



EEN ONAFHANKELIJKE BEOORDELING VAN HET AMSTERDAMSE BEHEERS- KADER VOOR ALGORITMEN

Aviva de Groot
Merel Noorman
Linnet Taylor
Johan Wolswinkel

Tilburg Law School
23 februari 2023

Samenvatting

In 2022 is binnen de gemeente Amsterdam een beheerskader voor algoritmen vastgesteld, getiteld *Programma Algoritme Lifecycle Aanpak*. Dit beheerskader bestaat uit zeven instrumenten die gezamenlijk gerealiseerd zijn om de ontwikkeling en het gebruik van algoritmen door de gemeente Amsterdam (beter) te beheersen. De instrumenten variëren van een publiek toegankelijk algoritmeregister, een *governance*-beschrijving van de diverse rollen en verantwoordelijkheden in de levenscyclus van een algoritme, modelbepalingen voor contracten over algoritmen en een handreiking voor bezwaarbehandelaars tot diverse *impact assessment tools* (mensenrechten, bias en risicoanalyse).

Dit rapport beoordeelt in hoeverre dit Amsterdamse beheerskader voor algoritmen toereikend is voor een verantwoorde toepassing van algoritmen. Daarvoor leunt dit rapport niet alleen op juridische inzichten die zijn ontleend aan het privacyrecht, de mensenrechten en het bestuursrecht, maar maakt het ook gebruik van ethische inzichten die soms verder gaan dan waartoe het geldende recht momenteel verplicht. Gezamenlijk laten deze verschillende perspectieven zien waarop in elk geval moet worden gelet bij de beheersing van algoritmen.

De verschillende instrumenten richten zich op uiteenlopende aspecten bij de ontwikkeling en het gebruik van algoritmen. Daarmee zijn deze instrumenten nuttige en waardevolle bouwstenen om de risico's van de inzet van algoritmen gedurende hun volledige levenscyclus te beheersen. Zo zet het beheerskader sterk in op het duidelijk beschrijven van de verschillende verantwoordelijkheden bij de ontwikkeling van algoritmen. Ook de noodzaak om transparant te zijn over de inzet van algoritmen wordt nadrukkelijk onderkend met het algoritmeregister en het motiveren van beslissingen in bezwaar. Tegelijk zijn de verschillende instrumenten nog duidelijk in ontwikkeling en dekt de naamgeving van het instrument (zoals 'mensenrechtenimpactassessment') niet altijd de lading van de inhoud van het instrument.

Als geheel heeft het huidige Amsterdamse beheerskader voor algoritmen momenteel het karakter van een lappendeken. Niet alleen is de onderlinge verhouding tussen de verschillende instrumenten niet altijd even duidelijk, maar ook zijn er nog duidelijk lacunes in het beheerskader. Zo wordt veel aandacht besteed aan de omgang met geautomatiseerde besluitvorming in bezwaar, maar ontbreekt een vergelijkbaar instrument voor de omgang met geautomatiseerde besluitvorming in de fase van de primaire besluitvorming. Ook is participatie van burgers of betrokken doelgroepen momenteel onvoldoende geborgd in het huidige beheerskader, terwijl dat element (hoewel soms lastiger te verwezenlijken dan een interne checklist van een beheersinstrument) vanuit ethisch perspectief juist cruciaal is voor een sociaal robuuste en maatschappelijk gedragen inzet van algoritmen.

Het Amsterdamse beheerskader voor algoritmen als geheel richt zich sterk op wet- en regelgeving die momenteel van toepassing is. Dit geldende recht is echter nog sterk in ontwikkeling, zowel op nationaal niveau (bijvoorbeeld in de bestuursrechtelijke jurisprudentie) als op internationaal niveau (zoals het voorstel voor een Europese AI-Verordening). Het is daarom van belang dat het beheerskader ook houvast biedt ten aanzien van vraagstukken die juridisch nog onvoldoende zijn beantwoord of uit een oogpunt van een verantwoorde inzet van algoritmen juist verder gaat dan waartoe het recht momenteel verplicht, bijvoorbeeld wat betreft participatie vanuit de maatschappij. Die verantwoorde inzet van algoritmen is niet beperkt tot het enkel beheersen van specifieke risico's op het terrein van privacy of mensenrechten, maar raakt uiteindelijk aan de voortdurende relevante vraag of de inzet van een algoritme (nog steeds) bijdraagt aan het verwezenlijken van het publieke belang dat wordt nagestreefd met het algoritme.

Inhoudsopgave

Samenvatting	2
1. Inleiding	4
1.1 Aanleiding en inbedding.....	4
1.2 Centrale vraag en aanpak.....	4
1.3 Leeswijzer	5
2. Perspectieven op het beheerskader	6
2.1 Inleiding.....	6
2.2 Perspectieven op de inzet van algoritmen.....	6
2.2.1 Privacy.....	6
2.2.2 (Overige) Mensenrechten	8
2.2.3 Bestuursrecht	9
2.2.4 Ethiek.....	10
2.3 Een integraal perspectief	11
3. Beoordeling van het Amsterdamse beheerskader.....	12
3.1 Inleiding.....	12
3.2 Afzonderlijke beheersinstrumenten	12
3.2.1 Algoritmeregister	12
3.2.2 Governancebeschrijving en levenscyclusmodel.....	15
3.2.3 Contractvoorwaarden.....	19
3.2.4 Handreiking bezwaar	22
3.2.5 Mensenrechtenimpactanalysemodel	25
3.2.6 Biasanalysemodel	27
3.2.7 Risicoanalysemodel	28
3.3 Beheersinstrumenten in onderlinge samenhang	30
4. Conclusies en aanbevelingen	33
4.1 Conclusies.....	33
4.2 Aanbevelingen	34

1. Inleiding

1.1 Aanleiding en inbedding

Begin 2022 is het Amsterdamse beheerskader voor algoritmen vastgesteld. Dit beheerskader, getiteld *Programma Algoritme Lifecycle Aanpak*, bestaat uit zeven instrumenten die gezamenlijk beogen (de risico's van) de inzet van algoritmen door de gemeente Amsterdam (beter) te beheersen, zowel tijdens de ontwikkeling als tijdens het gebruik hiervan.

Momenteel verricht de Rekenkamer Amsterdam onderzoek naar een verantwoorde toepassing van algoritmen door de gemeente Amsterdam.¹ Centraal in dat onderzoek staat de volgende, tweeledige vraag:

In hoeverre is het Amsterdamse beheerskader voor algoritmen toereikend voor een verantwoorde toepassing van algoritmen en welke lessen kunnen getrokken worden over het toepassen van algoritmen in de Amsterdamse praktijk?

Ter beantwoording van het eerste onderdeel van deze vraag zal de Rekenkamer Amsterdam het Amsterdamse beheerskader voor algoritmen toetsen aan andere normenkaders, waaronder het normenkader van de Algemene Rekenkamer² en het Europese voorstel voor een AI-Verordening.³ In aanvulling op haar eigen beoordeling heeft de Rekenkamer Amsterdam aan Tilburg Law School (Tilburg University) verzocht om een onafhankelijke beoordeling van het Amsterdamse beheerskader voor algoritmen. Dit rapport beoogt in die beoordeling te voorzien.

1.2 Centrale vraag en aanpak

In lijn met het eerste deel van de onderzoeksvraag van de Rekenkamer Amsterdam staat in dit rapport de vraag centraal in hoeverre het Amsterdamse beheerskader voor algoritmen toereikend is voor een verantwoorde toepassing van algoritmen.

