UWV is de geplande ethische commissie bij het CJIB geen vervanging van de bestaande privacy-structuren waar de FG en de *privacy officer* onderdeel van zijn, maar een toevoeging aan wat er nu is. Het CJIB is aangehaakt bij de werkgroep Ethiek van het ministerie van Justitie en Veiligheid.

Projectmanager betrokken bij de Commissie Data Ethiek Eveline Jacobson van het UWV over de toegevoegde waarde van de commissie:

'Zolang er geen jurisprudentie of andere gevoelens in de maatschappij zijn over een bepaald ethisch aspect, zou je als organisatie er toch iets van moeten vinden op het moment dat je een algoritme ontwikkelt.'

Uit gesprekken met de uitvoeringsorganisaties blijkt dat men vindt dat ethiek niet alleen bij bepaalde functies belegd moet zijn, maar een zaak is voor alle medewerkers. Dit brengt ons bij de volgende paragraaf, over expertise.

3.2.3 Expertise

Onder expertise verzamelen we de initiatieven en maatregelen die bij medewerkers de kennisbasis en vaardigheden ten aanzien van de bescherming van mensenrechten versterken. Weet men genoeg over biasdetectie of *fairness*, en hoe kunnen medewerkers hun kennis effectief in de praktijk brengen?

Het gesprek aangaan over ethiek, bias en *fairness*

Uit de gesprekken bleek dat de medewerkers van de onderzochte organisaties zich bewust zijn van de risico's van vooringenomenheid (bias) met betrekking tot profilerende algoritmische systemen. Niet alleen tijdens de ontwikkeling van de systemen bestaat er een risico dat er bias in de systemen sluipt, maar ook tijdens de implementatie van de systemen en de uitvoering van de daarop volgende beslissingen kan bias van medewerkers een ongewenste rol spelen.

Bij de SVB werd gedurende de pilotfase van het SWAN-algoritme duidelijk dat er vooral mannen werden geselecteerd. Bij een discussie over de mogelijke problematiek hiervan werd duidelijk dat de medewerkers van de afdeling Handhaving vertrouwen op het eigen oordeelsvermogen en werkervaring, en niet

altijd rekening leken te houden met onbewuste bias en blikvernaauwing. Mede op aangeven van de adviseur Recht en Ethiek, wordt binnen de SVB nu nagedacht over de manier waarop een gesprek hierover gestart kan worden, zonder dat medewerkers zich aangevallen voelen. Een eerste bewustwordingssessie heeft inmiddels plaatsgevonden.

Naast de sessies over vooringenomenheid zijn er bij de ontwikkeling van het SWAN-algoritme ethische sessies georganiseerd waarbij de *ethical matrix* wordt gebruikt. Deze methode is onder andere ontwikkeld door Marlies van Steenberg en collega's aan de Hogeschool Utrecht. De beleidsadviseur Recht en Ethiek van de SVB, Alex Corra, heeft in samenwerking met Marlies van Steenberg deze matrix omgezet in een specifiek instrument (workshopformat) voor de SVB. Tijdens zo'n sessie worden aan de hand van dit instrument verschillende vragen beantwoord, zoals: wie zijn de stakeholders, wat zijn hun belangen en hoe worden die geraakt, tussen welke waarden ontstaat er spanning en wat zijn de nadelige effecten en hoe kunnen die spanningen of effecten geadresseerd worden? Deze resultaten worden gekoppeld aan ethische stromingen, zoals utilisme, deontologie en deugdenethiek waardoor duidelijk wordt waar spanningen en keuzes liggen.

Senior beleidsadviseur Recht en Ethiek Alex Corra van de SVB over deze aanpak:

'Eigenlijk ga je bij het maken van een ethical matrix eerst op zoek naar de directe en indirecte stakeholders. En wat zijn dan de directe en indirecte belangen die geraakt kunnen worden? En wat zijn de mogelijke adverse effects? Je gaat op zoek naar de bedoelde en onbedoelde impact voor die verschillende groepen. Welke waarden en bijbehorende spanningen zitten daarachter en dan ga je kijken: kunnen we die met bepaalde oplossingen adresseren? Kunnen we daar uiteindelijk requirements omheen formuleren?'

Uit de gesprekken blijkt dat het voor de organisaties belangrijk is om 'het goede te doen' en hierbij zorgvuldige afwegingen te maken. Ook al zijn er bepaalde wettelijke voorschriften, dat wil nog niet zeggen dat hetgeen wat mag, ook maatschappelijk gewenst is.

De geïnterviewde medewerkers zijn zich bewust van de rol die werkcultuur kan spelen bij de totstandkoming van nadelige effecten voor burgers wanneer profilerende algoritmische systemen worden gebruikt. Om situaties als de kinderopvangtoeslagenaffaire te voorkomen, zet men zich in om het ethische bewustzijn bij medewerkers te verhogen en zo de cultuur binnen de organisatie te kunnen veranderen.

Een belangrijk element van deze cultuurverandering is het serieus nemen van ethische bezwaren. Wanneer deze bezwaren geuit worden, moet dat aangemoedigd worden en niet genegeerd, zo stelt een medewerker van het UWV. Wanneer kritiek niet wordt overgenomen, moet uitgelegd worden waarom. Het goed uitleggen waarom je als organisatie iets wel of niet doet, is de sleutel. In alle haast en drukte wordt dit nog wel eens vergeten.

Projectmanager betrokken bij de Commissie Data Ethiek Eveline Jacobson van het UWV over het belang van transparantie en toetsing:

‘Juist in de vaart der volkeren en door de snelheid van het “heilige moeten” zien we uitlegbaarheid en het grondig toetsen van “waarom doen we het ook alweer” over het hoofd.’

Kader 4 De Bedoeling bij de Sociale Verzekeringsbank

Bij de SVB bestaat het programma De Bedoeling, waarin 30 tot 50 medewerkers van verschillende afdelingen, meedenken en meedoen aan projecten die bijdragen aan de bedoeling van de SVB. Hier wordt bijvoorbeeld nagedacht over de ruimte van de ambtenaar en het leveren van maatwerk. Maatwerk blijkt in de praktijk moeilijk, omdat management en bestuur in de organisatie soms toch nog een te sterke nadruk blijven leggen op beheersen en optimaliseren. De vraag ‘waarom doen we dit eigenlijk?’ wordt dan te weinig gesteld. Ook het gebruik van algoritmes ter ondersteuning van maatwerk zijn binnen dit programma besproken.

Sectiemanager Intelligence en Risicogericht onderzoek Jeffrey Apon van de SVB over het zoeken naar houvast bij het ontwikkelen van algoritmische systemen:

‘De vraag is steeds: hoe kan je zorgvuldig genoeg zijn? Dat is steeds wikkelen en wegen, maar ik merk wel dat wij daar heel veel tijd en effort in steken.’

Bij het CJIB staan vergelijkbare inspiratiesessies over ethiek op de planning en voert men momenteel al gesprekken met medewerkers over *fairness*. Om een systeem te kunnen ontwikkelen dat mensen niet onterecht benadeelt, en dus zo

eerlijk mogelijk is, zijn medewerkers van het UWV op zoek gegaan naar wat mensen als eerlijk beschouwen. Uitgangspunt daarbij is dat men bij profilerende algoritmische systemen wel gedragskenmerken mag gebruiken waarop mensen over het algemeen invloed hebben, en geen kenmerken waaraan mensen niks kunnen doen. Als voorbeeld werd genoemd dat je er niks aan kunt doen dat je donker haar hebt, maar wel dat je meerdere keren naar de klantenservice belt.

Een andere manier om de *fairness* van een systeem te beoordelen, vraagt technische expertise. Dat gaat door het vaststellen van het aandeel foutpositieven. Van een foutpositief is bijvoorbeeld sprake wanneer iemand het label krijgt van een hoge kans op fraude, terwijl nader onderzoek uitwijst dat er geen fraude is. Als voor een bepaald kenmerk het aandeel foutpositieven in verschillende groepen gelijk is, dan is de wetenschappelijke hypothese dat er voor dat kenmerk geen bias zit in het systeem. Over de definitie van bias en de juiste technische biasdetectie is in de wetenschappelijke literatuur nog geen consensus bereikt. De geïnterviewde medewerkers zien veel verschillende toetsen, maar geen duidelijke standaard. Het UWV heeft een interne werkgroep die de wetenschappelijke literatuur over bias in de gaten houdt.

Opleidingen en certificering

De onderzochte uitvoeringsorganisaties zijn actief bezig met het ondersteunen van het ethisch bewustzijn binnen de organisatie. Het streven is dat ethiek een onderdeel wordt van het vakmanschap van de medewerkers. Het vormt dan een aanvulling op de bestaande wettelijke vereisten zoals die in de AVG. Zoals eerder genoemd, zullen een ethische commissie of ethisch adviseur niet de enige actoren zijn die moeten toezien op ethiek. Ook leidinggevenden, data-analisten en andere medewerkers dienen te werken met ethische principes in het achterhoofd.

Privacy officer Anja Tol van het CJIB over het karakter van de organisatie:

'De organisatie neemt de bescherming van persoonsgegevens heel serieus. Mensen in de hele organisatie zijn zich bewust van de risico's die komen kijken bij het werken met data. Het Innovatielab heeft het zelf vaak goed in de gaten, weet zelf wat wel en niet kan.'

Op dit moment werkt het UWV aan een ethische opleiding en voorlichting van medewerkers. Ook het CJIB zegt aandacht te besteden aan de (ethische) opleiding van medewerkers. In 2019 was er vanuit het ontwikkelteam een vraag naar een bepaalde ethiek-certificering.

Lead datascientist Innovatielab058 Joke de Boer van het CJIB over bewustwording binnen de organisatie:

‘Sinds wij zijn gestart met een ethiekcertificering merk je echt dat er zoveel awareness is bij de data scientists, dat vind ik een hele positieve ontwikkeling.’

Kennis verbreden en verdiepen door externe samenwerkingen

Naast interne sessies om de expertise van medewerkers te ondersteunen, zoeken de onderzochte organisaties contact met externen (ministeries, academici, expertisegroepen) om kennis te delen en inzichten op te doen.

Het CJIB neemt deel aan periodieke kennissessies, onder andere over bias en *fairness*, met het ministerie van Justitie en Veiligheid en alle andere organisaties die onder het ministerie vallen. Ook het JenV Datalab is hier een onderdeel van.

Het CJIB neemt deel aan de *Data&AI Board* bij JenV en deelt kennis binnen de *Manifest Groep* waar bestuurders van uitvoeringsorganisaties elkaar treffen.

De SVB onderhoudt nauw contact met externe partijen, zoals de Volksbank, die ook betrokken is bij het opzetten van een algoritme-expertisegroep, met onder andere de Erasmus Universiteit, Achmea en KPN. De SVB gaat hier in de toekomst misschien ook aan deelnemen. Daarnaast heeft de SVB een *Data Stakeholder Board* die getrokken wordt door de kwartiermaker *data-autoriteit CIO-office*.

Lead datascientist Innovatielab058 Joke de Boer van het CJIB ziet een behoefte bij uitvoeringsorganisaties aan kennisuitwisseling en samenwerking:

‘Wij hadden behoefte aan gesprekken met andere organisaties. Daarom hebben we kennissessies georganiseerd. Zowel ten aanzien van het onderwerp ‘bias’ als ‘fairness’ is de behoefte bij organisaties aanwezig om hierover met elkaar in gesprek te gaan en kennis uit te wisselen.’

Functionaris gegevensbescherming Hatice Dogan van de SVB onderschrijft het belang van samenwerkingen tussen verschillende functies binnen de organisatie:

‘Samenwerken met ethici is belangrijker dan het ontwikkelen van instrumenten als de ethische matrix, die dan via intranet worden verspreid. Dat is toch vrij theoretisch. Het werkt het beste als je bij vraagtekens in concrete situaties met mensen met expertise en verschillende perspectieven het gesprek aan kan gaan. Dan ontdekken mensen welke grenzen in acht moeten worden genomen.’

3.2.4 Hulpmiddelen

Onder hulpmiddelen verzamelen we de richtlijnen, checklists en assessments die uitvoeringsorganisaties gebruiken om mensenrechten te beschermen bij de ontwikkeling en implementatie van profilerende algoritmische systemen. Wat maakt een richtlijn effectief voor de uitvoeringsorganisatie? Hoe werkbaar zijn de wettelijke kaders?

Impact assessments

Door het grote aanbod van richtlijnen en codes die (mede door de Rijksoverheid) worden gepubliceerd, is het voor de geïnterviewde medewerkers soms lastig om hierin een weg te vinden. We zien dan ook dat organisaties vanuit een behoefte aan duidelijkheid en ‘concrete checklists’, naast de bestaande en wettelijke verplichte DPIA, hun eigen assessments ontwikkelen.

Een *data protection impact assessment* (DPIA), ook wel gegevensbeschermingseffectbeoordeling (GEB) genoemd, is een verplicht onderdeel van de Algemene verordening gegevensbescherming (AVG). Bij dit assessment wordt bekeken welke impact gegevensverwerking heeft op de persoon over wie de data gaan. Welke persoonsgegevens worden gebruikt? Hoe lang worden ze bewaard? Worden de gegevens geanonimiseerd? Doordat er verschillende afdelingen bij betrokken zijn, kan het uitvoeren van een DPIA wel twee tot drie maanden duren. Vanwege de twee datastromen waarmee het CJIB te maken heeft (AVG en Wjsg), heeft het CJIB een specifieke DPIA voor de eigen organisatie ontwikkeld, gebaseerd op die van de Rijksoverheid. Een DPIA moet zijn goedgekeurd door de functionaris gegevensbescherming.

Volgens de functionaris gegevensbescherming van de SVB zou het bij de verwerking van risicoprofielen in de handhaving of dienstverlening kunnen helpen als er bepaalde ethische vragen aan de DPIA worden toegevoegd, op basis van bijvoorbeeld de *Impact Assessment Mensenrechten en Algoritmes (IAMA)*. De IAMA kan dan gebruikt worden om systemen te classificeren op hoog en laag risico, waarna men besluit of er met een zwaar normenkader gewerkt gaat worden.

Het CJIB heeft de CJIB-AI impact assessment ontwikkeld. In dit assessment zijn elementen meegenomen uit de IAMA, het normenkader van de Auditdienst Rijk (ADR) en de richtlijnen van het ministerie van Justitie en Veiligheid. De *data scientists* hebben behoefte aan concrete checklists en concrete richtlijnen die goed toepasbaar zijn.

Lead data scientist Innovatielab058 Joke de Boer van het CJIB zegt daar het volgende over:

‘Door de ontwikkelingen en uitdagingen die toepassing van AI met zich meebrengt, zijn in relatief korte tijd veel kaders, richtlijnen en handreikingen verschenen. Dit leidde tot een onoverzichtelijke situatie en lastige toepasbaarheid in het dagelijks werk. Daarom hebben wij de verschillende kaders en richtlijnen zoveel mogelijk vertaald in een eigen werkwijze.’

Voor het hanteren van de juiste juridische normen, heeft het CJIB aan Pels Rijcken gevraagd een juridische handreiking te schrijven die toegespitst is op de context van het CJIB. De handreiking bevat een checklist waarin staat aan welke juridische vereisten het Innovatielab058 moet voldoen en welke stappen de medewerkers moeten zetten in de diverse fasen van ontwikkeling en toepassing van de algoritmes. Deze handreiking is gedeeld met het departement en andere uitvoeringsorganisaties.

Een andere vorm van *impact assessment* zijn de burgerpanels van de SVB. Deze worden ingezet bij de beoordeling van nieuwe systemen. Dit gebeurt zodra de organisatie impact van het systeem op de burger verwacht. De panels bestaan uit 200 tot 250 mensen. Daarnaast heeft de SVB een Usability Lab om te demonstreren wat het betekent voor de burger als de organisatie dingen op een bepaalde manier doet.

Ethische kaders

De open normen van de AVG ten aanzien van bijvoorbeeld transparantie, behoorlijkheid en menselijke tussenkomst, worden bij de onderzochte organisaties intern bediscussieerd. Het UWV heeft een Kompas Data Ethiek gemaakt waarmee deze normen verder gepreciseerd kunnen worden. Het kompas is opgesteld om het ethisch bewustzijn van medewerkers te ontwikkelen en om vragen uit de samenleving voor te zijn. De Commissie Data Ethiek zal toezien op de invulling van dit kompas en op de samenhang ervan met de al bestaande en wettelijk verplichte GEB.

Volgens een medewerker van het UWV zou de IAMA een bruikbare bron zijn om naast hun eigen ethische kompas te hanteren. Wel is het volgens de organisatie nog nodig om duidelijkheid te scheppen over de precieze toepassing van dit instrument.

Voorzitter raad van bestuur Maarten Camps van het UWV over de inzet van het Kompas Data Ethiek:

‘Het helpt met onze verantwoording naar buiten. Want we kunnen ook aan de samenleving laten zien van: kijk eens, dit zijn jullie data, die hebben wij, zo gaan wij ermee om, op een manier die ook ten voordele is van jullie, om onze dienstverlening te verbeteren.’

Normenkader

De SVB is bezig met het ontwikkelen van een eigen normenkader, een vertaling van wettelijke vereisten en ethische principes naar concrete vragen. Ook legt het vast wie in de organisatie verantwoordelijk is voor het beantwoorden van die vragen. Een medewerker van de SVB geeft aan dat dit op nationaal niveau nog ontbreekt en dat de organisatie niet wil wachten op kaders die gemaakt worden vanuit de Rijksoverheid of de Algemene Rekenkamer. Het normenkader kent de werkvorm van een audit of assessment. Daarin wordt gecontroleerd of concrete beheersmaatregelen zijn doorgevoerd door de verantwoordelijke partijen, zoals een *data scientist*, businessanalist, data-architect of manager. Zo moet een *data scientist* bijvoorbeeld de vraag beantwoorden of: ‘het algoritme creatieproces een structurele aanpak volgt, inclusief het gebruik van representatieve trainingdata, testdata en productiedata’. Ook moet de manager aangeven of: ‘de organisatie begrijpt dat algoritmes indicaties geven, geen zekerheid’.

De gesproken organisaties geven aan dat zij het niet altijd nodig vinden om een zware toetsing te doen, maar dat een bepaalde mate van toetsing altijd nodig zal zijn.

Sectiemanager Intelligence en Risicogericht onderzoek Jeffrey Apon van de SVB zegt hierover:

‘Met de huidige ontwikkelingen ten aanzien van het gebruik van algoritmes zijn we steeds meer gaan kijken naar hoe zorgvuldig wij bezig zijn. Wat we constateren, is dat ongeacht de complexiteit van een algoritme je altijd bepaalde aspecten van zo’n normenkader er tegenaan moet houden.’

3.2.5 Werkprocessen

Onder werkprocessen beschrijven we de belangrijkste stappen die plaats moeten vinden tijdens het ontwikkelen van een algoritmisch systeem om in lijn met mensenrechten te werken.

Werken aan de hand van *privacy by design*

De drie onderzochte instanties geven aan hun werkprocessen zo te hebben ingericht dat privacyaspecten zo vroeg mogelijk in het ontwerp van nieuwe systemen worden meegenomen. Het CJIB vat dit onder de noemer *privacy by design* – het nadenken over privacy vanaf het begin van het ontwerpen van een dienst of product. De werkwijze vloeit mede voort uit de wettelijke eisen die de AVG stelt en omvat zowel technische als organisatorische maatregelen. Zo beoordelen ontwikkelteams in een vroeg stadium aan welke privacy-eisen het systeem moet voldoen en geeft een *privacy officer* (zie functies) in een vroeg stadium van het ontwikkelproces advies, bijvoorbeeld over de data die gebruikt mogen worden. In de organisaties vindt na de eerste pilot-ontwikkeling een uitgebreide risicobeoordeling plaats, via een DPIA of GEB. De uitkomst van de DPIA wordt vervolgens getoetst door de functionaris gegevensbescherming.

Biastoetsing

De gesproken organisaties geven aan dat biastoetsing technisch complex is en dat nog niet uitgekristalliseerd is wat de beste manier is om deze toetsing uit te voeren. Elke organisatie ontwikkelt nu zelf een proces en methode. Ook organiseert iedere organisatie zelf kennisuitwisseling met wetenschappers om te leren van de nieuwste technieken en deze toe te kunnen passen in de eigen organisatie. Het signaleren (en voorkomen) van bias begint bij interne toetsing door het ontwikkelteam. Hiervoor gebruiken de onderzochte organisaties verschillende methoden. Het CJIB gebruikt bijvoorbeeld de methodiek van Tobias Baer (2019).

Twee organisaties laten ook een externe kwaliteitstoets uitvoeren, bijvoorbeeld door accountants- en adviesbureau KPMG of communicatieadviesbureau Blue Field Agency. KPMG toetst zowel de techniek van de modellen als het gehele ontwikkelproces. Hierbij komen elementen aan bod zoals het voorkomen van bias, of de juiste mensen zijn betrokken bij het ontwikkelproces en of de techniek aansluit bij het doel van het model.

De onderzochte organisaties zijn zich ervan bewust dat de inzet van profilerende systemen voortdurende toetsing op bias vereist. Het UWV voert daarom periodieke monitoring uit hoe de risicoscans presteren. Daar bovenop vindt een jaarlijkse controle plaats.

De organisaties proberen niet alleen via technische modellen en periodieke toetsing bias tegen te gaan. Ze kijken ook naar de manier waarop de uitkomsten van de profilering vervolgens door de organisatie worden gebruikt. Het UWV zorgt ervoor dat de medewerker die een burger gaat onderzoeken niet weet of de melding afkomstig is van een risicoscan (of van een interne melding of

steekproef). Zij noemen dit menselijke tussenkomst. De gedachte achter deze maatregel is dat een medewerker zo onbevooroordeeld mogelijk het onderzoek zal uitvoeren. De kans bestaat namelijk dat een medewerker een foutje in een aanvraag anders beoordeelt wanneer hij weet dat het systeem denkt dat er sprake kan zijn van fraude. Dit dient ervoor te zorgen dat de medewerker het algoritmische systeem niet blind gaat vertrouwen en het eigen onderzoek kritisch en zorgvuldig uitvoert.

De SVB hanteert een soortgelijke werkwijze: naast de selectie van het SWAN-algoritme zullen er willekeurige gevallen worden toegevoegd om de juiste werking van het algoritme te kunnen blijven toetsen.

We zien dat de gesproken organisaties maatregelen nemen om vooringenomenheid bij medewerkers te voorkomen. Data-analist Robin Philippo van het UWV zegt daarover:

'We hebben altijd een menselijke tussenkomst. De uitvoering weet niet op welke kenmerken een klant geselecteerd is. Ze weten niet hoe hoog die score was. Die melding ziet er hetzelfde uit als een interne melding of een externe melding van een burger. Het komt op dezelfde manier binnen in ons workflowsysteem.'

Voorzien in het recht op informatie en het procedureel recht

In het werkproces van de onderzochte organisaties is er in verschillende mate aandacht voor informatievoorziening aan de burger en voor het voorzien in een heldere klachtenprocedure. Zo neemt het CJIB bijvoorbeeld de klachtenprocedure opnieuw onder de loep met de ontwikkeling van nieuwe profilerende algoritmische systemen. Bij het UWV is men bezig met het vastleggen van de beslissingen die genomen worden tijdens de ontwikkelfase van nieuwe systemen. Wanneer er dan vragen vanuit de samenleving komen, wil het UWV zo transparant mogelijk kunnen zijn. Projectmanager betrokken bij de Commissie Data Ethiek Eveline Jacobson van het UWV ziet het als volgt:

'Ik ben altijd heel erg van het worstcasescenario. Stel, op een dag komt er iemand die zegt: mogen wij even kijken wat jullie dan gedaan hebben. Nou, dan wil ik wel met droge ogen kunnen laten zien dat we de zaken op orde hebben.'

Governance structuren

De drie organisaties ontwikkelen voor profilerende systemen een overkoepelende *governance structuur*. Ze bedoelen daarmee het geheel van beleid, regels, principes en functies die beschrijven hoe de organisatie de risico's van

profilerende systemen beheerst, en wie binnen de organisatie welke verantwoordelijk draagt. Rondom privacyrechten bestaat er al een duidelijke *governance structuur* die volgt uit de AVG: werken aan de hand van *privacy by design*, met een DPIA, het aanstellen van een functionaris gegevensbescherming en een *privacy officer*.

De structuur rondom de andere mensenrechten is nog in ontwikkeling. Elke organisatie heeft een hulpmiddel ontwikkeld (zoals een Kompas Data Ethiek, ethische matrix of een eigen checklist) dat belangrijke uitgangspunten vaststelt voor hoe de organisatie omgaat met profilerende algoritmes. Ook zijn er nieuwe functies ingesteld die de organisatie vroegtijdig adviseren bij ethische vraagstukken, zoals een adviseur Recht en Ethiek, die een soortgelijke functie heeft in het werkproces als een *privacy officer*. Alle organisaties zijn bezig met het ethisch bewustzijn versterken, en organiseren daarvoor ethische sessies. Verder zoeken ze naar manieren om het toezicht op dit 'ethische proces' te organiseren, zowel intern als extern.

Zo heeft het UWV in een *Model Risk Management-beleid* (het woord model verwijst naar een algoritme) uiteengezet welke functies verantwoordelijk zijn in de verschillende fasen van de ontwikkelingscyclus van een algoritmisch systeem. Een divisiedirecteur initieert bijvoorbeeld de ontwikkeling van een nieuw systeem en is modeleigenaar. De cyclus eindigt bij de goedkeuring van het systeem door diezelfde modeleigenaar. In de tussenliggende fase bouwen data-analisten het systeem en evalueert een externe controleur de kwaliteit ervan en de biastoetsing. Goedkeuring vindt plaats door de functionaris gegevensbescherming (en straks mogelijk ook een Commissie Data Ethiek). Dan wordt het systeem getest en na goedkeuring van de modeleigenaar uiteindelijk in gebruik genomen, en vindt er periodieke monitoring en jaarlijks onderhoud plaats.

Een aantal werkprocessen bij de SVB zijn gericht op het versterken van de ethische reflectie. Senior beleidsadviseur Recht en Ethiek Alex Corra zegt daarover:

'Nu zie je dat met het creëren van die governance-structuur [...] ethische reflectie ook geïnstitutionaliseerd wordt in de organisatie. Want ik wil niet "het" geweten van de organisatie zijn. Ik ben er om de organisatie te adviseren over dilemma's waar ze mee worstelen en af en toe ook ongevraagd eens te roepen van: goh, weet je het zeker, wil je dit wel? Maar uiteindelijk moet aan de ene kant iedereen het geweten van de organisatie zijn, en aan de andere kant moeten er natuurlijk ook problem owners op bestuursniveau zijn, die gewoon die verantwoordelijkheid

nemen. En die moeten kunnen zeggen: ja, ik heb hier mijn handtekening onder gezet, dit wordt uit mijn naam gebruikt, en daar sta ik achter.'

3.2.6 Technische maatregelen

In onze onderzoeksvraag maken we onderscheid tussen functies, expertise, hulpmiddelen en werkprocessen. In de interviews noemden de uitvoeringsorganisaties echter ook technische maatregelen om diverse mensenrechtelijke risico's te beheersen.

Bouwen in eigen beheer

Een technische maatregel die door de onderzochte organisaties wordt toegepast om *privacy by design* te bewerkstelligen, is het bouwen in eigen beheer. De onderzochte organisaties kiezen bij de inzet van profilerende systemen hier ook voor met het oog op controleerbaarheid en afhankelijkheid. Dat geldt zowel voor de drie huidige pilots, als toekomstige projecten waarbij profilering wordt ingezet.

Hoofd Risicomanagement en Intelligence bij de directie Handhaving Henk de Ruiter van het UWV licht de keuze om in eigen beheer lerende systemen te ontwikkelen als volgt toe:

'We staan stil bij het belang van uitlegbaarheid en het kunnen onderhouden van modellen. Vanuit mijn verantwoordelijkheid voor risicoscans en het toezichtdomein kan ik zeggen dat we geen enkele vendor lock-in of afhankelijkheid willen hebben van externen.'

Om de veiligheid en privacy van burgers te waarborgen tijdens de ontwikkeling van systemen werken de *data scientists* van het CJIB in een afgeschermd omgeving waar bijvoorbeeld geen verbinding is met het internet.

Geen *unsupervised learning*

Om transparant te kunnen zijn, is het voor de organisaties van belang dat zij uit kunnen leggen hoe hun systemen werken. Naast het bouwen in eigen beheer kiezen de gesproken organisaties ervoor om geen *unsupervised learning* toe te passen – ook niet als eerste experimenten tonen dat met deze technieken een hogere effectiviteit kan worden behaald. De organisaties willen kunnen controleren wat een algoritme doet, en dit kunnen uitleggen. De organisaties achten *unsupervised learning* hier niet mee verenigbaar.

Lead datascientist Innovatielab058 Joke de Boer van het CJIB zegt hierover:

‘Uitlegbaarheid van AI is erg belangrijk. Ervoor zorgen dat een algoritme geen black box is en geïnterpreteerd kan worden. Ook voor effectieve menselijke tussenkomst is dit cruciaal.’

Algoritmeregister

De SVB is onderdeel van het programma Publieke Controle op Algoritmes dat een algoritmeregister aan het ontwikkelen is. Dit register moet gaan helpen bij het verkrijgen van inzicht op de ontwikkeling van algoritmes bij uitvoeringsorganisaties.

Kwartiermaker data-autoriteit CIO-office Rolf Leijdekker van de SVB ziet de voordelen van een algoritmeregister:

‘Een nationaal algoritmeregister zou niet alleen helpen voor de transparantie naar buiten toe, maar ook om de samenwerking tussen uitvoeringsorganisaties te versterken.’

3.3 Blik op de toekomst

We eindigen de inventarisatie van maatregelen met een blik op de toekomst. Welke stappen willen de gesproken uitvoeringsorganisaties zelf zetten om de bescherming van mensenrechten te versterken bij de ontwikkeling van profilerende algoritmes?

Roep om centrale regie

Medewerkers van de onderzochte organisaties geven aan dat zij door de grote hoeveelheid (door de overheid ontwikkelde) hulpmiddelen en breed interpreteerbare wetten, soms door de bomen het bos niet meer zien.