Die toereikendheid van het Amsterdamse beheerskader kan op verschillende manieren worden beoordeeld. In dit rapport is ervoor gekozen om vier beoordelingsperspectieven met elkaar te combineren, namelijk die van de privacy, de mensenrechten, het bestuursrecht en de ethiek. Hoewel die perspectieven soms overlap vertonen en niet altijd even strikt van elkaar kunnen worden gescheiden, vallen ze niet volledig samen en hebben ze daarom toegevoegde waarde ten opzichte van aan elkaar.

Vanuit deze perspectieven kan niet alleen de toereikendheid van elk instrument afzonderlijk binnen het Amsterdamse beheerskader worden beoordeeld, maar juist ook de samenhang tussen de verschillende instrumenten. In dit rapport worden beide beoordelingen gecombineerd, omdat de toereikendheid van het Amsterdamse beheerskader als geheel uiteindelijk wordt bepaald door de wijze waarop de verschillende instrumenten op elkaar aansluiten.

¹ Zie <https://www.rekenkamer.amsterdam.nl/onderzoek/algoritmen/>.

² Zie het Toetsingskader Algoritmes, dat vijf verschillende perspectieven omvat: sturing en verantwoording, model en data, privacy, ITGC en ethiek (waarbij het laatste perspectief de andere vier perspectieven verbindt). Zie <https://www.rekenkamer.nl/onderwerpen/algoritmes/algoritmes-toetsingskader/>.

³ Voorstel van de Commissie voor een Verordening van het Europees Parlement en de Raad voor een verordening tot vaststelling van geharmoniseerde regels betreffende artificiële intelligentie (wet op de artificiële intelligentie), COM(2021) 206 final.

De vraag wat de beste aanpak is om de inzet van algoritmen adequaat te beheersen, laat zich niet eenvoudig en definitief beantwoorden.⁴ De technologische ontwikkelingen rond algoritmische toepassingen volgen elkaar in rap tempo op, terwijl bepaalde effecten van die toepassingen sowieso nog niet in volle omvang zichtbaar zijn. Elk beheerskader kan in dit verband onvolkomen blijken, maar dat neemt niet weg dat een beheerskader zeker een aantal belangrijke kwesties rond de inzet van algoritmen moet adresseren. Om die reden zal deze beoordeling zowel punten benoemen die reeds adequaat onder ogen worden gezien als punten identificeren die nadere aandacht verdienen.

De beoordeling in dit rapport is gebaseerd op documentanalyse. Deze documenten, die gezamenlijk het Amsterdamse beheerskader voor algoritmen vormen, zijn beschikbaar gesteld door de Rekenkamer Amsterdam. Een beperking van deze aanpak is dat de bredere context over en de concrete ervaringen met dit beheerskader, die niet zijn neergelegd in de documenten zelf, niet zijn meegenomen in de beoordeling, hoewel die achtergrondinformatie wellicht antwoord kan geven op de vraag waarom bepaalde keuzes in het beheerskader zijn gemaakt. Wel is tijdens een bijeenkomst op 15 december 2022 met medewerkers van de Rekenkamer Amsterdam en van de gemeente Amsterdam, waarbij de voorlopige bevindingen van dit onderzoek zijn gedeeld, van de zijde van de gemeente meer toelichting en achtergrond gegeven over het beheerskader en de inzet en ontwikkeling daarvan. Die achtergrondinformatie is waar nuttig meegenomen bij het formuleren van de definitieve bevindingen en aanbevelingen.

1.3 Leeswijzer

Dit rapport bevat eerst in hoofdstuk 2 een globaal overzicht van de verschillende perspectieven die in dit rapport worden gehanteerd: privacy, mensenrechten, bestuursrecht en ethiek. Daarbij wordt ook nadrukkelijk de overlap en samenhang tussen de verschillende perspectieven besproken, omdat die overlap en samenhang ook bepalend zullen zijn voor de beoordeling.

Vervolgens wordt in hoofdstuk 3 het Amsterdamse beheerskader voor algoritmen beoordeeld tegen de achtergrond van deze vier perspectieven gezamenlijk. Deze beoordeling ziet zowel op de zeven instrumenten afzonderlijk als op de samenhang tussen de verschillende instrumenten. De beoordeling per instrument wordt steeds gestructureerd aan de hand van een aantal kernbevindingen, waarna een samenvattende beoordeling volgt.

Het rapport eindigt in hoofdstuk 4 met een antwoord op de centrale vraag in hoeverre het Amsterdamse beheerskader voor algoritmen toereikend is voor een verantwoorde toepassing van algoritmen. Verder bevat deze concluderende paragraaf enkele aanbevelingen voor de verdere ontwikkeling van het Amsterdamse beheerskader voor algoritmen.

⁴ Zie ook Wetenschappelijke Raad voor het Regeringsbeleid, *Opgave AI. De nieuwe systeemtechnologie*, Den Haag: WRR 2021, p. 293, waarin wordt gesproken over 'de zoektocht naar de gewenste en noodzakelijke regulering'.

2. Perspectieven op het beheerskader

2.1 Inleiding

Het Amsterdamse beheerskader voor algoritmen is erop gericht de risico's van de inzet van algoritmen te beheersen.⁵ De term 'algoritme' is de laatste decennia aan verandering onderhevig geweest, waarbij in recente jaren de definitie steeds meer is opgerekt. Zo wordt de term in de media vaak als synoniem gebruikt voor AI, terwijl in de computerwetenschappen nauwere definities worden gehanteerd, zoals een stapsgewijze set van instructies om een taak uit te voeren.⁶ Dit rapport hanteert niet een eigen definitie, maar reflecteert op de wijze waarop de term 'algoritme' binnen het beheerskader wordt gedefinieerd en gehanteerd.

Ook de risicobeheersing van die algoritmen kan op verschillende wijzen worden beoordeeld. Het juridische perspectief is daarbij een belangrijk perspectief, maar zeker niet het enige perspectief, zoals ook de ontwikkelingen op Europees niveau laten zien. Zo wordt in de *Ethische richtsnoeren voor betrouwbare KI* van de AI High-Level Expert Group,⁷ die fungeren als *guidelines* voor de inzet van AI, benadrukt dat betrouwbare AI⁸ niet alleen wettig (rechtmatig), maar ook ethisch en robuust moet zijn.⁹ De term 'legitiem' wordt wel gebruikt om te verwijzen naar een inzet van technologie die met al deze dimensies voldoende rekening houdt.

In dit rapport staan vier perspectieven centraal, die van de privacy, de (overige) mensenrechten, het bestuursrecht en de ethiek. Deze perspectieven overlappen deels, maar zijn ook complementair aan elkaar. Hieronder volgt eerst een korte typering van elk van die perspectieven (paragraaf 2.2). Daarna zal de samenhang en interactie tussen de verschillende perspectieven worden beschreven (paragraaf 2.3). Deze integratie van perspectieven zal ook het uitgangspunt zijn voor de beoordeling van het Amsterdamse beheerskader voor algoritmen in hoofdstuk 3.

2.2 Perspectieven op de inzet van algoritmen

2.2.1 Privacy

De bescherming en eerbiediging van de persoonlijke levenssfeer, vaak samengevat tot 'privacy', is een sterk ontwikkeld mensenrecht. De individuele burger kan hiermee opkomen tegen onrechtmatige inmenging in diens autonome bewegings-, ontwikkelings- en keuzevrijheid. AI vroeg bestond vanuit dit perspectief van privacy veel aandacht voor 'geautomatiseerd beslissen' door de overheid vanwege het risico van niet-individuele behandeling dat daarmee wordt gecreëerd. Het individu wordt in die gevallen niet 'gezien' en hun persoonlijke omstandigheden worden niet meegewogen. Daarmee wordt niet alleen de privacy, maar uiteindelijk ook de onderliggende menselijke waardigheid geschonden.