Zij hebben behoefte aan consistentie en overzicht. De vele richtlijnen worden als nuttig ervaren, maar zorgen ook voor onzekerheid: is wat de organisatie doet goed en is het voldoende? Een medewerker van de SVB geeft aan dat het behulpzaam zou zijn als het ministerie van Binnenlandse Zaken en Koninkrijksrelaties aan zou geven welke richtlijnen de juiste zijn. Een medewerker van het CJIB heeft behoefte aan richtlijnen vanuit Europa.

Privacy officer Anja Tol van het CJIB geeft aan dat haar organisatie in de toekomst gebaat zou zijn bij een ander soort richtlijnen:

‘Het CJIB zou in de toekomst geholpen zijn met centrale richtlijnen vanuit de Europese Commissie of Europese toezichthouders. Er zijn nu namelijk verschillende richtlijnen en er is behoefte aan uniformiteit.’

Een medewerker van het UWV ziet kansen voor verbetering bij de omgang met bewaartermijnen van data. Die bewaartermijnen zijn losjes geregeld in de selectielijst van de organisatie, die door de minister geaccordeerd wordt en de status van wet heeft. Maar die bewaartermijnen zijn soms arbitrair – vaak is het helemaal niet nodig om zo lang gegevens te bewaren.

Verbeteren van passende dienstverlening

We zien dat de onderzochte uitvoeringsorganisaties niet alleen profilerende algoritmische systemen ontwikkelen voor fraudebestrijding, maar deze juist willen inzetten om burgers te helpen. Met het verbeteren van hun dienstverlening willen de organisaties zorgen dat burgers tijdig krijgen waar zij recht op hebben. De organisaties geven aan dat zij in de toekomst profilerende algoritmische systemen willen inzetten om burgers sneller en efficiënter te kunnen helpen. Een medewerker van de SVB geeft aan dat algoritmes gebruikt zouden kunnen worden om te voorspellen waar maatwerk en aandacht nodig is, en waar niet. Zo kan worden voorkomen dat burgers herhaaldelijk blijven bellen naar de organisatie zonder dat zij gelijk goed geholpen worden.

Senior beleidsadviseur Recht en Ethiek Alex Corra van de SVB denkt na over het met AI verbeteren van de dienstverlening:

‘Kijk, in een ideale wereld hadden wij een computersysteem in de organisatie dat kan aangeven: klant x komt nu binnen, wij vermoeden dat dit misschien een maatwerkgeval kan zijn. Dit zijn de tien oplossingen die het relevantst zouden kunnen zijn voor deze situatie. Zo’n soort kennisondersteuningssysteem zou je idealiter willen hebben.’

Een medewerker van het CJIB geeft aan dat men ernaar streeft om zo vroeg mogelijk in het inning- en incassoproces te komen met een passende inzet. Zo wordt ook voorkomen dat er een middel wordt ingezet dat de maatschappij en de burger niets oplevert maar wel veel geld kost.

Plaatsvervangend algemeen directeur en directeur IT en Bedrijfsvoering
Henkjan Derks van het CJIB zegt hierover:

‘In het kader van passende, persoonsgerichte dienstverlening streven we ernaar zo vroeg mogelijk in het inning- en incassoproces te komen met een passende inzet. Als de kans op geslaagde inzet er niet is, of een ander middel is beter, dan wil je dat vooraf weten en niet achteraf tot de conclusie komen dat het de burger en/of maatschappij weinig tot niets heeft opgeleverd.’

Een medewerker van het CJIB geeft aan dat er nog meer aandacht kan worden besteed aan hoe het CJIB burgers kan helpen. De burger zal altijd een uitleg moeten kunnen krijgen over het algoritme en de uitkomst en de mogelijkheid moeten hebben om in bezwaar te gaan, of anderszins de mogelijkheid moeten krijgen om gegevens te corrigeren als er iets gewijzigd is. De procedures moeten beschikbaar en gebruiksvriendelijk zijn.

Als het gaat om gerichte en tijdige dienstverlening zou volgens een medewerker van de SVB een andere manier van denken nuttig zijn. Denk bijvoorbeeld aan een data-ecosysteem, een overzicht waarin duidelijk wordt wanneer uitvoeringsorganisaties gegevens met elkaar mogen delen om de burger te geven waarop hij of zij op dat moment recht heeft. Het kan daarbij helpen om op een hoger en abstracter niveau naar het leven van burgers te kijken. Met dit inzicht kan men vooraf bedenken welke organisaties die burger nodig gaat hebben en kan je de burger tijdig geven waarop hij of zij recht heeft. Zo kan men gemakkelijker kijken naar welke data de uitvoeringsorganisatie nodig heeft om de burger te mogen helpen. Daardoor kunnen datastromen van bovenaf geregeld worden en blijft er meer ruimte over voor de uitvoering. Het ecosysteem kan zorgen voor rust en overzicht doordat de burger niet meerdere malen dezelfde data met verschillende organisaties moet delen, zoals nu wel het geval is.

Sectiemanager Intelligence en Risicogericht onderzoek Jeffrey Apon van de SVB zegt hierover:

‘Als je het hebt over burgerperspectief en transparantie en verklaarbaarheid, dan zit je aan de voorkant, die burger snapt met een ecosysteem veel beter waarom gegevens worden gedeeld met allemaal organisaties. En voor welke toepassingen, dat is dan veel beter uit te leggen.’

Investeren in samenwerking

De onderzochte uitvoeringsorganisaties investeren in het ontwikkelen van expertise bij hun medewerkers. Een manier om ethiek (meer) op de kaart te krijgen als organisatie, is het aangaan van samenwerkingen, zowel binnen de organisatie als daarbuiten. Op extern niveau geven de organisaties aan dat zij meer inzicht willen hebben in de werkwijze van andere uitvoeringsorganisaties, zodat zij complexe situaties sneller en beter kunnen oplossen. Zo pleit een medewerker van de SVB voor meer overleggen tussen uitvoeringsorganisaties en functionarissen gegevensbescherming.

Op intern niveau zien de organisaties dat de communicatie tussen verschillende afdelingen en functies verbeterd kan worden: verschillende perspectieven moeten bijeen komen. Een medewerker van het UWV legt uit dat medewerkers zich soms in andere werelden begeven – denk aan de programmeur en iemand op managementniveau. Zij denken te vaak nog niet in een gemeenschappelijke taal over de risico's die zouden kunnen ontstaan bij de ontwikkeling en implementatie van (profielerende) algoritmische systeem. Een medewerker van de SVB pleit voor een betere multidisciplinaire samenwerking tussen juristen, programmeurs en ethici rondom de AVG, omdat deze wet complex en breed interpreteerbaar is.

3.4 Samenvattend overzicht

Alle onderzochte uitvoeringsorganisaties ontwikkelen momenteel profilerende algoritmes. Het gaat om pilots die vanaf eind 2021 en begin 2022 structureel in gebruik worden genomen. De onderzochte uitvoeringsorganisaties zijn zich bewust van mogelijke risico's, en noemen onder meer zaken als privacy, transparantie, uitlegbaarheid en menselijke tussenkomst. De nadruk lijkt te liggen op privacyrechten en het recht op non-discriminatie. Het valt echter op dat de organisaties de risico's in termen van ethische en specifieke wettelijke eisen bespreken, en niet in termen van mensenrechtelijke verdragen. Investerings in ethiek zien de gesproken organisaties als een belangrijk onderdeel in het ondervangen van risico's.

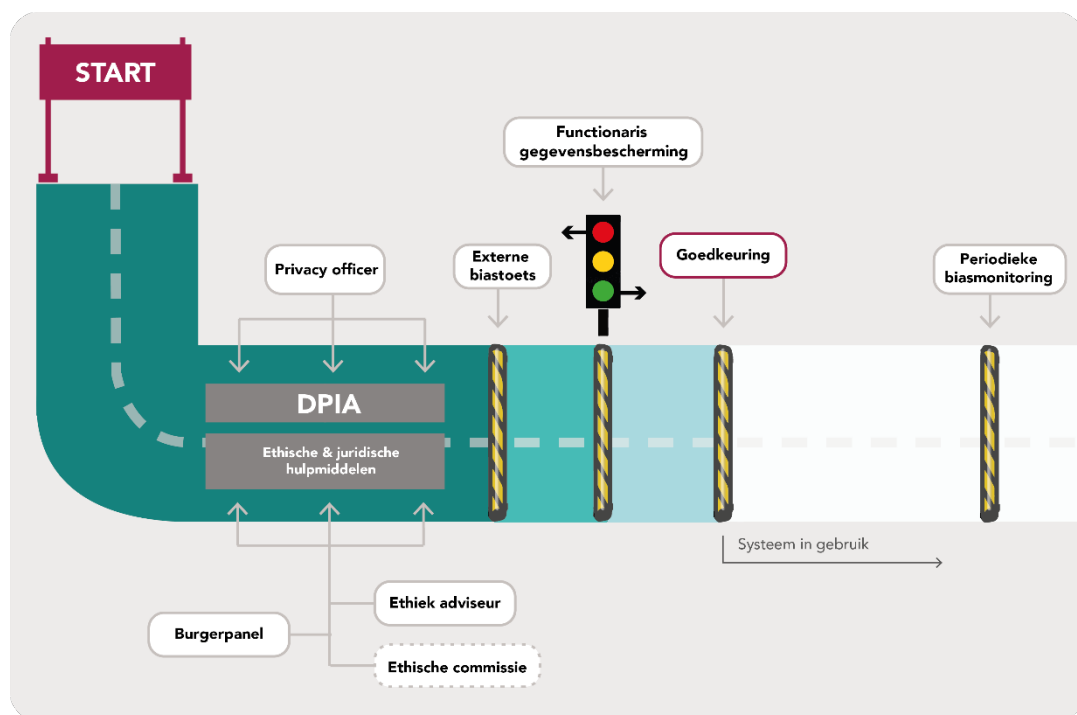
Tabel 1 geeft een overzicht van de genoemde maatregelen. Deze zijn in ontwikkeling. Globaal lijken de drie onderzochte organisaties vergelijkbare type maatregelen te hebben ingesteld. We zien wel verschillen tussen de maatregelen, bijvoorbeeld in de mate van detail – zoals de ontwikkelde normenkaders – maar ook wat betreft de aandacht voor verschillende mensenrechten. We gaan in hoofdstuk vier nader in op deze bevindingen.

Tabel 1 Overzicht belangrijkste maatregelen

Type	Specifieke maatregel
Werkprocessen	• Privacy by design • Model Risk Management beleid • Externe bias- en kwaliteitstoets • Periodieke bias monitoring • Jaarlijks onderhoudsinterval • Vastleggen beslissingen tijdens ontwikkelproces • Klachtenprocedure opnieuw bekijken
Functies	• Functionaris gegevensbescherming • Privacy officer • Adviseur Recht en Ethiek • Kwartiermaker data-autoriteit • Ethische commissie
Hulpmiddelen	• DPIA/GEB • CJIB-AI impact assessment • Burgerpanels en Usability Lab • Kompas Data Ethiek • Normenkader • Juridische handreiking
Expertise	• Meer expertise in ontwikkelteams • Ethische (bewustwordings)sessies over bias, fairness met o.a. ethische matrix en aan de hand van interne werkgroep • Ethische certificering en opleiding • Kennissessies met externe experts • (in de toekomst) overleg tussen verschillende functies • Programma 'de Bedoeling' • Serieus nemen ethische bezwaren
Technische maatregelen	• Bouwen in eigen beheer • Ontwikkelen in afgeschermd omgeving • Geen unsupervised learning toepassen • Opzetten van algoritmeregister

De vijf typen maatregelen geven samen een beeld van hoe de ontwikkeling van profilerende systemen binnen de onderzochte organisaties tot stand komt. De ontwikkeling doorloopt verschillende stappen, waarbinnen verschillende hulpmiddelen, processen, en functies een rol hebben. Figuur 1 toont de ontwikkeling die een systeem globaal in een organisatie doorloopt, met verschillende *checks en balances*. De figuur dient om een globale indruk te geven en is geen exacte weergave van de processen bij de drie organisaties.

Figuur 1 De weg van een profilerend systeem: van plan naar uitvoering



NB: Dit is een impressie van de ontwikkeling. Het is geen exacte weergave van alle stappen die gezet worden.

Met het oog op de toekomst noemen de gesproken organisaties een brede reeks uitdagingen, die we hebben ingedeeld in drie categorieën: pleidooien voor meer centrale regie met betrekking tot de beschikbare hulpmiddelen, meer inzet op het verbeteren van de dienstverlening, waaronder het met profilering helpen van burgers, en verdere investeringen in de ethische en juridische expertise van medewerkers.

4 Analyse

4.1 Inleiding

In het vorige hoofdstuk beschreven we de belangrijkste maatregelen die drie onderzochte uitvoeringsorganisaties hebben genomen om de relevante mensenrechten te waarborgen. Dit hoofdstuk zal op basis van die inventarisatie en de gehouden externe consultaties (zie bijlage 3), deze onderzoeksresultaten analyseren en plaatsen in de maatschappelijke context waarbinnen uitvoeringsorganisaties gebruik maken van algoritmes. We doen dat aan de hand van de volgende tien bevindingen:

1. De onderzochte uitvoeringsorganisaties staan aan het begin van de inzet en ontwikkeling van profilerende algoritmische systemen.
2. Hoewel andere doelen onderzocht worden, ligt de nadruk bij de inzet van profilering tot dusver op het versterken van de handhaving.
3. De onderzochte uitvoeringsorganisaties zien risico's van profilerende systemen en nemen daar maatregelen voor.
4. Het raamwerk voor privacybescherming is het meest volwassen, maar de uitvoeringsorganisaties zien ruimte voor verbetering.
5. De onderzochte uitvoeringsorganisaties zoeken bij het mitigeren van de overige mensenrechtelijke risico's naar houvast.
6. De onderzochte uitvoeringsorganisaties zoeken bij de oprichting van een ethische commissie naar de voorwaarden waaronder deze succesvol kan zijn.
7. Bias-detectie krijgt veel aandacht, en is decentraal georganiseerd.
8. De uitvoeringsorganisaties zoeken naar de beste omgang met informatierechten en brengen een verschillende balans aan tussen actieve en passieve informatievoorziening.
9. Ten aanzien van de procedurele rechten geven de uitvoeringsorganisaties aandacht aan menselijke tussenkomst en uitlegbaarheid. De organisaties verschillen in de beoordeling of de klachtenprocedures aanpassing behoeven.
10. De onderzochte uitvoeringsorganisaties geven aan de inzet van profilering proportioneel te vinden.

Tijdens de externe consultaties zijn deze bevindingen in voorlopige vorm voorgelegd. Op basis van de gevoerde discussies zijn de bevindingen verfijnd. De geraadpleegde experts gaven aan dat de bevindingen een herkenbaar beeld

vormen. We geven in dit hoofdstuk ook weer waar tussen de experts verschil van inzicht bestond.

4.2 Bevindingen

1. De onderzochte uitvoeringsorganisaties staan aan het begin van de inzet en ontwikkeling van profilerende algoritmische systemen.