⁵ De instrumenten van het beheerskader worden gepresenteerd als onderdeel van de *Aanpak risicobeheersing algoritmes*.

⁶ Zie uitgebreider over de ontwikkeling van de begrippen 'algoritme' en 'AI': Wetenschappelijke Raad voor het Regeringsbeleid, *Opgave AI: De nieuwe systeemtechnologie*, Den Haag: WRR 2021, p. 43-49. <https://www.wrr.nl/publicaties/rapporten/2021/11/11/opgave-ai-de-nieuwe-systeemtechnologie>

⁷ De AI High-Level Expert Group is een onafhankelijke deskundigengroep die in juni 2018 door de Europese Commissie is opgericht om advies uit te brengen over de omgang met AI. Zie voor meer informatie: <https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai>.

⁸ AI verwijst naar artificiële intelligentie. In dit rapport komen zowel de begrippen 'algoritme' als 'AI' voor. Over de verhouding daartussen, zie onder meer par. 3.3.

⁹ High-Level Expert Group on AI, *Ethische richtsnoeren voor betrouwbare KI* (gepubliceerd op 8 april 2019).

Digitale ontwikkelingen hebben alle beschermde levenssferen in toenemende mate digitaal toegankelijk gemaakt. 'Informatieele' bescherming van al deze sferen is hierdoor noodzakelijk geworden.¹⁰ Dit heeft de bescherming van persoonsgegevens een prominente plaats in het overkoepelende privacyrecht gegeven. Uiteindelijk is de bescherming van persoonsgegevens uitgegroeid tot een zelfstandig mensenrecht binnen de EU¹¹ en de Raad van Europa.¹² Met name de Algemene Verordening Gegevensbescherming (AVG) heeft een belangrijke impuls gegeven aan de normering van geautomatiseerde besluitvorming.¹³ Centraal hierin staat het verbod op volledig geautomatiseerde besluitvorming. Tegelijk benoemt de AVG een aantal mogelijke uitzonderingen op dit verbod, mits een aantal specifieke waarborgen wordt verwezenlijkt.¹⁴ Deze bestaan onder meer uit verschillende transparantieplichtingen voor de verwerkingsverantwoordelijke (waaronder de overheid) en rechten ('data subject rights') voor de betrokkene die met verwerking van persoonsgegevens worden geconfronteerd.¹⁵ Twee belangrijke nuanceringen van dit recht zijn belangrijk om hier te noemen. Allereerst is nog onduidelijk wat precies als betekenisvolle informatie en uitleg over gebruikte algoritmen heeft te gelden binnen de context van bovengenoemde transparantieplichtingen.¹⁶ Daarnaast is in Nederland in artikel 40 van de Uitvoeringswet AVG een zeer ruime uitzonderingsgrond gecreëerd voor volledig geautomatiseerde besluitvorming (behoudens moderne vormen van profilering), waardoor binnen de Nederlandse context relatief weinig bescherming uitgaat van het verbod op volledig geautomatiseerde besluitvorming van artikel 22 AVG en dus ook de bescherming belangrijk is die door andere (juridische) kaders wordt geboden, zoals het (hierna te bespreken) non-discriminatierecht en het bestuursrecht.

De nauwe verwevenheid die vanouds bestaat tussen de rechten van privacy en bescherming van persoonsgegevens, komt ook nadrukkelijk naar voren in recente rechtspraak die laat zien dat de beginselen die aan de AVG ten grondslag liggen, ook de invulling en interpretatie van het recht op bescherming van privacy kleuren. Zo heeft de rechtbank Den Haag in de *SyRI*-uitspraak¹⁷ inspiratie geput uit de AVG bij de toetsing van het Systeem Risico Indicatie aan het grondrecht op bescherming van de persoonlijke levenssfeer (artikel 8 EVRM). Toch blijven ook zeker in het kader van algoritmische besluitvorming belangrijke verschillen bestaan, waardoor beide rechten niet zonder meer met elkaar kunnen worden vereenzelvigd. Zo is voor een schending van het recht op privacy bijvoorbeeld niet nodig dat er persoonsgegevens zijn verwerkt, terwijl het kader van 'persoonlijke' data geen recht doet aan de mogelijkheid van manipulatie van individuen op basis van niet-persoonlijke data.

Zowel het privacyrecht als het gegevensbeschermingsrecht heeft zich sterk ontwikkeld als *individueel* mensenrecht dat aan individuen bepaalde bescherming biedt. Juist in het licht van digitale ontwikkelingen is dit sterk individuele karakter van het privacyrecht en het gegevensbeschermingsrecht een beperking gebleken. Zo kan sprake zijn van onrechtmatige

¹⁰ B.-J. Koops e.a., 'A Typology of Privacy', 38 *University of Pennsylvania Journal of International Law* 2017, p. 483-575.

¹¹ Artikel 8 van het Handvest van de grondrechten van de Europese Unie betreft de bescherming van persoonsgegevens, terwijl artikel 7 van het Handvest van de grondrechten van de Europese Unie ziet op de eerbiediging van het privé-leven en van het familie- en gezinsleven.

¹² Council of Europe Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data (CETS No. 108)

¹³ Zie uitgebreider N. Jak en S. Bastiaans, 'De betekenis van de AVG voor geautomatiseerde besluitvorming door de overheid. Een black box voor een black box?', *Nederlands Juristenblad* 2018, afl. 40, p. 3018-3025.

¹⁴ Zie artikel 22 AVG.

¹⁵ Art. 13 lid 2 aanhef en onder f AVG; art. 14 lid 2 aanhef en onder g AVG; art. 15 lid 1 aanhef en onder h AVG.

¹⁶ De vraag over 'betekenisvolle informatie' ligt momenteel voor bij het Hof van Justitie van de EU (zaak C-203/22).

¹⁷ Rb. Den Haag 5 februari 2020, ECLI:NL:RBDHA:2020:865 (SyRI).

inmenging op basis van groepsprofilering en 'big data processing', terwijl dit voor individuen op onnavolgbare wijze gebeurt. Steeds duidelijker is geworden dat algoritmische besluitvorming juist ook 'systemische' en op institutioneel niveau ingebedde ongelijkwaardigheid bevordert, aangezien algoritmen bestaande sociaal-economische ongelijkheden in de samenleving kunnen reproduceren en zelfs versterken,¹⁸ zoals de recente kinderopvangtoeslagaffaire¹⁹ en de Syri-uitspraak laten zien. Wanneer rechtsbescherming in zulke gevallen beperkt is tot individueel ingestoken processen, kan die tekort schieten.²⁰ Het is daarom belangrijker geworden om het privacyrecht te ontwikkelen op een meer relationeel (minder individueel) ingestoken manier, en om dit recht in te bedden in beheersinstrumenten die beter passen bij de bovengenoemde systemische en institutionele invloed van algoritmische ontwikkelingen en niet enkel gericht zijn op het bieden van individuele rechtsbescherming.

2.2.2 (Overige) Mensenrechten

Het recht op eerbiediging van de persoonlijke levenssfeer (artikel 8 EVRM) is weliswaar een mensenrecht dat al snel wordt getroffen door de inzet van algoritmen, maar de inzet van algoritmen raakt ook aan andere mensenrechten. In dit verband kunnen gelijkheidsrechten, vrijheidsrechten en procedurele rechten worden onderscheiden. Bij vrijheidsrechten kan onder meer worden gedacht aan de impact van algoritmen op de vrijheid van meningsuiting (waaronder demonstreren en zich anderszins organiseren) en het recht op betrouwbare informatie (waaronder medische informatie, bij ontbreken waarvan ook de persoonlijke levenssfeer kan worden aangetast, zoals bij misinformatie over COVID-vaccins). Gelijkheidsrechten staan vooral in het teken van non-discriminatie bij de inzet van algoritmen. Procedurele rechten hebben onder meer betrekking op het recht op een effectief rechtsmiddel (waaronder transparantie van besluiten, uitleg, etc.) en het recht op een eerlijk proces.²¹

De opmerkingen hierboven onder 'privacy' met betrekking tot het ontbreken van een voldoende sterk relationeel en institutioneel karakter treffen ook de overige mensenrechten. In hun huidige karakter zijn zij niet voldoende goed in staat gebleken om te beschermen tegen (vooral) de wijze waarop algoritmen bestaande sociaal-economische ongelijkwaardigheidspatronen versterken. Non-discriminatiewetgeving vraagt bijvoorbeeld om het aantonen van individuele schade, hetgeen niet altijd mogelijk is. Bovendien is deze wetgeving niet gericht op intersectionele discriminatie ofwel discriminatie die niet te herleiden is tot een specifieke factor (bv. nationaliteit), maar juist voortvloeit uit een combinatie van factoren.²² De bestaande kaders hebben ook moeite met ongelijkheid die op nieuwe manieren wordt bewerkstelligd, zoals bij 'voorspellende' systemen die worden ingezet om 'ongewenst' gedrag en -situaties te voorkomen (zoals schooluitval, schulden, kinderverwaarlozing en criminaliteit). Technisch gezien wordt door zulke systemen niets voorspeld, omdat alleen correlaties worden gepresenteerd, gebaseerd op bevindingen uit het verleden. Wanneer echter daarop wordt geacteerd door in de levens van mensen in te grijpen (zoals door bemoeienis van jeugdzorg/buurtteam/politie), dreigt stigmatisering en andere

¹⁸ P. Williams e.a., 'Surfacing Systemic (In)Justices: A Community View' (2022) (<https://systemicjustice.ngo/communityview/>).

¹⁹ Autoriteit Persoonsgegevens, *Belastingdienst/Toeslagen. De verwerking van de nationaliteit van aanvragers van kinderopvangtoeslag*, Den Haag 2020.

²⁰ Zie ook L. Taylor, F. Floridi, en B. Van der Sloot (red.) *Group Privacy: New Challenges of Data Technologies*, Cham: Springer 2016.

²¹ Zie uitgebreider M.J. Vetzo, J.H. Gerards en R. Nehmelman, *Algoritmes en grondrechten*, Den Haag: Boom juridisch 2017.

²² Zie A.L. Hoffmann, 'Where fairness fails: data, algorithms, and the limits of antidiscrimination discourse', 22(7) *Information, Communication & Society* 2019, p. 900-915.

benadeling, zoals het moeten voldoen aan versterkte controles of langduriger aanvraagprocedures.²³

2.2.3 Bestuursrecht

Zowel het recht op privacy in het bijzonder als de mensenrechten in het algemeen zijn onderdeel van het publiekrecht dat de burger tegen de overheid beschermt. De betekenis van het publiekrecht reikt echter verder dan de mensenrechten. Dit blijkt in het bijzonder wanneer de inzet van algoritmen wordt benaderd vanuit het bestuursrecht. Omdat op veel overheidsterreinen sectorspecifieke wetgeving ontbreekt die grenzen stelt aan de inzet van algoritmen, blijkt momenteel een belangrijke rol weggelegd voor de zogeheten 'algemene beginselen van behoorlijk bestuur', zoals het zorgvuldigheidsbeginsel en het motiveringsbeginsel. De overheid moet deze algemene beginselen steeds in acht nemen in aanvulling op de specifieke wetgeving die van toepassing is. Toegepast op 'algoritmisch overheidshandelen' moeten deze beginselen niet alleen worden 'gecontextualiseerd', maar volgens de Afdeling advisering van de Raad van State ook 'verscherpt' worden geïnterpreteerd.²⁴ In de jurisprudentie heeft de Afdeling bestuursrechtspraak van de Raad van State in de zogeheten *AERIUS*-jurisprudentie invulling gegeven aan het beginsel van *equality of arms* door het bestuur te verplichten om in geval van geautomatiseerde besluitvorming de gemaakte keuzes en de gebruikte gegevens en aannames volledig, tijdig en uit eigen beweging openbaar te maken op een passende wijze zodat deze keuzes, gegevens en aannames voor derden toegankelijk zijn. Deze volledige, tijdige en adequate beschikbaarstelling moet het procespartijen mogelijk maken de gemaakte keuzes en de gebruikte gegevens en aannames te beoordelen of te laten beoordelen en zo nodig gemotiveerd te betwisten. Alleen op deze wijze is reële rechtsbescherming mogelijk tegen besluiten die op deze keuzes, gegevens en aannames zijn gebaseerd, en is de rechter in staat om aan de hand van deze informatie de rechtmatigheid van deze besluiten te toetsen.²⁵

In de recente publicatie *Digitalisering. Wetgeving en bestuursrechtspraak* laat de Raad van State duidelijk doorschemeren dat de rechtsontwikkeling op dit terrein zeker nog niet ten einde is.²⁶ Een belangrijke uitdaging voor het bestuursrecht ligt daarom in het concreet invulling geven aan die verscherpte interpretatie van de algemene beginselen van behoorlijk bestuur, juist nu de in een algoritmische context belangrijke beginselen van een 'zorgvuldige voorbereiding van besluiten' en van een 'deugdelijke motivering' slechts in zeer algemene bewoordingen zijn verankerd in de Algemene wet bestuursrecht.²⁷ De focus ligt binnen het bestuursrecht bovendien sterk op het legitimeren van het uiteindelijk genomen besluit, dat immers kan worden aangevochten bij de bestuursrechter, maar minder op de totstandkoming daarvan. Daarnaast dreigt bij geautomatiseerde besluitvorming het risico dat eerst in bezwaar gericht aandacht wordt besteed aan een voor individuen begrijpelijke motivering van het geautomatiseerd genomen besluit. Juist voor burgers in kwetsbare posities die een beperkt 'doenvermogen'²⁸ hebben en niet kunnen beschikken over gesubsidieerde rechtsbijstand, kan het instellen van bezwaar inderdaad soms een onneembare horde zijn, hetgeen het belang van een inzichtelijke motivering in eerste instantie nog groter maakt.

²³ Zie in dit verband <https://www.vpro.nl/argos/lees/onderwerpen/artikelen/2022/in-arme-wijken-voorspelt-de-overheid-nog-altijd-fraude.html>.

²⁴ Afdeling advisering van de Raad van State, 'Ongevraagd advies over de effecten van de digitalisering voor de rechtsstatelijke verhoudingen' (2018), <https://www.raadvanstate.nl/@112661/w04-18-0230/>

²⁵ ABRvS 17 mei 2017, ECLI:NL:RVS:2017:1259 (*AERIUS I*), *AB Klassiek* 2022/42, m.nt. C.J. Wolswinkel, r.o. 14.4. Deze jurisprudentie is later verfijnd in ABRvS 18 juli 2018, ECLI:NL:RVS:2018: 2454 (*AERIUS II*).

²⁶ Raad van State, *Digitalisering. Wetgeving en bestuursrechtspraak*, Den Haag: Raad van State 2021.

²⁷ Zie de artikelen 3:2 en 3:46 Algemene wet bestuursrecht.