De onderzochte uitvoeringsorganisaties zetten profilerende systemen nog beperkt in. Ze zijn bezig er een klein aantal te ontwikkelen. Het gaat om pilots die vanaf eind 2021/begin 2022 in de reguliere dienstverlening worden ingezet. Deze pilots kenden een ontwikkeltijd van rond de twee à drie jaar, waarbij, zoals we hieronder zullen zien, over tal van technische, wettelijke en ethische aspecten werd nagedacht. Het gaat om lerende systemen, *supervised learning*, waarbij profilerende algoritmes de besluitvorming ondersteunen. In het geval van fraude-detectie helpen de systemen met het selecteren van gevallen voor nader onderzoek. Ze vormen een aanvulling op steekproeven die al werden uitgevoerd, of komen daarvoor in de plaats. Een mens gaat vervolgens nader onderzoek doen.

De pilots met profilerende systemen vormen een nieuwe technische stap die de onderzochte uitvoeringsorganisaties nemen. Voor het merendeel gebruiken deze organisaties namelijk systemen die direct de uitvoering van rechtsregels automatiseren, bijvoorbeeld het bepalen en uitkeren van de kinderbijslag of de werkloosheidsuitkering. Dit gaat om zogenoemde *rule-based* algoritmes. De uitvoeringsorganisaties verwachten in de toekomst profilerende systemen meer te zullen gaan gebruiken.

Deze verwachting past in de brede maatschappelijke trend dat AI als de technologie van de toekomst wordt gezien. Tegelijkertijd valt de inzet van profilerende systemen bij de onderzochte uitvoeringsorganisaties mee, zeker gezien de enorme wetenschappelijke, commerciële en maatschappelijke aandacht voor de technologie. Het is lastig om te zeggen in hoeverre de uitvoeringsorganisaties daarmee voor- of juist achterlopen op de rest van de overheidsorganisaties. Een precies overzicht valt uit de huidige rapporten niet op te maken (ARK 2021, WRR 2016; TNO 2019; TNO 2021). Wel wijzen de wetenschappelijke rapporten erop dat ook andere uitvoeringsorganisaties experimenteren met profilerende algoritmes en dat er sprake is van een stijgende trend.

2. Hoewel andere doelen onderzocht worden, ligt de nadruk bij de inzet van profilering tot dusver op het versterken van de handhaving.

De onderzochte organisaties ontwikkelen profilering voor verschillende doeleinden: voor het handhaven van de wet, het ondersteunen van burgers in de dienstverlening en het tegengaan van schulden. De meeste van de besproken pilots zijn gericht op handhaving. Tijdens de gesprekken en gehouden consultaties kwam naar voren dat de behoefte bestaat binnen de uitvoeringsorganisaties om profilering vaker in te zetten om burgers gericht te ondersteunen.

De pilots gericht op handhaving moeten de uitvoeringsorganisaties efficiëntie bieden. Een handhavingsafdeling zou minder willekeurige steekproeven hoeven te doen en onderzoeken uitvoeren. Hetzelfde aantal handhavers zou op een meer gerichte manier kunnen werken en daardoor meer fraude detecteren. Ook geven medewerkers in de gesprekken aan dat profilering kan helpen bij een tijdige detectie van een onterechte aanvraag. Daarmee wordt voorkomen dat de organisatie pas in een laat stadium een onterechte aanvraag opmerkt, en de burger geconfronteerd wordt met hoge terugvorderingen. Deze beloftes spreken de onderzochte uitvoeringsorganisaties aan. Ze voelen zich genoodzaakt om de mogelijkheden van profilering te verkennen, vanwege de roep vanuit de politiek en de samenleving om enerzijds effectieve fraudedetectie en anderzijds een soepele dienstverlening.

In tegenstelling tot handhaving, levert de profilering die directer gericht is op het helpen van burgers niet altijd efficiëntie op voor de desbetreffende uitvoeringsorganisatie. Tijdens de externe consultaties kwam naar voren dat het soms moeilijk is de businesscase rond te krijgen: de organisatie neemt meer dienstverlenende taken op zich. Zonder profilerend systeem lagen de kosten om burgers te helpen bij andere organisaties. Overigens zou de profilering om burgers te helpen de overheid als geheel wel kosten kunnen besparen – problemen kunnen worden voorkomen, of burgers kunnen eerder worden geholpen.

3. De onderzochte uitvoeringsorganisaties zien risico's van profilerende systemen en nemen daar maatregelen voor.

Bij het ontwikkelen van de pilots zijn de uitvoeringsorganisaties niet over een nacht ijs gegaan. Uit de gesprekken komt naar voren dat er over tal van aspecten uitvoerig is nagedacht, ook over de mogelijke risico's. De ontwikkeling van de pilots nam, zoals hierboven al aangegeven, enkele jaren in beslag. In de interviews werden deze risico's vaak niet gearticuleerd als mensenrechtelijke risico's, maar in termen van ethiek, of in termen van concrete wettelijke

verplichtingen, zoals opgenomen in de AVG. Hieronder gaan we op dat aspect apart in.

De door de uitvoeringsorganisaties geïdentificeerde risico's komen grotendeels overeen met de mensenrechtelijke risico's beschreven in hoofdstuk 2: mogelijke schendingen van privacy, non-discriminatie, uitlegbaarheid, transparantie en *fairness*. Ook zagen we in hoofdstuk 3 dat de organisaties vele maatregelen nemen om risico's tijdig te identificeren en te mitigeren. Zo kiezen ze ervoor de profilerende systemen volledig zelf te ontwikkelen. Hiervoor hebben ze de afgelopen jaren *data-scientists* in dienst genomen en ontwikkelafdelingen opgericht.

Dit in tegenstelling tot *rule-based* systemen, die vaak worden ingekocht. In de externe consultaties werd juist de impact van *rule-based* algoritmes benadrukt. Daarbij werd aangegeven dat voor deze categorie diverse beheersingsmaatregelen (zoals een algoritmeregister) lastiger zijn uit te voeren, omdat deze systemen al jaren oud zijn en niet vanaf het begin is bijgehouden wat is ontwikkeld en hoe.

Ook blijkt uit de gesprekken dat voor bepaalde mensenrechten meer maatregelen worden genomen dan voor andere. Bevindingen 4 en 5 gaan hier verder op in.

4. Het raamwerk voor privacybescherming is het meest volwassen, maar de uitvoeringsorganisaties zien ruimte voor verbetering.

Het privacyraamwerk is het verst ontwikkeld en ingevoerd, en dat mag niet verbazen. Specifieke privacywetgeving kent met de Wet persoonsregistratie (ingevoerd in 1989) en de Wet bescherming persoonsgegevens (uit 2001) al een lange traditie. De AVG uit 2018 verfijnt deze kaders en schrijft onder meer voor wanneer DPIA's, *privacy officers* en functionarissen gegevensbescherming nodig zijn. Wel blijkt uit de gesprekken en externe consultaties dat dit privacyraamwerk in de praktijk nog niet optimaal wordt toegepast en ingevuld. Zo ziet de functionaris gegevensbescherming van SVB graag meer expertise ten aanzien van privacykwesties bij de ontwikkelteams en de *privacy officers* – zodat meer steekproefsgewijs gewerkt kan worden.

De geïnterviewde functionaris gegevensbescherming schrijft dit toe aan de complexiteit van de AVG. Zo is het niet eenvoudig te beoordelen of een bepaalde manier van data verwerken wel of niet is toegestaan. Ook is lastig te beoordelen of een organisatie voldoende wettelijke grondslag heeft om een bepaalde dienst op te zetten. Daarbij speelt ook de scope van de beoordeling een rol: kijk je of een specifiek systeem voldoet aan de AVG, of kijk je naar de impact van een reeks

systemen op de privacy van mensen? Door breder te kijken naar de impact van systemen, komt meer aandacht voor wat bij de SVB ‘de bedoeling’ wordt genoemd – wat voor impact wil een uitvoeringsorganisatie hebben op de levens en rechten van burgers?

Ten aanzien van de onderzochte profilerende systemen is het meermaals voorgekomen dat een functionaris gegevensbescherming de plannen had afgekeurd. Enkele keren ontstond er stevige discussie binnen de organisatie. Functionarissen gegevensbescherming kunnen dus een cruciale rol spelen bij het waarborgen van privacy binnen een organisatie. De FG’s en ethiek adviseurs gaan ook samenwerken, om de nodige afwegingen te kunnen maken.

Uit de gesprekken en consultaties blijkt bovendien dat er druk op de ketel staat. De uitvoeringsorganisaties staan voor complexe uitdagingen en de computertechnologie belooft hieraan veel bij te dragen. Vanuit de kant van de *data scientists*, maar ook vanuit de hoek van handhaving en de dienstverlening ontstaat druk om die technologie te gaan gebruiken. Zo gaf een medewerker van het UWV aan dat het management bij *data scientists* moest aangeven dat *unsupervised learning* ongewenst was vanuit het oogpunt van uitlegbaarheid, ook al betoogden de *data scientists* dat deze technologie hogere effectiviteitscores kon opleveren.

Ten slotte blijkt uit de gesprekken dat de positionering van de functionarissen gegevensbescherming en de *privacy officers* verschilt. De functionarissen gegevensbescherming bij het UWV en de SVB zitten intern en toetsten vrijwel alle ingevulde DPIA’s. Zij geven aan soms kritische feedback te geven en bepaalde voorgestelde toepassing te hebben geblokkeerd. De functionaris gegevensbescherming bij het CJIB staat op grotere afstand en vervult dezelfde functie bij andere organisaties die onder het ministerie van Justitie en Veiligheid vallen. Uit navraag blijkt dat de *privacy officer* in deze constructie een grotere rol speelt bij het toetsen van DPIA’s.

5. De onderzochte uitvoeringsorganisaties zoeken bij het mitigeren van de overige mensenrechtelijke risico’s naar houvast.

Uit het overzicht van de maatregelen blijkt dat de onderzochte organisaties niet alleen bezig zijn om de bescherming van privacy te versterken, maar ook die van een aantal andere mensenrechtelijke aspecten, met name het voorkomen van bias en het voorzien in uitlegbaarheid. De organisaties geven daarbij herhaaldelijk aan dat deze maatregelen, zoals het opzetten van een ethische commissie en het ontwikkelen van een ethische code, pas recent worden ontwikkeld.

Bij het adresseren van deze overige risico's zoeken de gesproken organisaties naar houvast: welke combinatie van maatregelen past hierbij? Niemand twijfelt eraan dat bias voorkomen moet worden en een mate van transparantie betracht moet worden. Maar men zoekt naar de beste manier om deze bescherming vorm te geven en te begrijpen. Daarbij worden in twee richtingen maatregelen genomen.

Allereerst investeert men in het zo goed mogelijk toepassen van de bestaande juridische kaders, die volgens de onderzochte uitvoeringsorganisaties ruimte voor interpretatie overlaten, zoals hierboven al genoemd. Daarbij ligt de nadruk op het interpreteren van de AVG. Zo werkt de functionaris gegevensbescherming van de SVB bij het afhandelen van lastige interpretatiekwesties regelmatig samen met de adviseur Recht en Ethiek. Soms wordt er over een bepaald thema advies gevraagd over het toepassen van de rechtsregels. Zo schreef de landsadvocaat Pels Rijcken op verzoek van het CJIB een handreiking over de juridische vereisten waaraan het innovatielab moet voldoen en welke stappen de medewerkers in de diverse fasen van ontwikkeling en toepassing van het algoritmen moeten zetten. Uit deze voorbeelden blijkt dat de uitvoeringsorganisaties soms een interpretatieslag moeten maken om de regels in concrete situaties toe te passen. Op die momenten helpen additionele hulpmiddelen en expertise om de wet weer hanteerbaar te maken. Tegelijkertijd zeggen sommige uitvoeringsorganisaties dat ze door alle hulpmiddelen die de afgelopen periode zijn verschenen, door de bomen het bos soms niet meer zien.

Ten tweede ontwikkelen de uitvoeringsorganisaties nieuwe normenkaders en trainingsprogramma's, onder de noemer van ethiek. Zo liet het UWV een Kompas Data Ethiek opstellen, dat als leidraad moet dienen voor het handelen van alle medewerkers. Uit dit soort ethische kaders spreekt een verlangen om niet alleen binnen de grenzen van de wet te handelen, maar ook te voldoen aan wat maatschappelijk wenselijk is. Op de achtergrond speelt hier de kinderopvangtoeslagenaffaire mee, waarbij dikwijls binnen de grenzen van de wet maatschappelijk onaanvaardbare besluiten zijn genomen. De onderzochte uitvoeringsorganisaties willen in staat zijn om uit te maken wanneer het wettelijk toegestane niet langer het fatsoenlijke is. Ze willen medewerkers voldoende equiperen om ethische afwegingen met elkaar te maken.

Opvallend is dat het in de interviews meer gegaan is over ethiek dan over mensenrechten. Tijdens deze gesprekken vroegen we naar de mogelijke mensenrechtelijke risico's, waarop de onderzochte uitvoeringsorganisaties vaak antwoorden in termen van specifieke wettelijke vereisten, zoals opgenomen in de AVG, of in termen van ethiek. Ze noemden bijvoorbeeld het ontwikkelen van ethische hulpmiddelen (Kompas Data Ethiek, ethische matrix) of het aanstellen

van een ethische adviseur of ethische commissie. De onderzochte organisaties gaven in dit kader het belang aan van investeringen in het ethische karakter van de uitvoeringsorganisatie. Ook werd in de gesprekken minder gesproken over andere juridische kaders dan de AVG, zoals de Algemene wet bestuursrecht en de Algemene beginselen van behoorlijk bestuur. Tijdens de consultaties werd opgemerkt dat juist deze kaders de kern vormen van verantwoord handelen door overheidsorganisaties, ook bij de inzet van algoritmes. Er is discussie gaande over hoe deze beginselen toegepast zouden moeten worden, en of nadere wettelijke normering nodig is (Van Eck et al. 2018; Wolswinkel, 2021).

De bevinding om te investeren in ethiek past binnen een bredere trend bij uitvoeringsorganisaties. Zo richten andere uitvoeringsorganisaties, zoals politie en gemeentes ethische commissies en klankbordgroepen op. Bij tal van organisaties worden ethische waardenkaarten opgesteld.⁴

6. De onderzochte uitvoeringsorganisaties zoeken bij de oprichting van een ethische commissie naar de voorwaarden waaronder deze succesvol kan zijn.

Zowel het UWV als het CJIB zijn betrokken bij de oprichting van ethische commissies. Uit de interviews komt naar voren dat zij zoeken naar de juiste invulling van deze organen. Hierbij hebben zij een reeks risico's in het vizier. Zo willen deze organisaties dat de ethische commissies goed samenhangen met het bestaande toezicht door de functionaris gegevensbescherming en geen dubbel werk doen. Ook wordt gezocht naar een taakstelling die goed aansluit bij de beschikbaarheid en capaciteit van een ethische commissie. Want als de commissie slechts eens in het kwartaal samenkomt, moet ze een bepaald takenpakket natuurlijk wel uit kunnen voeren. Bovendien moet duidelijk zijn op basis van welke normen de ethische commissie haar werk doet.

Tijdens de externe consultaties ging veel aandacht uit naar de mogelijkheid om ethische commissies als additionele toezichthouders te laten functioneren. De commissie geeft dan een bindend oordeel over de specifieke inzet van een systeem. Een dergelijke opzet is bijvoorbeeld in de wetenschap gebruikelijk, waarbij een ethische commissie een bindend oordeel geeft over het wel of niet uitvoeren van een specifiek wetenschappelijk onderzoek. Dit idee werd echter kritisch ontvangen. Toezicht zou aan de hand van bindende juridische normen moeten verlopen en ethische veroordelingen zouden onnodig stigmatiserend kunnen werken. Diverse experts zagen er meer in om een ethische commissie in te zetten als een adviesorgaan en aanjager van ethisch debat binnen de

⁴ <https://vng.nl/artikelen/de-ethiek-van-datahonger-de-opmars-van-de-ethische-commissie>

organisatie. In ieder geval bleek uit de consultaties de wens om te voorkomen dat ethiek belegd wordt bij (externe) experts, in plaats van een verantwoordelijkheid van de organisatie en haar medewerkers.

7. Biasdetectie krijgt veel aandacht en is decentraal georganiseerd.

De maatregelen voor het opsporen en verhelpen van bias krijgen binnen de onderzochte organisaties veel aandacht, van het opzetten van biastrainingen en het raadplegen van wetenschappelijke experts tot het uitvoeren van interne en externe toetsen. Daarbij valt op dat de biasmaatregelen in twee opzichten decentraal zijn georganiseerd.

Ten eerste hebben de uitvoeringsorganisaties ieder hun eigen samenwerking opgezet met een externe accountant of toetsingsbureau, die weer met een eigen toetsingsmethode bias probeert te achterhalen. De uitvoeringsorganisaties lijken op dit gebied allemaal zoekende naar de beste manier om bias tegen te gaan, en zijn daarbij aan het pionieren. Zo worden bepaalde bias-detectiemethodes geselecteerd, en proberen ontwikkelteams expliciet te maken op basis van welke principes iets eerlijk is, of juist discriminerend. Ten tweede ontbreekt een centraal gremium waar biasdetectie onderling wordt afgestemd, en gezamenlijk gezocht wordt naar de beste methodiek. Iedere organisatie zoekt nu zelf contact met wetenschappers om hierover te overleggen.

8. De uitvoeringsorganisaties zoeken naar de beste omgang met informatierechten en brengen een verschillende balans aan tussen actieve en passieve informatievoorziening.

De uitvoeringsorganisaties zijn zich bewust van het belang de burger mee te nemen bij de ontwikkeling van profilerende systemen. Daarbij is het zaak zowel passief als actief in informatie te voorzien. De uitvoeringsorganisaties verschillen in de manier waarop ze deze twee aspecten combineren en vormgeven. Een medewerker van het UWV gaf aan dat beslissingen die genomen worden tijdens het ontwikkelproces, of die genomen worden door het profilerende algoritme, vooral goed vastgelegd worden, en uitgelegd kunnen worden als men ernaar vraagt. Denk bijvoorbeeld aan journalisten die een WOB-verzoek doen, of een rechter die bepaalde vragen stelt. Tijdens de externe consultatie stond discussie over de mogelijkheid dat meer transparantie fraudeurs in de hand zou werken. Daarbij gaven diverse experts aan dat een bepaalde vorm van transparantie mogelijk is, en belangrijk is. De mate van passieve of actieve informatievoorziening, kan ook afhangen van de fase waarin een pilot zich bevindt.

In algemene zin zijn de organisaties zoekende naar de beste manier om over profilerende algoritmes te communiceren. Zo wil het UWV niet spreken over algoritmes, maar over risicoscans omdat die eerste term verontrustende associaties zou kunnen oproepen. Ook ten aanzien van dit aspect blijkt dat de uitvoeringsorganisaties zich bewust zijn van het onrustige maatschappelijke vaarwater waarin het gebruik van algoritmische profilering terecht is gekomen. Ze zoeken naar manieren om op een vertrouwenwekkende manier over profilering te communiceren.

9. Ten aanzien van de procedurele rechten geven de uitvoeringsorganisaties aandacht aan menselijke tussenkomst en uitlegbaarheid. Ze verschillen in de beoordeling of de klachtenprocedures aanpassing behoeven.

Uit de gesprekken bleek dat de onderzochte uitvoeringsorganisaties zoekende zijn naar de beste manier om het principe van uitlegbaarheid in te vullen. Dit sluit aan bij de huidige wettelijke situatie. Nederland kent beperkte specifieke wettelijke bepalingen met betrekking tot informatie en uitleg over een algoritmisch besluit of een algoritme dat een besluit voorbereidt (Wolswinkel 2021). Zo is niet precies duidelijk wat voor menselijke tussenkomst in welke situatie eigenlijk vereist is.

Het UWV vult uitlegbaarheid mede in door in een specifieke vorm van menselijke tussenkomst te voorzien. Zo gebruikt een medewerker van het UWV de term ‘menselijke tussenkomst’ om het proces te omschrijven van medewerkers op de handhavingsafdeling die niet weten welke burgers vanwege een algoritmische profilering onderzocht moeten worden. De handhavers interveniëren niet in de werking van het algoritme (zoals de data scientists dat doen), maar gaan aan de slag met de uitkomsten ervan. De menselijke handhaver kan op basis van het eigen onderzoek uitleggen waarom er sprake is van fraude, en ook *hoe* de handhaver tot dat oordeel is gekomen. Het handelen van de uitvoeringsorganisatie wordt daarom uitlegbaar geacht. Dat is opvallend, omdat bij deze vorm van menselijke tussenkomst de medewerker geen uitleg geeft over de precieze werking van het algoritme.

De uitvoeringsorganisaties trachten ook op andere manieren uitlegbaarheid te waarborgen. De organisaties kiezen er bewust voor om geen *unsupervised learning* te gebruiken, zelfs niet als deze systemen betere resultaten boeken, omdat de technologie niet verenigbaar is met dit principe. Tegelijkertijd is duidelijk dat de ontwikkelaars wel experimenteren met deze technieken en ze spoedig zouden kunnen inzetten als de opvattingen over uitlegbaarheid zouden veranderen.

Ten slotte denken uitvoeringsorganisaties verschillend over de vraag of het vanwege de invoering van profilering nodig is om opnieuw naar de klachtenprocedures te kijken. Het CJIB heeft hiervoor gekozen, de SVB en het UWV noemden dit in de gevoerde gesprekken niet.

10. De uitvoeringsorganisaties geven aan de inzet van profilering proportioneel te vinden.

Om te bepalen of een systeem wordt ingezet in lijn met mensenrechten, speelt het afwegen van proportionaliteit een belangrijke rol. Is het mogelijke risico van het systeem in balans met de maatschappelijke opbrengst ervan, en zijn er geen alternatieven die hetzelfde resultaat opleveren zonder inbreuk te maken op mensenrechten (het subsidiariteitsbeginsel)?

Uit de interviews en externe consultaties komt naar voren dat de uitvoeringsorganisaties de inzet van profilerende systemen beoordelen als risicovol, maar de moeite waard. Sterker nog: het wordt gezien als de logische consequentie van de eisen die de maatschappij en de politiek aan hen stellen. Ze hebben immers beperkte middelen, maar moeten toch in hoge mate fraude opsporen en tijdige diensten verlenen. Als de profilerende systemen eenmaal zijn ontwikkeld, kunnen ze er volgens de organisaties voor zorgen dat die met een gelijk aantal handhavers meer fraude kunnen opsporen. De bandbreedte van dit positieve effect werd in de interviews breed ingeschat: bijvoorbeeld 2,5 tot 8 keer effectiever dan steekproeven. Het was voor de onderzoekers niet mogelijk de verschillende vermelde cijfers met elkaar te vergelijken en daarover een oordeel te vellen.

Tijdens de externe consultaties werd er verschillend gedacht over proportionaliteit van de systemen. Een deel van de experts ziet, net als de uitvoeringsorganisaties, de systemen als proportioneel. Mogelijke risico's op schending van mensenrechten worden zo klein mogelijk gemaakt door alle genomen maatregelen en zijn in balans met de maatschappelijke opbrengst.

Tegelijkertijd is duidelijk dat de uitvoeringsorganisaties veel maatregelen moeten nemen en naleven om algoritmes op verantwoorde wijze in te zetten. Daarbij gaat het bijvoorbeeld om het aannemen van *datascientists*, het opstellen van een ethisch kompas, het organiseren van ethische commissies, het laten uitvoeren van impact assessments, investeren in biasdetectie en het geven van ethiektrainingen. Ook die kosten zijn onderdeel van een volledige weging. Meerdere experts gaven in dit verband aan dat de zogenoemde nuloptie bij het beoordelen van de proportionaliteit van een systeem op tafel moet liggen.

Overigens verschilt het antwoord op de proportionaliteitsvraag per systeem. Tijdens de externe consultaties werd ingebracht dat bij de weging van systemen die bedoeld zijn om de burger te helpen, de proportionaliteit in verhouding tot handhavingssystemen gunstiger zou kunnen uitvallen.

5 Conclusies en aanbevelingen

5.1 Inleiding

Het tijdperk van de algoritmische uitvoeringsorganisatie is al enkele decennia gaande. Organisaties als het UWV, de SVB en het CJIB gebruiken inmiddels op grote schaal algoritmes om diensten uit te voeren. Daarbij draaide het lange tijd hoofdzakelijk om het toepassen van *rule-based* algoritmes. Het structureel in gaan zetten van algoritmische profilering, vormt in deze ontwikkeling de volgende stap. Door het gedrag van burgers met lerende algoritmes te voorspellen, zouden uitvoeringsorganisaties beter regels kunnen handhaven en burgers sneller kunnen helpen met maatwerk.

De nieuwe stap die uitvoeringsorganisaties zetten, verdiept ook de mensenrechtelijke uitdagingen. Het is bijvoorbeeld door de inzet van lerende systemen lastiger om discriminatie op te sporen en de keuzes van die systemen goed uit te leggen of volledig te kennen. Het is daarom van belang om te kijken hoe de bescherming van mensenrechten bij de inzet van profilerende systemen versterkt kan worden. Te meer omdat in het recente verleden uitvoeringsorganisaties bij de bescherming van mensenrechten soms ernstig tekort zijn geschoten, zoals bij de kinderopvangtoeslagenaffaire.

Uit deze misstanden bleek ook dat het voorkomen van mensenrechtenschendingen niet alleen een kwestie is van deugdelijke algoritmes. Algoritmes worden ontwikkeld en toegepast in een complexe overheidspraktijk met een bepaalde cultuur, onderlinge verhoudingen, expertise en organisatie. Deze gehele socio-technische context zal versterkt moeten worden. Daarom inventariseerde deze studie maatregelen op het gebied van functies, expertise, hulpmiddelen en werkprocessen. Hoewel er nationaal en internationaal veel richtlijnen, handreikingen en toetsingskaders zijn gepubliceerd, is er weinig zicht op hoe uitvoeringsorganisaties de bescherming van mensenrechten in de praktijk vormgeven. Daarom heeft het Rathenau Instituut deze studie uitgevoerd, op verzoek van het ministerie van Binnenlandse Zaken en Koninkrijksrelaties.

Op basis van interviews inventariseerden we welke maatregelen drie landelijke uitvoeringsorganisaties momenteel nemen. We spraken met dertien medewerkers van verschillende afdelingen bij het UWV, het CJIB en de SVB. De voorlopige bevindingen en suggesties voor mogelijke versterking van mensenrechten, zijn besproken tijdens twee externe consultaties (zie bijlage 3). Dit hoofdstuk vat de

belangrijkste bevindingen samen en formuleert zeven aanbevelingen om de bescherming van mensenrechten bij de inzet van profilerende systemen in de uitvoering te versterken.

5.2 Maatregelen voor profilerende systemen

Van pilots naar productie

De onderzochte uitvoeringsorganisaties ontwikkelen momenteel profilerende algoritmes. Het gaat om pilots die vanaf eind 2021/begin 2022 in de reguliere werkwijze van organisatie worden ingezet, en een ontwikkeltijd kenden van circa twee tot drie jaar. De doelen van de profilering lopen uiteen. De profilerende algoritmes van het UWV en de SVB worden ontwikkeld om fraude en foutief gebruik van regelingen te detecteren. Het UWV kijkt met de sollicitatiescan ook hoe burgers gericht kunnen worden ondersteund bij het vinden van werk. Het profilerende algoritme van het CJIB is gericht op het innen van openstaande boetes en het voorkomen van schuldenoploop. De nadruk ligt bij de pilots vooralsnog op het versterken van de handhaving. De onderzochte uitvoeringsorganisaties willen profilering in de toekomst meer gaan inzetten om burgers te ondersteunen. Tegelijkertijd kan het lastig zijn voor uitvoeringsorganisaties om dit financieel rond te krijgen.

Omgang met mensenrechtelijke risico's

Het valt op dat de organisaties de risico's vooral in termen van ethische en specifieke wettelijke eisen bespreken en daarbij niet of nauwelijks verwijzen naar mensenrechtenverdragen. Dat wil niet zeggen dat zij zich bij de ontwikkeling van hun profilerende algoritmes niet bewust zijn van mogelijke risico's. In de gesprekken werden zaken als privacy, transparantie, uitlegbaarheid en menselijke tussenkomst genoemd. Grofweg komen de genoemde kwesties overeen met de indeling die we in hoofdstuk twee beschreven: privacyrechten, gelijkheidsrechten, vrijheidsrechten en procedurele rechten. In de gesprekken leek de nadruk vooral te liggen op privacyrechten en het recht op non-discriminatie. De onderzochte uitvoeringsorganisaties beoordelen de inzet van profilerende systemen als proportioneel. Ze zien risico's, maar concluderen dat die voldoende te mitigeren zijn en dat er voldoende baten tegenover staan.

Ethische en wettelijke eisen

De genomen maatregelen dienen twee doelen. Enerzijds zijn ze gericht op het zo goed mogelijk toepassen van bestaande juridische kaders. De onderzochte organisaties proberen met de maatregelen de vertaalslag te maken van de soms abstracte wetgeving naar hun complexe uitvoeringspraktijk. Anderzijds investeren de organisaties in de ontwikkeling van ethische expertise, vanuit een verlangen

om niet alleen binnen de grenzen van de wet te handelen, maar ook te voldoen aan wat maatschappelijk wenselijk is. De maatregelen dienen dus ook de organisatie te stimuleren om verder te kijken dan strikt de wettelijke vereisten.

Breed scala van maatregelen

De drie organisaties werken aan een scala van maatregelen om in lijn te werken met mensenrechten. Tabel 1 geeft een overzicht van de genoemde maatregelen. Voorbeelden van ingestelde functies zijn een Commissie Data Ethiek of een adviseur Recht en Ethiek. De organisaties ontwikkelen specifieke hulpmiddelen, zoals een Kompas Data Ethiek, een normenkader of een impact assessment. Ook wordt vastgelegd wie binnen de organisatie waarvoor verantwoordelijk is. Op het gebied van werkprocessen is bijvoorbeeld bias-toetsing een standaard onderdeel geworden van de systeemontwikkeling. Tot slot investeren de organisaties in ethische en juridische expertise, onder meer bij de ontwikkelteams.

De onderzochte organisaties besteden veel aandacht aan het detecteren van bias. Opvallend is dat deze toetsing decentraal is georganiseerd. De onderzochte uitvoeringsorganisaties starten hun eigen samenwerkingen met externe partijen om onder andere bias te toetsen. De kennisontwikkeling vindt ook per organisatie plaats. Ten aanzien van het recht op informatie, zoeken de onderzochte uitvoeringsorganisaties vooral naar de balans tussen actieve en passieve informatievoorziening, waarbij ze verschillen wat betreft de genomen maatregelen. Op het gebied van procedurele rechten nemen de organisaties maatregelen om te zorgen voor menselijke tussenkomst en uitlegbaarheid. Ze verschillen van mening over de noodzaak de bestaande klachtenprocedures aan te passen.