²⁸ Wetenschappelijke Raad voor het Regeringsbeleid, *Weten is nog geen doen. Een realistisch perspectief op zelfredzaamheid*, Den Haag: WRR 2017.

De toegenomen belangstelling voor de algemene beginselen van behoorlijk bestuur, die overigens ook buiten de context van algoritmische besluitvorming zichtbaar is,²⁹ biedt daarmee ook een kans om een ambitieuzer normatief kader voor zowel 'handmatige' (menselijke) als algoritmische besluitvorming te ontwerpen. Daarbij komt dat het steeds complexer wordt om geautomatiseerde en handmatige stappen in besluitvorming adequaat te onderscheiden, zodat ook om die reden aanbeveling verdient om integraal naar het geheel aan besluitvorming te kijken vanuit de algemene beginselen van behoorlijk bestuur.³⁰

2.2.4 Ethiek

De algemene beginselen van behoorlijk bestuur kunnen als bestuursrechtelijk vangnet dienen wanneer specifieke wetgeving over de inzet van algoritmen ontbreekt of ontoereikend is. Toch kan ook de bescherming die voortvloeit uit dit juridische kader soms onvoldoende zijn om richting te geven aan de beheersing van algoritmen. Om die reden is het noodzakelijk om verder te kijken dan het geldende recht, zoals dat voortvloeit uit de drie bovengenoemde perspectieven. Hier heeft het perspectief van de ethiek toegevoegde waarde, omdat dit ethische perspectief in staat stelt om te reflecteren op de onderliggende waarden die aan juridische regels ten grondslag liggen en ook richting kan bieden als bestaande wetgeving in het licht van de voortschrijdende technologie ontoereikend is. Zo worden in het reeds aangehaalde rapport 'Ethische richtsnoeren voor betrouwbare KI' vier ethische beginselen onderscheiden: respect voor de menselijke autonomie, preventie van schade, rechtvaardigheid en verantwoording. Deze ethische beginselen kunnen in de woorden van de AI High-Level Expert Group "inspiratie opleveren voor nieuwe en specifieke regelgevingsinstrumenten, hulp bieden bij de uitleg van grondrechten naarmate onze sociaal-technische omgeving in de loop van de tijd verandert, en richting geven aan de motivering voor de ontwikkeling, het gebruik en de toepassing van KI-systemen – waarbij ze zich dynamisch aanpassen naarmate de samenleving zelf verandert."³¹ Vervolgens operationaliseert de AI High-Level Expert Group deze vier ethische beginselen in zeven concrete vereisten voor betrouwbare kunstmatige intelligentie (AI): (1) menselijke controle en menselijk toezicht, (2) technische robuustheid en veiligheid, (3) privacy en datagovernance, (4) transparantie, (5) diversiteit, non-discriminatie en rechtvaardigheid, (6) maatschappelijk en milieuwelzijn en (7) verantwoording.³² Een dergelijke benadering zou ook behulpzaam kunnen zijn bij het contextualiseren van algemene rechtsbeginselen bij algoritmische overheidshandelen.

Een specifiek voorbeeld van een ethisch perspectief dat bijdraagt aan 'sociaal robuuste' AI is de expliciete aandacht voor het nadrukkelijk betrekken van verschillende doelgroepen in de ontwikkeling en het gebruik van algoritmen. Een mogelijke vorm hiervoor betreft de figuur van het burgerpanel, waarin externe partijen (zoals academici) samen met personen in buurthuizen, scholen, etc. nagaan wat hun ervaringen zijn met discriminatie en de rol die technologie daarin speelt. Het is ook mogelijk om vertegenwoordigers van specifieke doelgroepen samen te brengen, hoewel daarin tegelijk een enorme uitdaging kan zijn gelegen. Niet alleen kan de inzet van algoritmen niet bekend zijn bij de betreffende doelgroepen, maar ook kan wantrouwen bestaan bij

²⁹ Een bekend voorbeeld betreft de rol van het evenredigheidsbeginsel bij de terugvordering van kinderopvangtoeslag in de Toeslagenaffaire. Zie in dit verband ook de preconsultatie van het wetsvoorstel tot versterking van de waarborgfunctie van de Algemene wet bestuursrecht, waarbij onder meer de eis wordt gesteld dat een motivering 'voor de belanghebbende redelijkerwijs begrijpelijk' is. Zie <https://www.rijksoverheid.nl/documenten/kamerstukken/2023/01/20/aanbiedingsbrief-tk-toezending-ter-kennisname-teksten-pre-consultatie>.

³⁰ R. Binns & M. Veale, 'Is that your final decision? Multi-stage profiling, selective effects, and Article 22 of the GDPR', 11(4) *International Data Privacy Law* 2021, 319-332.

³¹ AI High Level Expert Group, *Ethische richtsnoeren voor betrouwbare KI*, p. 14.

³² AI High Level Expert Group, *Ethische richtsnoeren voor betrouwbare KI*, p. 17-18.

die doelgroepen jegens de overheid die algoritmen inzet, hetgeen hun betrokkenheid kan belemmeren. Bovendien hoeven de onderscheidingen die het algoritme hanteert niet noodzakelijk samen te vallen met de onderscheidingen van die doelgroepen. Met het oog op sociaal robuuste AI is het niettemin vanuit ethisch perspectief essentieel om die dialoog met betrokken groepen op te zoeken en zo goed mogelijk te organiseren.

2.3 Een integraal perspectief

De vier verschillende perspectieven van privacy, mensenrechten, bestuursrecht en ethiek zijn sterk met elkaar verweven en kunnen niet altijd scherp van elkaar worden onderscheiden. Zo kunnen concrete rechtsregels of rechtsbeginselen binnen de mensenrechten of het bestuursrecht uitdrukking geven aan bovenliggende ethische waarden. Verder kan het recht op privacy (en daarmee nauw verweven het recht op bescherming van persoonsgegevens) worden beschouwd als een (dominant) onderdeel van de mensenrechten wanneer het gaat om de inzet van AI.

In concrete reguleringsinitiatieven van algoritmen of AI komt die verwevenheid tussen de verschillende perspectieven sterk naar voren. Zo onderscheidt het rapport 'Ethische richtsnoeren voor betrouwbare KI' drie dimensies van betrouwbare AI: wettige AI, ethische AI en robuuste AI (uit zowel technisch als sociaal oogpunt). In het rapport wordt erkend dat het aspect van wettige AI, ofwel het voldoen aan alle toepasselijke wet- en regelgeving, kan overlappen met de aspecten van ethische en robuuste AI, omdat die laatste twee aspecten tot op zekere hoogte al in de bestaande wetgeving zijn opgenomen. Tegelijk kan de volledige realisatie van ethische en robuuste AI verder gaan dan de bestaande wettelijke verplichtingen, zodat het wettige aspect op zichzelf tekort schiet voor het realiseren van betrouwbare AI.³³

Ook bij de interpretatie van algemene beginselen van behoorlijk bestuur vindt – al dan niet expliciet – een oriëntatie plaats op andere perspectieven. Zo overweegt de Afdeling bestuursrechtspraak van de Raad van State in haar recente publicatie *Digitalisering. Wetgeving en bestuursrechtspraak* nadrukkelijk dat de AVG een bron van inspiratie is voor rechtsvorming en rechtsbescherming over algoritmische besluiten, en een nuttige 'proeftuin' om algemeen aanvaarde principes als transparantie, uitlegbaarheid en verantwoording te concretiseren.³⁴

Het recente voorstel van de Europese Commissie voor een AI-Verordening geeft eveneens blijk van die verwevenheid van perspectieven. Zo zijn de wettelijke vereisten die ten aanzien van zogeheten 'hoog risico' AI-systemen zijn geformuleerd, grotendeels afgeleid van de hierboven genoemde ethische richtsnoeren³⁵ en ligt aan de beoogde AI-Verordening niet alleen het waarborgen van de werking van de interne markt (artikel 114 VwEU) ten grondslag, maar ook de bescherming van natuurlijke personen in verband met de verwerking van persoonsgegevens (artikel 16 VwEU).

Deze verwevenheid van perspectieven heeft tot gevolg dat het Amsterdamse beheerskader voor algoritmen in dit rapport niet vanuit elk perspectief afzonderlijk wordt beoordeeld, maar vanuit de vier perspectieven gezamenlijk. Dit integrale of gecombineerde perspectief stelt ook beter in staat om eventuele lacunes in het beheerskader te identificeren wanneer instrumenten uit het beheerskader slechts vanuit een bepaald perspectief zijn ontwikkeld.

³³ AI High-Level Expert Group, *Ethische richtsnoeren voor betrouwbare KI*, p. 8 en 14.

³⁴ Raad van State, *Digitalisering. Wetgeving en bestuursrechtspraak*, Raad van State 2021, p. 66.

³⁵ Voorstel van de Commissie voor een Verordening van het Europees Parlement en de Raad voor een verordening tot vaststelling van geharmoniseerde regels betreffende artificiële intelligentie (wet op de artificiële intelligentie), COM(2021) 206 final, p. 15.

3. Beoordeling van het Amsterdamse beheerskader

3.1 Inleiding

Het Amsterdamse beheerskader voor algoritmen bestaat uit zeven verschillende instrumenten, die begin 2022 zijn vastgesteld door het Gemeentelijk Management Team (GMT) van de gemeente Amsterdam:

1. Algoritmeregister
2. Governancebeschrijving en levenscyclusmodel
3. Contractvoorwaarden
4. Handreiking bezwaar
5. Mensenrechtenimpactanalysemodel (MRIA)
6. Biasanalysemodel (BAM)
7. Risicoanalysemodel (RAM)

In dit hoofdstuk zal dit Amsterdamse beheerskader voor algoritmen worden beoordeeld vanuit het gecombineerde perspectief van de privacy, de mensenrechten, het bestuursrecht en de ethiek. Het relatieve gewicht van de verschillende perspectieven zal daarbij enigszins verschillen per instrument. De beoordeling richt zich zowel op de instrumenten afzonderlijk (paragraaf 3.2) als op de instrumenten in onderlinge samenhang (paragraaf 3.3). Per instrument zal eerst een korte beschrijving worden gegeven van het betreffende instrument. Vervolgens zullen de bevindingen ten aanzien van dat instrument thematisch worden gepresenteerd. De bespreking per instrument eindigt met een korte overkoepelende beoordeling, waarin op basis van de verschillende bevindingen de belangrijkste punten worden samengevat.

3.2 Afzonderlijke beheersinstrumenten

3.2.1 Algoritmeregister

Beschrijving

Het algoritmeregister van de gemeente Amsterdam is een publiek toegankelijk website (algoritmeregister.amsterdam.nl) waarop een overzicht wordt gegeven van de algoritmen die de gemeente Amsterdam gebruikt bij 'gemeentelijke dienstverlening'.³⁶ In het register wordt niet ieder algoritme of systeem individueel geregistreerd, maar wordt de informatie over het gebruik van algoritmen verzameld op het niveau van de beleidsdoelstelling (bijvoorbeeld 'parkeercontrole'). Naast beknopte informatie in algemene zin over algoritmen en de impact daarvan op het leven van burgers,³⁷ bevat het algoritmeregister per algoritme (i) algemene informatie over het doel waarvoor en de context waarbinnen het algoritme wordt ingezet, (ii) contactgegevens en (iii) gedetailleerde informatie over het systeem. Die gedetailleerde informatie is onderverdeeld naar (a) datasets, (b) gegevensverwerking, (c) non-discriminatie, (d) menselijk toezicht en (e) risicobeheer.³⁸

Blijkens de white paper *Public AI Registers. Realising AI transparency and civic participation in government use of AI*, waarnaar wordt verwezen op de website van het algoritmeregister, liggen aan het algoritmeregister verschillende doelen ten grondslag. Allereerst beoogt het register

³⁶ In de Engelstalige versie van het algoritmeregister staat het gebruik neutraler geformuleerd: "The Algorithm Register is an overview of the artificial intelligence systems and algorithms used by the City of Amsterdam."

³⁷ Zie <https://algoritmeregister.amsterdam.nl/meer-informatie/>.

³⁸ *Accountability* lijkt in het algoritmeregister gereduceerd tot contactgegevens, *explainability* krijgt geen afzonderlijke aandacht, tenzij het algoritmeregister als zodanig wordt beschouwd als een bijdrage aan *explainability*.

transparantie te bieden aan burgers over de inzet van algoritmen. Daarnaast beoogt het algoritmeregister participatie door burgers te bevorderen door informatie over algoritmen te delen met burgers. Tot slot kan het algoritmeregister ook een interne functie vervullen in het ontwikkelproces van algoritmen, door ontwikkelaars van algoritmen en ambtenaren kaders voor *good governance* te verschaffen, onder meer ten aanzien van de vraag hoe adequate transparantie te bieden.

Bevindingen

[Transparantie] Het algoritmeregister is momenteel nog in ontwikkeling, zoals ook wordt gecommuniceerd op de website. Dat betekent dat het inzicht dat wordt geboden in de inzet van algoritmen door de gemeente Amsterdam, nog beperkt is. Momenteel bevat het algoritmeregister in de Nederlandstalige versie negen algoritmen en in de Engelstalige versie drie algoritmen. De bezoeker van het register weet daardoor niet over welke algoritmen geen informatie wordt verschaft.

Positief aan het algoritmeregister is dat per algoritme eerst het doel en de context worden beschreven waarvoor respectievelijk waarbinnen het algoritme wordt ingezet. Verder is positief dat de informatie die over de verschillende algoritmen wordt gepresenteerd, steeds in dezelfde categorieën wordt aangeboden. Tegelijk loopt de informatie die binnen die categorieën per algoritme wordt verstrekt, sterk uiteen. Zo bevat het ene algoritme wel een privacyverklaring en het andere algoritme niet, terwijl bovendien in sommige (maar niet alle) gevallen wordt verwezen naar de broncode van het algoritme. Zelfs waar uniformiteit in de presentatie van de verschillende algoritmen niet mogelijk is, zou de burger in elk geval houvast moeten worden geboden waarom in het ene geval wel bepaalde informatie wordt verstrekt en in het andere geval niet.

De informatie die wordt verstrekt per algoritme, wordt op een toegankelijke wijze in (zoveel mogelijk) begrijpelijke taal gecommuniceerd. Wel rijst de vraag in hoeverre de gehanteerde formuleringen ook voldoende geloofwaardig en betrouwbaar zijn voor burgers. Daarmee wordt bedoeld dat rekening wordt gehouden met bij hen aanwezige (of afwezige) kennis van een algoritme, maar ook met (doelgroep-) ervaring met een algoritme en/of met de context waarin een algoritme wordt ingezet en/of het beleid waarvan het algoritme deel uitmaakt. Ook met de 'reputatie' van algoritme moet rekening worden gehouden. Zo zal de bewering "Er is geen risico op discriminatie. De hulpverleners zijn opgeleid om burgers te helpen en worden hierop getraind. Overigens zijn er geen specifieke richtlijnen voor het tegengaan van discriminatie." bij het algoritme Vroeg Eropaf door degenen die te maken krijgen met vroegtijdige interventie terecht tot wantrouwen kunnen leiden in het licht van hun ervaringen of reeds aanwezig wantrouwen vergroten ten aanzien van een verantwoorde omgang met algoritmen.