De waarde van de genomen maatregelen zal zich in de toekomst nog moeten bewijzen. Het bestaan van maatregelen als een ethische commissie of de ontwikkeling van eigen hulpmiddelen biedt op zichzelf geen garantie dat mensenrechten bij het gebruik van profilerende algoritmische systemen in de uitvoeringspraktijk beschermd zijn. Uiteindelijk dient elke organisatie de inzet van algoritmen te leren afwegen.

Tabel 1 Overzicht belangrijkste maatregelen

Type	Specifieke maatregel
Werkprocessen	• Privacy by design • Model Risk Management beleid • Externe bias- en kwaliteitstoets • Periodieke bias monitoring • Jaarlijks onderhoudsinterval • Vastleggen beslissingen tijdens ontwikkelproces • Klachtenprocedure opnieuw bekijken
Functies	• Functionaris gegevensbescherming • Privacy officer • Adviseur Recht en Ethiek • Kwartiermaker data-autoriteit • Ethische commissie
Hulpmiddelen	• DPIA/GEB • CJIB-AI impact assessment • Burgerpanels en Usability Lab • Kompas Data Ethiek • Normenkader • Juridische handreiking
Expertise	• Meer expertise in ontwikkelteams • Ethische (bewustwordings)sessies over bias, fairness met o.a. ethische matrix en aan de hand van interne werkgroep • Ethische certificering en opleiding • Kennissessies met externe experts • (in de toekomst) overleg tussen verschillende functies • Programma 'de Bedoeling' • Serieus nemen ethische bezwaren
Technische maatregelen	• Bouwen in eigen beheer • Ontwikkelen in afgeschermd omgeving • Geen unsupervised learning toepassen • Opzetten van algoritmeregister

Maatregelen in ontwikkeling

De genomen maatregelen zijn nog in ontwikkeling. De onderzochte organisaties hebben bijvoorbeeld vragen over de rol en invulling van een op te richten ethische commissie en over het bespreekbaar maken van mogelijke bias bij de medewerkers. Met het oog op de toekomst noemen de organisaties een brede reeks uitdagingen, die we hebben ingedeeld in drie categorieën: pleidooien voor meer centrale regie met betrekking tot de beschikbare hulpmiddelen, meer inzet op het verbeteren van de dienstverlening, waaronder het met profilering helpen van burgers, en verdere investeringen in de ethische en juridische expertise van medewerkers.

Ook het meest volwassen raamwerk, gericht op privacy, kent ruimte voor verbetering. De functionaris gegevensbescherming vervult een cruciale rol bij het waarborgen van de AVG binnen de organisatie. De door ons gesproken functionarissen zien bijvoorbeeld graag meer expertise over de juridische eisen bij de ontwikkelteams. Verder zagen we dat de positionering van de functionarissen gegevensbescherming verschilt per organisaties. Dat geldt ook voor hoe actief ze betrokken zijn bij de toetsing van profilerende systemen. Tenslotte blijft het zaak om niet alleen per systeem te toetsen, maar oog te houden voor de impact van de gehele verzameling systemen op de levens en rechten van burgers.

Organisaties kunnen van elkaar leren. Over het geheel genomen treffen de organisaties waarmee we gesproken hebben, vergelijkbare maatregelen. Tegelijkertijd merken we verschillen op. De SVB besteedt bijvoorbeeld via burgerpanels en een algoritmeregister meer aandacht aan begrijpelijke en toegankelijke informatie. Ook de ontwikkelde normenkaders verschillen per organisatie in de mate van detail. Het CJIB kijkt vanwege de ontwikkeling van profilerende systemen opnieuw naar de klachtenprocedure. Het UVW en de SVB doen dat niet. Het UWV en het CJIB denken na over een ethische commissie, waarbij er verschillende ideeën zijn over de rol en invulling van zo'n commissie.

5.3 Algoritmes leren afwegen

Hoewel de onderzochte uitvoeringsorganisaties risico's doorgaans niet beschrijven in termen van mensenrechtelijke verdragen, blijkt uit ons onderzoek dat bij de ontwikkeling van profilerende systemen wel degelijk oog is voor mensenrechtelijke aspecten, zoals privacy en non-discriminatie. Deze aandacht is begrijpelijk in de huidige maatschappelijke omstandigheden. Uit de kinderopvangtoeslagenaffaire, maar ook uit andere misstanden, is gebleken dat uitvoeringsorganisaties bij de verzameling, analyse en verwerking van gegevens in de fout kunnen gaan en de rechten van mensen kunnen schenden. Tegelijkertijd voelen ze zich ook genooddaakt om profilerende systemen te gebruiken, vanwege de maatschappelijke en politieke wens om enerzijds effectief fraude op te sporen en anderzijds te zorgen voor soepele dienstverlening.

De onderzochte organisaties zien dat er risico's bij de inzet van profilerende systemen kunnen ontstaan, en nemen maatregelen om deze risico's zoveel mogelijk te verkleinen. Daarbij zijn ze op zoek naar houvast. Ze vragen zich steeds af hoe ze hun verantwoordelijkheid het beste kunnen nemen: welke maatregelen zorgen ervoor dat ze zowel in wettelijk als in ethisch opzicht voldoen aan de gerechtvaardigde verwachtingen van de politiek en de samenleving?

De genomen maatregelen dienen voor meer duidelijkheid te zorgen. Desondanks blijven de onderzochte organisaties zich afvragen of ze daarmee voldoende beantwoorden aan wat de politiek en de samenleving van hen verlangen. Daarbij is ethiek een wezenlijk onderdeel geworden van hoe binnen de onderzochte organisaties wordt gesproken en gewerkt. Steeds meer sijpelt het besef door dat ze niet alleen moeten handelen binnen en volgens de wet, maar ook wettelijk toegestane toepassingen moeten vermijden die maatschappelijk onwenselijk zijn.

Zoals gezegd dienen de genomen maatregelen twee doelen: het vertalen van wettelijke vereisten naar een hanteerbare praktijk en het stimuleren van ethische expertise binnen de organisatie. Op beide vlakken is veel gebeurd. Tegelijkertijd zijn de maatregelen nog in ontwikkeling en kunnen ze verder versterkt worden. De eerste twee aanbevelingen zijn daarop gericht: het verbeteren van de hanteerbaarheid van het gebruikte normenkader en het versterken van expertise in de organisatie. De vier aanbevelingen daarna richten zich op het versterken van specifieke mensenrechten: privacy, biastoetsing, informatierechten en procedurele rechten.

Uiteindelijk is het realiseren van een verantwoorde algoritmische uitvoeringspraktijk een gedeelde taak van uitvoeringsorganisaties, beleid en politiek. De leidende vragen daarbij moeten zijn hoe maatschappelijke doelen het beste kunnen worden bereikt en hoe de rechten van burgers het beste worden beschermd. De laatste aanbeveling richt zich hierop.

De geformuleerde aanbevelingen vloeien voort uit de gesprekken over de genomen maatregelen ten aanzien van profilerende algoritmes bij het UWV, het CJIB en de SVB, maar dienen als uitgangspunt voor elke uitvoeringsorganisatie die profilerende systemen ontwikkelt of inzet. Sommige aanbevelingen zijn niet geadresseerd aan uitvoeringsorganisaties, maar aan andere delen van de overheid zoals de wetgever en verschillende ministeries.

1. Verbeter de hanteerbaarheid van de juridische vereisten en ethische beginselen.
2. Stimuleer ethische expertise in de gehele organisatie.
3. Blijf investeren in privacytoetsing.
4. Versterk de manier waarop bias-toetsing plaatsvindt.
5. Investeer in een actieve informatievoorziening.
6. Versterk de waarborgen voor procedurele rechten.
7. Waak voor verkoking bij de ontwikkeling en beoordeling van algoritmische toepassingen.

1. Verbeter de hanteerbaarheid van bestaande juridische kaders en ethische beginselen.

Het ontwikkelen van profilerende systemen is niet alleen technisch complex, maar levert ook complexe juridische en ethische vraagstukken op. Omdat de technologie in ontwikkeling is, en veel afhangt van de specifieke toepassing, zijn deze vraagstukken niet vooraf volledig te vangen in regulering. Juridische eisen zijn – om goede redenen – vaak in algemene termen omschreven. Dat vraagt in de praktijk om een vertaalslag, bijvoorbeeld in de vorm van impact assessments, de ontwikkelde normenkaders of handreikingen die juridische eisen verhelderen. Het ministerie van BZK heeft om die reden onder meer het IAMA, CODIO en de handreiking non-discriminatie laten ontwikkelen. Het College van de Rechten van de Mensen ontwikkelde het toetsingskader voor risicoprofielen (CvdRM 2021). Ook internationaal zijn er diverse richtsnoeren verschenen met hetzelfde doel. Tegelijkertijd is duidelijk dat hulpmiddelen niet mogen worden tot checklists waarbij de consequenties van een systeem voor de rechten van burgers niet daadwerkelijk meer worden doordacht.

Uit de gesprekken concluderen we dat men onduidelijkheid ervaart over hoe bestaande juridische eisen en ethische beginselen, hanteerbaar in de eigen organisaties kunnen worden toegepast. De door BZK ontwikkelde hulpmiddelen hebben deze onduidelijkheid niet geheel weggenomen.

De onderzochte organisaties overbruggen deze afstand nu, tenminste ten dele, door het zelf ontwikkelen van impact assessments, normenkaders, ethische kompassen, handreikingen en biastoetsen, naast het toepassen van de AVG en de bijbehorende DPIA. Deze zelf ontwikkelde instrumenten maken een vertaalslag van bestaande juridische eisen en ethische beginselen naar gedetailleerde vragen over het systeem. Een normenkader legt zo vast wie binnen de organisatie verantwoordelijk is voor het beantwoorden van welke vragen en verbindt functies, werkprocessen, hulpmiddelen en benodigde expertise expliciet met elkaar. Hoe specifieker de organisatie die heeft vastgelegd, hoe duidelijker voor alle betrokkenen wat van hen wordt verwacht.

Actie 1.1 Het is zaak dat iedere uitvoeringsorganisatie een normenkader opstelt, waarin verantwoordelijkheden van functionarissen, en mensenrechtelijke aspecten zijn belegd.

Verder is het belangrijk na te gaan of alle relevante ethische beginselen en juridische vereisten in de ontwikkelde normenkaders zijn meegenomen. Zo noemden we al dat de onderzochte organisaties over deze vraagstukken spreken in termen van ethiek en juridische vereisten – en minder in termen van

mensenrechten. Dat kan betekenen dat bepaalde mensenrechten en codificeringen daarvan, of relevante overwegingen, niet even expliciet worden meegenomen. Tijdens de externe consultaties werd de vraag gesteld of bijvoorbeeld de algemene beginselen van behoorlijk bestuur explicieter moeten worden meegenomen, zoals in het IAMA al gebeurt. Ook de *Code goed digitaal openbaar bestuur* (CODIO) neemt verschillende relevante aspecten mee voor uitvoeringsorganisaties. Zijn de kaders compleet? Zijn er nadere, specifieke vertalingen nodig van vragen in bijvoorbeeld IAMA, CODIO, het toetsingskader voor risicoprofielen van het College voor de Rechten van de Mens?

Actie 1.2 Controleer als uitvoeringsorganisatie of het gebruikte normenkader volledig is. Maak daarbij gebruik van algemene hulpmiddelen zoals IAMA, CODIO, het toetsingskader voor risicoprofielen van het College voor de Rechten van de Mens en de handreiking non-discriminatie. Voer deze controle periodiek uit.

Ook is het belangrijk dat uitvoeringsorganisaties met elkaar in gesprek gaan over het gewenste normenkader. Het is namelijk nu onduidelijk in hoeverre standaardisatie van de gebruikte normenkaders wenselijk is, of dat daarmee de toepasbaarheid voor de eigen organisatie verloren gaat. Bovendien kunnen uitvoeringsorganisaties op deze manier van elkaar leren. Tot slot kan uit het overleg blijken dat een nadere codificering van bepaalde normen gewenst is.

De aankomende AI-verordening, zoals voorgesteld door de Europese Commissie, zal voor hoog-risico-systemen bepaalde maatregelen die de onderzochte uitvoeringsorganisaties nemen, verplicht stellen, zoals impact assessments, systemen voor kwaliteits- en risicobeheer en het geven van begrijpelijke informatie over het gebruikte systeem. De verordening zegt echter weinig over hoe dat precies dient te gebeuren. Juist daar zijn bovenstaande acties op gericht.

Actie 1.3 Stel als uitvoeringsorganisaties een overleg in over het ontwikkelde normenkader om te komen tot *best practices*. Het ministerie van BZK kan hierin een coördinerende en stimulerende rol spelen.

2. Stimuleer ethische expertise in de gehele organisatie

Hoe belangrijk de discussie over normenkaders ook is, de kaders vallen of staan bij een doortastende toepassing van normen in de praktijk. Daartoe is naar aanleiding van de AVG een raamwerk opgericht met *privacy officers* en functionarissen gegevensbescherming, die ontwikkelteams begeleiden en toetsen

bij het beschermen van privacy. Ons onderzoek wijst uit dat de functionarissen gegevensbescherming nog regelmatig moeten ingrijpen om nieuwe toepassingen in goede banen te leiden (zie aanbeveling 3), ook bij profilerende systemen. Dat geeft aan dat er nog meer geïnvesteerd moet worden in de juridische en ethische vaardigheden van de ontwikkelteams. Het is zaak dat ontwikkelteams beter om kunnen gaan met de normatieve complexiteit van algoritmische systemen, zodat ze ingewikkelde discussies vruchtbaar kunnen voeren en afwegingen kunnen maken. De onderzochte uitvoeringsorganisaties investeren in deze expertise, door trainingssessies te organiseren en het behalen van een certificaat aan te moedigen. Het is belangrijk deze investeringen te blijven doen en waar mogelijk uit te breiden, bijvoorbeeld door uitvoerige juridische en ethische opleidingen verplicht te stellen voor nieuwe medewerkers. De rijksoverheid kan hierbij een ondersteunende rol spelen. Verder zou het een pré moeten zijn bij sollicitaties als een *data scientist* naast technische vaardigheden ook beschikt over juridische en ethische competenties.

Actie 2.1 Zorg als uitvoeringsorganisatie dat ontwikkelteams beschikken over de relevante juridische en ethische vaardigheden. Investeer in interne opleidingen en train alle medewerkers in de organisatie.

Net als *privacy officers* die vroeg betrokken zijn bij de ontwikkeling van een systeem om advies te geven, is het zinvol om een ethisch adviseur met een vergelijkbare taak aan te stellen. De adviseur kan in samenspraak met de *privacy officer* en de functionaris gegevensbescherming de ontwikkelteams adviseren over een breed aspect ethische en mensenrechtelijke risico's.

Actie 2.2 Stel een ethisch adviseur in om de interne expertise-ontwikkeling te versterken en een doordacht gebruik van normenkaders te ondersteunen.