Blijkens de *infographic* voor het invullen van het algoritmeregister moet vooraf een afweging plaatsvinden of bepaalde informatie kan worden verstrekt of dat de Wet open overheid zich daartegen verzet. Het (publiek toegankelijke) algoritmeregister bevat momenteel echter geen informatie over algoritmen die niet kunnen worden geopenbaard. De *infographic* suggereert dat er misschien wel sprake is van een intern 'schaduwregister' dat niet publiek toegankelijk is. Als het doel van het register is om burgers meer inzicht te geven in het gebruik van algoritme en daarover in dialoog te gaan, dan ligt het in de lijn der verwachting dat ook informatie over waarom sommige algoritmen niet verder kunnen worden toegelicht, onderdeel is van het publiek toegankelijke deel van het register. Op die manier kunnen burgers in elk geval nagaan om welke reden het algoritme wordt ingezet en wie toezicht op de inzet van deze algoritmen zou moeten houden als zij zelf geen toegang hebben tot deze algoritmen.

De vraag wat uit een oogpunt van transparantie wel en niet in het algoritmeregister moet worden opgenomen, raakt aan verschillende onderwerpen:

- Wel/niet algoritme: Blijkens de instructie voor het gebruik van het algoritmeregister wordt als algoritme opgevat ‘software die – door gebruik van data-analyse, statistiek of zelflerende logica – geautomatiseerd voorspellingen doet, besluiten neemt of adviezen geeft, die leiden tot directe of indirecte impact op burgers of bedrijven’. Dat roept direct de vraag op of een voorbeeld zoals de Top 400/600 als een algoritme moet worden beschouwd. Op basis van de beschrijving op de website lijkt er vooral sprake te zijn van koppeling van data.³⁹ Het risico van een te smalle definitie van algoritmen is dat de suggestie wordt gewekt dat impact en risico enkel ontstaan bij ‘moderne’ en ‘geavanceerde’ algoritmen, terwijl die juist ook bij ‘eenvoudige’ algoritmen aan de orde kan zijn (bv. verkeersboetes of toeslagenaffaire). De nadruk zou daarom meer moet liggen op de maatschappelijke effecten van een geautomatiseerd systeem bij het beoordelen of het systeem wel of niet opgenomen moet worden in het register, dan op de vraag of het om data-analyse, statistiek of zelflerende logica gaat.
- Negatieve beleidskeuzes: Bevat het algoritmeregister ook informatie over projecten waar de inzet van algoritmen weliswaar is overwogen, maar uiteindelijk niet is ingezet (of aanvankelijk wel is ingezet, maar uiteindelijk is stopgezet)? Hiermee kan werkelijk inzicht worden gegeven in het functioneren van het beheerskader en aan burgers verantwoording worden afgelegd over een weloverwogen inzet van algoritmen. Bovendien geeft dit een mogelijkheid om aan burgers te laten zien dat participatie zin heeft, bijvoorbeeld wanneer uit een impact assessment (of een ander instrument) informatie naar voren komt dat de beslissing om het algoritme te staken of niet te implementeren heeft ondersteund.
- Algoritmen in ontwikkeling: Bevat het algoritmeregister ook informatie over algoritmen die nog in ontwikkeling zijn (*pilots*)? Kan daarbij mogelijk ook iets worden gezegd over op te halen feedback en toekomstige inspraak- of participatiemomenten?
- Niet-openbare algoritmen: Bevat het algoritmeregister een (bepaalde vorm van) transparantie over algoritmen waarover niet onbepaald informatie kan worden verschaft?
- Onderliggende documentatie: Bevat het algoritmeregister ook verwijzingen naar onderliggende totstandkomingsdocumenten, zoals risico-inschattingen en uitkomsten van impactanalyses, en wellicht zelfs naar gedocumenteerde discussies in het totstandkomingstraject? Een dergelijke gedocumenteerde discussie geeft burgers daadwerkelijk inzicht in de werking van het beheerskader.

[Burgerparticipatie en belangengroep participatie] Hoewel ook het bevorderen van burgerparticipatie wordt genoemd als een van de doelen die wordt nagestreefd met het algoritmeregister, is hiervan in de praktijk nog weinig sprake. De website nodigt weliswaar uit tot het geven van feedback, maar onduidelijk is waar die feedback terecht komt en wat met eventuele feedback wordt gedaan. Evenmin biedt het algoritmeregister inzicht in de vraag hoe bij de ontwikkeling van een algoritme (of het staken daarvan) sprake is geweest van burgerparticipatie en wat de inbreng/feedback daarvan is geweest. Overigens kan die burgerparticipatie langs verschillende lijnen vorm krijgen. In sommige gevallen zal een algemeen burgerpanel uitkomst kunnen bieden, maar in andere gevallen zal juist ervoor moeten worden gezorgd dat (vertegenwoordigers van) doelgroepen (*affected groups*) specifiek een stem krijgen (belangengroep participatie).

³⁹ “Alle binnenkomende data wordt vervolgens gecontroleerd door analisten op juistheid, om zeker te zijn dat de personen in de ene dataset dezelfde zijn als in de andere dataset(s). Daarna worden de datasets naast elkaar gelegd, zodat zichtbaar wordt welke personen aan elk onderdeel van de criteria voldoen. Als dit met alle relevante onderdelen van de criteria is gedaan, blijven de personen over die op dat moment aan de criteria voldoen.” Zie <https://algoritmeregister.amsterdam.nl/top-400-600/>.

[Algoritmeregister als onderdeel van beheerskader] Blijkens de handleiding voor het gebruik van het algoritmeregister komt het algoritmeregister pas in beeld zodra het algoritme is ontwikkeld. Daarmee kan het register dus (nog) niet de beoogde *governance* functie binnen het ontwikkelproces vervullen. Het risico bestaat zelfs dat het voldoen aan de transparantie-eisen die het algoritmeregister stelt, wordt opgevat als een rechtvaardiging voor de inzet van het algoritme als zodanig. Het algoritmeregister beantwoordt echter niet de voorvraag of überhaupt de inzet van een algoritme gerechtvaardigd is. In plaats daarvan vooronderstelt het register de inzet van een algoritme en wordt ingegaan op de (risico)beheersmaatregelen die in dit verband zijn getroffen.

In het verlengde hiervan rijst de vraag hoe het algoritmeregister zich verhoudt tot juridische verplichtingen om transparantie te betrachten over de inzet van algoritmen. Zo bevat de AVG ten aanzien van de verwerking van persoonsgegevens de verplichting om de betrokkene te informeren over het bestaan van geautomatiseerde besluitvorming en om de betrokkene nuttige informatie te verschaffen over de onderliggende logica, alsmede het belang en de verwachte gevolgen van die geautomatiseerde verwerking voor de betrokkene.⁴⁰ Ook buiten de context van persoonsgegevens geldt voor het bestuur een bestuursrechtelijke verplichting om in geval van geautomatiseerde besluitvorming de gemaakte keuzes en de gebruikte gegevens en aannames volledig, tijdig en uit eigen beweging openbaar te maken op een passende wijze zodat deze keuzes, gegevens en aannames voor derden toegankelijk zijn.⁴¹ Het is niet duidelijk of het Amsterdamse beheerskader aan deze verplichting beoogt te voldoen door middel van het algoritmeregister en dus of het opnemen van een algoritme in het algoritmeregister wordt beschouwd als voldoende om te voldoen aan de informatieverplichtingen uit de AVG respectievelijk het bestuursrecht. In elk geval moet niet te snel worden aangenomen dat het algoritmeregister in zijn huidige opzet die functie kan vervullen: weliswaar bevat het register bepaalde informatie over het algoritme (hoewel uiteenlopend per algoritme), maar die informatie kan onvoldoende zijn voor een individu om in zijn concrete geval te begrijpen waarom een bepaalde beslissing is genomen of hoe zijn persoonsgegevens zijn verwerkt.