Verder blijkt dat het oprichten van ethische commissies door de onderzochte uitvoeringsorganisaties wordt gezien als een belangrijke versterking van het ethische karakter van de organisatie. Het idee dat een ontwikkelteam ook langs een ethische commissie moet, kan de behandeling en beantwoording van de normenkaders verdiepen. Ook de betrokkenheid van externe experts kan de blik van de organisatie verruimen. Tegelijkertijd bestaan er zorgen over de rol en invulling die zo'n commissie zou moeten krijgen, bijvoorbeeld meer agenderend of juist toezichthoudend. Dat leidt tot verschillende vragen. Bij een toezichthoudende rol is de vraag hoe de commissie zich verhoudt tot de taken van de functionaris gegevensbescherming en of ze voldoende capaciteit heeft om tijdig en volwaardig te toetsen. Ook zijn er vragen over of het voldoende duidelijk is op welke normen

de commissie toetst. Als het gaat om een meer agenderende invulling, is het onwenselijk als het instellen van een commissie afleidt van de al aanwezige expertise en de al belegde verantwoordelijkheden. Het is zaak dat de commissie met gezag advies kan geven. Er is geen kant-en-klaar recept voor een succesvolle ethische commissie. Allerlei partijen, waaronder gemeentes, zoeken naar een goede invulling.

Actie 2.3 Zorg voor overleg tussen uitvoeringsorganisaties over hoe zij hun ethische commissie willen invullen en waar geleerd wordt van elkaars ervaringen.

Tot slot is het van belang dat uitvoeringsorganisaties oog blijven houden voor degene op wie het systeem uiteindelijk een effect heeft: burgers. Door hen structureel te betrekken en te bevragen bij de ontwikkeling van systemen, ontstaat een meer voldragen risicobeoordeling dan wanneer de blik uitsluitend intern gericht is. Naast overleggen met cliëntenraden, de Ombudsman en maatschappelijke organisaties, verdient het instellen van een burgerpanel aanbeveling.

Actie 2.4 Stel als uitvoeringsorganisatie een burgerpanel in waarin de gevolgen van algoritmische systemen worden bekeken.

3. Blijf investeren in privacytoetsing

Het beschermen van privacy kent al een lange traditie. De genomen maatregelen zijn op dit punt dan ook het meest volwassen. Toch is er ruimte voor verbetering. We noemden hierboven al het belang van aanwezige expertise over mensenrechtelijke risico's bij ontwikkelteams en andere betrokken medewerkers. Uit het onderzoek bleek dat een functionaris gegevensbescherming een bepaalde opzet van een profilerend systeem heeft tegengehouden. De functionarissen gegevensbescherming vormen – zoals de wet bedoeld heeft – dus een cruciale schakel om privacyrechten te waarborgen. We zijn verschillen tegengekomen bij de onderzochte organisaties in hoe de functionarissen gegevensbescherming zijn gepositioneerd en de mate waarin zij betrokken zijn bij de toetsing van profilerende systemen.

Actie 3.1 Zorg als uitvoeringsorganisaties dat de functionaris gegevensbescherming zo gepositioneerd is dat hij waar nodig doortastend toezicht kan houden op profilerende systemen.

Daarnaast kwam in sommige gesprekken met de onderzochte uitvoeringsorganisaties naar voren dat het privacybeleid op punten verder aangescherpt zou kunnen worden. Bijvoorbeeld als het gaat om de omgang met bewaartermijnen of het opschonen van data en het weghalen van niet-kloppende gegevens. Er kan dus nog meer gedaan worden om vastgestelde privacykaders in de praktijk te realiseren.

Actie 3.2 Blijf het gevoerde privacybeleid kritisch evalueren en waar nodig aanscherpen.

4. Versterk de manier waarop bias-toetsing plaatsvindt

Het toetsen van bias in lerende algoritmes vormt bij de onderzochte organisaties een standaardonderdeel van de ontwikkeling en de gehele levenscyclus van algoritmische systemen. Het geeft aan hoe serieus men het risico op discriminatie acht. Maar we merkten op dat bij het proces van biascontrole de onderzochte uitvoeringsorganisaties een eigen partij inhuren, en deze rapportages niet publiekelijk beschikbaar zijn. Daardoor is niet transparant voor de samenleving en politiek hoe op bias wordt getoetst.

Actie 4.1 Geef als uitvoeringsorganisatie meer inzicht over hoe biastoetsing plaatsvindt.

Daarnaast gaven de onderzochte uitvoeringsorganisaties aan dat zij ieder hun eigen methoden en technieken hanteren om het proces van biastoetsing vorm te geven. Daarbij richten ze een eigen uitwisseling op met wetenschappers om de gebruikte methode aan te laten sluiten bij de laatste wetenschappelijke inzichten. Tegelijkertijd zijn er algemene hulpmiddelen voorhanden, zoals de handreiking non-discriminatie by design, waarin noodzakelijke processtappen en voorbeeldtechnieken worden omschreven. Er speelt hier dus een vergelijkbare kwestie als bij aanbeveling 1 over hanteerbare normenkaders: welke mate van standaardisatie, en toepasbaarheid voor de organisatie, is wenselijk? Is het nodig om de handreiking een dwingender karakter mee te geven? Ook op het punt van databeheer geeft de voorgestelde AI-verordening eisen mee: biastoetsing dient plaats te vinden. Daarmee is nog weinig gezegd over het hoe. Het is dus zaak om de expertise op het gebied van biastoetsing bij elkaar te brengen, zodat alle uitvoeringsorganisaties daarvan de vruchten kunnen plukken.

Actie 4.2 Zet een nationaal kennisplatform op, waar de relevante expertise kan worden ontwikkeld en gedeeld. Bepaal welke mate van standaardisatie gewenst en of nadere wettelijke eisen nodig zijn.

Ten slotte benadrukken we dat het van belang is en blijft dat toezichthouders op actieve wijze op bias controleren.

5. Investeer in een actieve informatievoorziening

Het is om verschillende redenen van belang dat uitvoeringsorganisaties zorgen voor een goede informatievoorziening. Gekozen volksvertegenwoordigers moeten op de hoogte zijn van de praktijk van uitvoeringsorganisaties om hun controlerende taak uit te kunnen voeren en beleid te sturen. De journalistiek moet met de juiste informatie het publieke debat kunnen voeden. De burger heeft op basis van de AVG een recht om te weten hoe zijn of haar persoonsgegevens door uitvoeringsorganisaties worden verwerkt. Het is daarom zaak dat uitvoeringsorganisaties doen wat ze kunnen om deze informatievoorziening tot stand te brengen. Over sommige aspecten van een algoritme, zoals het doel en de mogelijke impact op burgers, kan vrijwel altijd informatie gegeven worden. Juist vanwege het na verschillende misstanden geschonden vertrouwen van burgers verdient dit extra aandacht.

Daarnaast lag bij sommige gesprekken de nadruk op passieve informatievoorziening, de informatie die een burger krijgt als hij daarom vraagt. Dit is belangrijk, maar om draagvlak en vertrouwen te creëren, is het zaak burgers actief en toegankelijk te informeren over nieuwe algoritmische systemen. Algoritmeregisters kunnen daar een rol in spelen. In 2021 werd al een motie van het lid Dassen c.s. in de Tweede Kamer aangenomen, die de regering verzocht om overheden te verplichten een algoritmeregister te gebruiken.⁵ Deze ontwikkeling juichen we van harte toe. Het is zaak dat het bouwen van een dergelijk register niet te lang op zich laat wachten.

Actie 5.1 Versnel de ontwikkeling en het gebruik van een algoritmeregister. Zowel uitvoeringsorganisaties als de rijksoverheid dragen hiervoor verantwoordelijkheid.

Ook is het van belang de informatievoorziening af te stemmen met een representatieve groep burgers, die kan inschatten of de informatie nuttig en begrijpelijk is en aansluit bij de vragen die bij burgers leven.

⁵Kamerstuk 35 925 VII, 2021.

Actie 5.2 Ontwikkel de informatievoorziening in samenspraak met representatieve burgerpanels.

Tot slot merken we op dat het woord algoritme soms vermeden werd, vanwege de zorgen die dat in politiek en maatschappij oproept. Ook zien we dat in het maatschappelijke en politieke debat onduidelijkheid bestaat over wat algoritmes zijn en over welke kenmerken leiden tot mensenrechtelijke vraagstukken. Deze onduidelijkheid kan gevolgen hebben voor het publieke debat over en de controle op algoritmes.

Actie 5.3 Ontwikkel als overheid helder en consistent taalgebruik over algoritmes. Het ministerie van Binnenlandse Zaken en Koninkrijksrelaties kan hier een voortrekkersrol spelen.**6. Versterk de waarborgen voor procedurele rechten**

We zagen in het onderzoek dat uitvoeringsorganisaties verschillend denken over de vraag of het vanwege de invoering van profilering nodig is om opnieuw naar de klachtenprocedures te kijken. Voor ieder algoritmisch systeem is het van belang dat er voldoende mogelijkheden zijn om fouten in het proces te herstellen en dat burgers op een toegankelijke manier klachten kunnen indienen.

Actie 6.1 Zorg als uitvoeringsorganisatie ervoor dat de klachtenprocedure bij een nieuw profilerend systeem opnieuw wordt geanalyseerd en zo nodig wordt verbeterd.

We concluderen dat er onduidelijkheid bestaat over hoe het begrip ‘menselijke tussenkomst’ kan worden ingevuld, en wanneer een systeem voldoende uitlegbaar is. Bij het gebruik van profilerende systemen is vaak geen sprake van geautomatiseerde besluitvorming: een menselijke handhaver kijkt op basis van het lerende algoritme of er bijvoorbeeld een onderzoek moet worden gestart. De menselijke handhaver kan op basis van het eigen onderzoek uitleggen waarom er sprake is van fraude, en ook hoe tot dat oordeel is gekomen. Maar deze motivering geeft nog geen uitleg over de werking van het algoritme – en waarom een burger geselecteerd is.

Actie 6.2 Verhelder als wetgever de eisen aan de uitlegbaarheid van beleidsondersteunende lerende algoritmes.

De uitvoeringsorganisaties trachten ook op andere manieren uitlegbaarheid te waarborgen. Omdat ze *unsupervised learning* niet verenigbaar achten met dit

principe, kiezen ze er bewust voor om deze technologie niet te gebruiken, zelfs niet als dat betere resultaten zou opleveren. Tegelijkertijd is duidelijk dat de ontwikkelaars wel experimenteren met deze technieken en ze spoedig zouden kunnen inzetten als de opvattingen over uitlegbaarheid zouden veranderen. Dit geeft aanleiding om te kijken naar de verdere concretisering van juridische vereisten.

Actie 6.3 Overweeg als wetgever om nadere wettelijke eisen aan het toepassen van lerende systemen te stellen. Kijk in het bijzonder naar de toelaatbaarheid van *unsupervised learning*.

7. Waak voor verkokering bij de ontwikkeling en beoordeling van algoritmische toepassingen

Uit dit onderzoek blijkt dat de onderzochte uitvoeringsorganisaties van mening zijn met profilering aan de slag te *moeten* gaan, om tegemoet te komen aan de eisen van de politiek en de maatschappij. De moderne uitvoeringsorganisatie moet immers effectief fraude bestrijden, maar ook snel en nauwkeurig diensten verlenen. En dat alles mag niet te veel geld kosten. In deze omstandigheden kan het moeilijk zijn om de proportionaliteitsvraag in alle aspecten te stellen. Uit de interviews en externe consultaties bleek dat de onderzochte uitvoeringsorganisaties bij het beoordelen van proportionaliteit niet altijd de kosten van de mitigatiemaatregelen meenemen, zoals het aannemen van *data-scientists*, ethische adviseurs en ethische commissies. Ook bleek uit de gesprekken en consultaties niet dat het scenario met profilering altijd werd afgewogen tegen een scenario waarin een alternatieve oplossing gezocht wordt, die weinig tot geen mensenrechtelijke risico's kent (subsidiariteit). Toch is het essentieel deze aspecten wel mee te nemen. Het draait erom hoe maatschappelijke doelen het beste kunnen worden bereikt, waarbij technologie niet automatisch de beste oplossing hoeft te zijn. Daarbij is het van belang om ook te kijken naar de gezamenlijke impact van profilerende systemen op de rechten van burgers.

Actie 7.1 Zorg dat binnen en buiten uitvoeringsorganisaties de *nuloptie* op tafel ligt bij het toetsen van de proportionaliteit van algoritmische profilering.

Bij het voeren van deze discussie is meer inzicht nodig in de behaalde effectiviteit van de profilerende systemen. De in dit onderzoek genoemde cijfers lopen uiteen, en waren moeilijk met elkaar te vergelijken. Een helder beeld van de effectiviteitsscores is nodig om de proportionaliteit te beoordelen. Het is zaak dat de uitvoeringsorganisaties overleggen hoe deze helderheid geboden kan worden.

Actie 7.2 Zorg dat de effectiviteit van de profilerende algoritmes op een vergelijkbare manier wordt geëvalueerd en centraal bekend is.

Voorts merken we op dat de proportionaliteit per algoritmische toepassing anders kan uitvallen. Zo zou met name de proportionaliteitsbeoordeling van profilerende algoritmes die burgers ondersteunen, en minder kans hebben om de rechten van burgers te schaden, positief uit kunnen vallen. Tegelijkertijd kan het lastig zijn voor uitvoeringsorganisaties om dit financieel rond te krijgen. Juist als een toepassing burgers ondersteuning biedt, of kan helpen hun rechten te laten gelden, is het van belang de mogelijkheden serieus te nemen. Ook als een toepassing niet direct in het belang is van de uitvoeringsorganisatie zelf of als dit geld kost. Algoritmische toepassingen die van waarde zouden kunnen zijn voor de burger, mogen niet het slachtoffer worden van verkokerde uitvoeringsorganisaties.

Actie 7.3 Verken als uitvoeringsorganisatie profilerende toepassingen die de burger meer ondersteuning bieden. Het ministerie van Binnenlandse Zaken en Koninkrijksrelaties kan deze ontwikkeling stimuleren.

Tot slot

Uit dit onderzoek blijkt dat de onderzochte uitvoeringsorganisaties veel aandacht besteden aan het mitigeren van de risico's van profilerende algoritmes. Ook in maatschappij en politiek zien we veel aandacht uitgaan naar de kansen en risico's van lerende systemen. Maar de aandacht voor profilerende algoritmes mag niet leiden tot een zwakkere focus op *rule-based* systemen, die nog steeds het merendeel van de uitvoering uitmaken. Deze systemen hebben namelijk dikwijls ook een grote impact op de rechten van burgers omdat hun besluiten vaak geautomatiseerd worden doorgevoerd, doorgaans direct financiële consequenties hebben en fouten soms lastig te corrigeren blijken.

Opvallend is dat bepaalde mensenrechten, zoals het geven van informatie, juist bij *rule based* systemen lastiger lijken te realiseren. De lerende systemen zijn pas recent geïntroduceerd en de ontwikkeling ervan, gemaakte keuzes en de werking lijken beter gedocumenteerd dan bij veel oudere *rule-based* systemen. Bovendien zien we bij de organisaties die we gesproken hebben dat ze bij lerende systemen er bewust voor kiezen om deze *zelf* te ontwikkelen. De organisaties hebben daar dus meer grip op. *Rule based* systemen worden vaak als laag risico ingeschat en vaker ingekocht.

Het is daarom zaak om de lessen die geleerd zijn bij de ontwikkeling van profilerende algoritmes en de aanbevelingen van dit onderzoek, ook toe te passen op de *rule-based* systemen. Uiteindelijk is het streven immers dat uitvoeringsorganisaties bij alle algoritmische toepassingen mensenrechten afdoende beschermen.

Literatuurlijst

Algemene Rekenkamer (2021). *Aandacht voor algoritmes*. Den Haag: Algemene Rekenkamer

Baer, T. (2019) *Understand, Manage and Prevent Algorithmic Bias: A Guide for Business Users and Data Scientists*. New York: Apress

Barocas, S. en A. Selbst (2016). "Big Data's Disparate Impact." *Californian Law Review* 104, 671.

CAHAI Secretariat (2020). *Towards Regulation of AI Systems*. Council of Europe

College voor de Rechten van de Mens (2021) *Discriminatie door risicoprofielen: een mensenrechtelijk toetsingskader*. 's-Gravenzande: Van Deventer

Considerati (2020). *Onvoorziene effecten van zelflerende algoritmen*. Amsterdam: Considerati

Gerards, J. T. Schafer, A. Vankan & I. Muis (2021) *Impact Assessment Mensenrechten en Algoritmes*. Den Haag: Ministerie van Binnenlandse Zaken en Koninkrijksrelaties

Grimmelikhuijsen, S. & Meijer, A. (2020) Verantwoorde algoritmisering: zorgen waardengevoeligheid en transparantie voor meer vertrouwen in algoritmische besluitvorming? *Bestuurskunde* 30(4), 7-20

Hofman, E. 'Wie stopt de datahonger?' In: *De Groene Amsterdammer* 9 juni 2021. <https://www.groene.nl/artikel/wie-stopt-de-datahonger>

Hooghiemstra & Partners (2021) *Hoe gemeenten besluiten over algoritmen & mensenrechten*. Den Haag: Hooghiemstra & Partners

Kamerstukken 2019-2020, 26643, nr. 641

Kamerstukken II 2019-2020. 26643, nr. 642

Kamerstukken II 2020-2021. 26643, nr. 765

Kulk, S. en van Deursen, S. (2020) *Juridische aspecten van algoritmen die besluiten nemen. Een verkennend onderzoek.* Universiteit Utrecht.

Meijer, A. & Ruijter, E. (2021) *Code Goed Digitaal Openbaar Bestuur.* Utrecht: Universiteit Utrecht

Nationale ombudsman (2012) *Mijn onbegrijpelijke overheid.* Den Haag: Nationale ombudsman

Nationale ombudsman (2016) *Wie niet past, loopt vast.* Den Haag: Nationale ombudsman

PwC (2021) *Onderzoek effecten FSV Toeslagen.* Amsterdam: PwC

Raad van State (2017) *Uitspraak 201600614/3/R2, 201600617/3/R2, 201600618/3/R2, 201600620/3/R2, 201600622/4/R2, 201600630/3/R2*
<https://www.raadvanstate.nl/@115602/201600614-3-r2/>

Raad van State (2021) *Digitalisering: Wetgeving en bestuursrechtspraak.* Den Haag: Raad van State.

Raad voor het Openbaar Bestuur (2021). *Sturen of gestuurd worden? Over de legitimiteit van sturen met data.* Den Haag: Raad van het Openbaar Bestuur

Rathenau Instituut (2010) *Databases: Over ICT-beloftes, informatiehonger en digitale autonomie.* Den Haag: Rathenau Instituut

Rathenau Instituut (2012) *Voorgeprogrammeerd: Hoe internet ons leven leidt.* Den Haag: Boom Lemma uitgevers

Rathenau Instituut (2017) *Human rights in the robot age. Challenges arising from the use of robotics, artificial intelligence, and virtual and augmented reality.* Den Haag: Rathenau Instituut

Rathenau Instituut (2017) *Opwaarderen. Borgen van publieke warden in de digitale samenleving.* Den Haag: Rathenau Instituut

TNO (2019) *Quick scan AI in de publieke dienstverlening.*
<https://www.rijksoverheid.nl/documenten/rapporten/2019/04/08/quick-scan-in-de-publieke-dienstverlening>

TNO (2021) *Quick scan AI in de publieke diensverlening II* Den Haag: TNO

Tweede Kamer der Staten Generaal (2021a). *Klem tussen balie en beleid*. Rapport van de Tijdelijke Commissie Uitvoeringsorganisaties (TCU).

Tweede Kamer der Staten-Generaal (2021b) *Ongekend onrecht*. Rapport van de Parlementaire ondervragingscommissie

Van Eck, M., M. Bovens en S. Zouridis (2018). "Algoritmische rechtstoepassing in de democratische rechtstaat." *Nederlands juristenblad* 93(40), 3008-3017.

Vetzo, M.J., J. Gerards en R. Nehmelman (2018). *Algoritmes en grondrechten*. Den Haag: Boom Juridisch.

Widlak, A. & R. Peeters (2018). *De digitale kooi. Over (on)behoorlijk bestuur door informatiearchitectuur*. Den Haag: Boom Bestuurskunde.

Widlak, A. (2021) *Volwassen digitale overheid*. Den Haag: Boom Bestuurskunde

Wolswinkel, J. (2021). Algoritmische besluitvorming door de overheid: naar een (betere) wettelijke verankering? Notitie Eerste Kamer ten behoeve van deskundigenbijeenkomst.

WRR (2011). *iOverheid*. Den Haag: WRR.

WRR (2016). *Big data in een vrije en veilige samenleving*. Den Haag: WRR.

Wetten en handreikingen

Algemene beginselen behoorlijk bestuur: <https://www.infomil.nl/vaste-onderdelen/onderwerpen/lucht-water/handboek-water/handreiking-lozingen/procedure-keuze-rechtsbescherming/algemene-beginselen-behoorlijk-bestuur/>

Algemene verordening gegevensbescherming:
https://autoriteitpersoonsgegevens.nl/sites/default/files/atoms/files/verordening_2016_-_679_definitief.pdf

Algemene wet bestuursrecht:
<https://wetten.overheid.nl/BWBR0005537/2022-01-28>

Europees Verdrag voor de Rechten van de Mens:

https://www.echr.coe.int/documents/convention_nld.pdf

Handreiking non discriminatie BZK 2021:

Van der Sloot, B., E. Keymolen, M. Noorman, het College voor de Rechten van de Mens, H. Weerts, Y. Wagenveld en B. Visser.

<https://open.overheid.nl/repository/ronl-3f9fa69c-acf4-444d-96e1-5c48df00eb3c/1/pdf/bijlage-1-handreiking-non-discriminatie-by-design.pdf>

Wet open overheid:

<https://zoek.officielebekendmakingen.nl/stb-2021-499.html>

Bijlage 1: interviews

Medio augustus tot medio september 2021 heeft het onderzoeksteam interviews afgenomen met medewerkers van het UWV, de SVB en het CJIB. Deze interviews vormen de bron voor hoofdstuk 3. De respondenten bestaan uit medewerkers van verschillende lagen uit de organisaties: data-analisten, projectmanagers, *privacy officers*, functionarissen gegevensbescherming, beleidsadviseurs en directieleden.

Respondenten	Rol en organisatie
Eveline Jacobson	Projectmanager betrokken bij Commissie Data Ethiek UWV
Gerjan van Schaik	Functionaris gegevensbescherming UWV
Henk de Ruiter	Hoofd Risicomanagement en Intelligence bij Directie Handhaving UWV
Maarten Camps	Voorzitter raad van bestuur UWV
Robin Philippo	Data-analist UWV
Alex Corra	Senior beleidsadviseur Recht en Ethiek SVB
Berend van der Molen	Bestuursadviseur Strategie en Externe Betrekkingen SVB
Hatice Dogan	Functionaris gegevensbescherming SVB
Jeffrey Apon	Sectiemanager Intelligence en Risicogericht onderzoek SVB
Rolf Leijdekker	

	Kwartiermaker data-autoriteit CIO-office SVB
Anja Tol	<i>Privacy officer</i> CJIB
Henkjan Derks	Plaatsvervangend algemeen directeur en directeur IT en Bedrijfsvoering CJIB
Joke de Boer	<i>Lead datascientist</i> bij Innovatielab058 CJIB

Bijlage 2: interviewprotocol

Voor het onderzoek zijn semigestructureerde interviews gevoerd aan de hand van negen vragen verdeeld over drie thema's.

De stand van zaken van het gebruik van profilerende algoritmes

1. Voor welke vraagstukken gebruikt of ontwikkelt de uitvoeringsorganisatie momenteel profilerende algoritmes?

Wat zijn de doelen van deze systemen? Welke taken worden overgenomen of ondersteund door het gebruik van algoritmes? Wat leveren de systemen concreet op?

2. Hoe zijn die algoritmische systemen tot stand gekomen?

Heeft de organisatie systemen zelf ontwikkeld of is de dienst/het product ingekocht? Wie beslist of gebruik gemaakt gaat worden van algoritmes? Wie wordt er betrokken bij de besluitvorming? Vond in dit proces een impact assessment plaats/ werden impact assessment tools gebruikt?

3. Om wat voor type systemen en welke soorten data gaat het?

Van welke bronnen maakt het algoritme gebruik? Welke mate van menselijke tussenkomst is aanwezig? Welk type AI wordt gebruikt? Wat zijn de technische uitdagingen? Hoe goed zijn de systemen?

4. Vooruitblik: wat kunnen de algoritmische systemen – en risicoprofilering in het bijzonder – in de toekomst betekenen voor deze uitvoeringsorganisatie?

Waar kunnen ze de grootste impact hebben en wat heeft het meeste belofte?

Mensenrechtenbescherming

5. Kan het gebruik van algoritmes in uw organisatie leiden tot mensenrechtelijke vraagstukken?

Voorbeelden: gelijkheidsrechten zoals non-discriminatie, privacyrechten omtrent datagebruik, -verzameling en -opslag, procedurele rechten zoals recht op informatie, recht op herstelprocedure bij fouten. Welke risico's heeft de organisatie op de radar?

6. Op welke manier vindt binnen de organisatie het gesprek plaats over het wegen van risico's?

Welke methodes zijn gehanteerd om risico's te identificeren? Hoe worden verschillende aspecten geborgd en bij wie in de organisatiestructuur ligt de verantwoordelijkheid? Is er een duidelijk aanspreekpunt; worden er hulpmiddelen ingezet?

7. Welke maatregelen neemt de uitvoeringsorganisatie om risico's te vermijden en potentiële negatieve gevolgen tegen te gaan?

In welke kennis en ervaring is en wordt er geïnvesteerd om risico's in te schatten en te

mitigeren? Zijn er speciale richtlijnen voor werknemers voor het ontwerpen/inkopen of gebruiken van algoritmes? Wat kan een burger doen als die geraakt is door een handeling van de uitvoeringsorganisatie op basis van een algoritmische toepassing?

Mogelijke knelpunten

8. Ervaart de uitvoeringsorganisatie bij de bescherming van mensenrechten knelpunten, en zo ja, welke zijn dat?

Wellicht capaciteit om risico's te mitigeren, te weinig kennis of budget, technologische problemen, organisatiecultuur. Is er voldoende ruimte om intern problemen aan te kaarten, is er voldoende uitwisseling tussen expertises (IT-juridisch).

9. Welke mogelijke oplossingen en maatregelen ziet u om praktijken van uitvoeringsorganisaties te versterken?

Op het terrein van regelgeving, financieel, technologisch of organisatorisch? In hoeverre helpen de impact assessments, handreikingen en richtlijnen vanuit BZK, EZK, JenV en vele andere (maatschappelijke) organisaties?

Afsluiting

- Zijn er nog andere personen waarvan u denkt dat het relevant is om ze te spreken voor dit onderzoek?
- Zijn er nog zaken onbesproken gebleven waarvan u denkt dat ze belangrijk zijn voor dit onderzoek?

Bijlage 3: externe consultaties

Op 18 en 22 november 2021 organiseerde het onderzoeksteam een externe consultatie waarin de voorlopige bevindingen vanuit de interviews aan experts uit verschillende vakgebieden werden voorgelegd. De bevindingen zijn gepresenteerd waarna de experts hun feedback zowel schriftelijk (m.b.v. Ahaslides) als mondeling konden toelichten. Hun bijdrages zijn verwerkt in hoofdstuk 4 (Analyse). Bij de externe consultaties waren ook afgevaardigden van de onderzochte organisaties en de opdrachtgever aanwezig.

De volgende vragen werden bij de externe consultaties besproken:

1. Herkent u zich in de gepresenteerde bevindingen?
2. Hoe kan een ethische omgang met het wettelijk kader het beste geborgd worden?
3. Hoe kunnen de hulpmiddelen optimaal worden ingezet?
4. In hoeverre is er behoefte aan nadere wetgeving?
5. Welk mensenrecht heeft nog verdere versterking nodig?
6. Wat betekenen de kosten en baten voor het proportionaliteitsvraagstuk?

Experts	Organisatie
Daan Kolkman	Jheronimus Academy of Data Science
Albert Meijer	Universiteit Utrecht
Daan van der Neut	Ministerie van Justitie en Veiligheid
Henk de Ruiter	UWV
Marc Woltering	UWV
Peter van Leeuwen	Landelijke Cliëntenraad
Danny Hanse	Nationale Ombudsman
Michel van Leeuwen	Directie AI Ministerie van Justitie en Veiligheid
Arthur Vankan	Dialogic

Guido Terpstra	College voor de Rechten van de Mens
Rolf Leijdekker	SVB
Joke de Boer	CJIB
Linda van der Fliert	Gemeente Amsterdam
Marc Gerrard	ICTU
Patrick Ozer	KPMG

© Rathenau Instituut 2022

Verveelvoudigen en/of openbaarmaking van (delen van) dit werk voor creatieve, persoonlijke of educatieve doeleinden is toegestaan, mits kopieën niet gemaakt of gebruikt worden voor commerciële doeleinden en onder voorwaarde dat de kopieën de volledige bovenstaande referentie bevatten. In alle andere gevallen mag niets uit deze uitgave worden verveelvoudigd en/of openbaar gemaakt door middel van druk, fotokopie of op welke wijze dan ook, zonder voorafgaande schriftelijke toestemming.

Open Access

Het Rathenau Instituut heeft een beleid voor open access. Rapporten, achtergrondstudies, wetenschappelijke artikelen, software worden vrij beschikbaar gepubliceerd. Onderzoeksgegevens komen beschikbaar met inachtneming van wettelijke bepalingen en ethische normen voor onderzoek over rechten van derden, privacy en auteursrecht.

Contactgegevens

Anna van Saksenlaan 51
Postbus 95366
2509 CJ Den Haag
070-342 15 42
info@rathenau.nl
www.rathenau.nl

Bestuur van het Rathenau Instituut

Drs. Maria Henneman - voorzitter
Prof. dr. Noelle Aarts
Drs. Felix Cohen
Dr. Laurence Guérin
Dr. Janneke Hoekstra MSc
Prof. mr. dr. Erwin Muller
Drs. Rajash Rawal
Prof. dr. ir. Peter-Paul Verbeek
Drs. Henk de Jong - secretaris

Het Rathenau Instituut stimuleert de publieke en politieke meningsvorming over de maatschappelijke aspecten van wetenschap en technologie. We doen onderzoek en organiseren het debat over wetenschap, innovatie en nieuwe technologieën.

Rathenau Instituut