Beoordeling

Het algoritmeregister is een belangrijk instrument om bij te dragen aan transparantie over de inzet van algoritmen door de gemeente Amsterdam. Omdat het algoritmeregister nog in ontwikkeling is, kan zowel de verstrekte informatie per algoritme als algemene informatie over wat wel en niet in het algoritmeregister wordt opgenomen, worden verbeterd. In de huidige uitwerking nodigt het algoritmeregister nog onvoldoende uit tot het bevorderen van burger- of belangengroep participatie, terwijl de rol die het algoritmeregister kan vervullen in het ontwikkelproces van het algoritme, eveneens nog onderontwikkeld is.

3.2.2 Governancebeschrijving en levenscyclusmodel

Beschrijving

Het tweede instrument van het beheerskader betreft een beschrijving van de *governance*-structuur in het document 'Taken en verantwoordelijkheden bij het gebruik van Algoritmen door de gemeente Amsterdam'. Dit document gaat vergezeld van een presentatie, waarin allereerst de levenscyclus van een algoritme visueel wordt gepresenteerd. Verder wordt aangegeven wie wat doet, waarbij onderscheid wordt gemaakt tussen de opdrachtgever (de directeur of eventueel *product owner*), de opdrachtnemer (het ontwikkelteam of de leverancier) en stakeholders (waaronder juridische experts, *security officers*, *privacy officers*, ethische experts, etc.). In geval

⁴⁰ Zie de artikelen 13 en 14 AVG.

⁴¹ ABRvS 17 mei 2017, ECLI:NL:RVS:2017:1259 (*AERIUS I*), r.o. 14.4.

van een algoritme met een 'hoog risico' dienen de stakeholders blijkens de *visual* ook in het ontwikkelteam zelf te zijn vertegenwoordigd. Tot slot worden de verschillende verantwoordelijkheden visueel gepresenteerd in de figuur van een boom met vertakkingen, waarbij onderscheid wordt gemaakt tussen de bestuurlijke verantwoordelijkheid, de ambtelijke verantwoordelijkheid en de verantwoordelijkheid voor advies en toezicht. Het document 'Taken en verantwoordelijkheden bij het gebruik van Algoritmen door de gemeente Amsterdam' (hierna: het *governance*-document) sluit aan bij deze visuele weergave en beschrijft uitgebreid de verschillende verantwoordelijkheden, zowel bestuurlijk als ambtelijk als op het terrein van advies en toezicht.

Bevindingen

[Doel: beperking van risico's of meer?] Het *governance*-document is bedoeld zowel om innovatie door middel van slimme technologie te stimuleren als om de mogelijk negatieve effecten te beperken door het beheersen van risico's. Hoewel het document in de uitwerking ervan met name gericht lijkt op het beheersen van risico's, wordt in het document terecht ook op meerdere punten in de levenscyclus de verantwoordelijkheid belegd om het (beleids)doel dat met de inzet van het algoritme zou moeten worden gediend, vast te leggen en vervolgens hieraan te toetsen. Wel dient te worden beseft dat het (beleids)doel geen statisch gegeven hoeft te zijn tijdens de ontwikkeling en de inzet van algoritmen, zodat ook de noodzaak kan bestaan om het beleidsdoel zelf (dat aan het begin mogelijk nog onvoldoende scherp is gedefinieerd) tijdens het ontwikkelproces kritisch tegen het licht te houden. Een belangrijk aandachtspunt is daarom, mede gelet op de bestuurlijke verantwoordelijkheid, wie gedurende de levenscyclus van het algoritme toeziet op naleving (en eventueel herijking) van eerder genomen besluiten op basis van juridische en ethische afwegingen. Met name de verantwoordelijkheid voor het actief incorporeren van externe feedback van burgers is minder nadrukkelijk uitgewerkt in het *governance*-document.

[Bestuurlijke verantwoordelijkheid] Het *governance*-document benoemt nadrukkelijk de bestuurlijke verantwoordelijkheid naast de ambtelijke verantwoordelijkheid. Dit is zinvol om te verzekeren dat de inzet van algoritmen niet louter een ambtelijke exercitie is. Het college van burgemeester en wethouders is verantwoordelijk voor de toepassing van algoritmen die door of namens de gemeente worden uitgevoerd. Binnen het college is de portefeuillehouder ICT verantwoordelijk voor het door het college laten vaststellen van kaders voor het werken met algoritmen. Verder zijn de vakportefeuillehouders verantwoordelijk voor de betrouwbaarheid van de algoritmen die worden 'toegepast' binnen hun vakportefeuilles en voor het ter besluitvorming aan het college voorleggen van algoritmen met een hoog risico binnen hun vakportefeuille. Het *governance*-document maakt echter niet duidelijk waarom enkel dergelijke algoritmen ter besluitvorming moeten worden voorgelegd aan het college, terwijl de term 'toepassing' de vraag opwerpt welke verantwoordelijkheid het college heeft voor het ontwikkelen van en het experimenteren met algoritmen. Deze vraag is temeer van belang, aangezien voor algoritmen die worden ontwikkeld als deel van bijvoorbeeld *living labs* of andere experimentele omstandigheden, niet altijd duidelijk is wanneer een algoritme is 'uitontwikkeld', zeker niet als dit algoritme intern wordt ontwikkeld.

[Ambtelijke verantwoordelijkheid: stelselverantwoordelijkheid] Als het gaat om de ambtelijke verantwoordelijkheid, maakt het *governance*-document onderscheid tussen de lijnverantwoordelijkheid en de stelselverantwoordelijkheid. Binnen de stelselverantwoordelijkheid neemt de Chief Information Officer (CIO) een centrale rol in; hij is verantwoordelijk voor het laten vaststellen van kaders voor het toepassen van algoritmen en de regie als het gaat om het beheersen van risico's. Het *governance*-document laat echter zien dat juist de ontwikkeling en

inzet van algoritmen buitengewoon veelzijdig is en daardoor raakt aan de (stelsel)verantwoordelijkheid van verschillende personen binnen de ambtelijke organisatie. Dit maakt het *governance*-model complexer en vergt dus waarborgen voor voldoende afstemming tussen de verschillende stelselverantwoordelijken (zoals CIO, Chief Information Security Officer (CISO), Informatiecommissaris (IC) en Directeur Digitalisering & CTO). Bovendien lijkt het (beleids)doel of maatschappelijk nut waarvoor het algoritme wordt ingezet, bij deze stelselverantwoordelijken geen belangrijke rol te vervullen, waardoor deze stelselverantwoordelijkheid ten onrechte wordt gereduceerd tot risicobeheersing.

[Ambtelijke verantwoordelijkheid: lijnverantwoordelijkheid] De meeste aandacht gaat in het *governance*-document uit naar de lijnverantwoordelijkheid, waarbij een onderscheid wordt gemaakt tussen de directeur als opdrachtgever en het ontwikkelteam als opdrachtnemer. De verantwoordelijkheid voor de betrouwbaarheid van algoritmen die worden ontwikkeld, ingekocht of gebruikt, berust primair bij de directeur van de betreffende directie (of het dagelijks bestuur van het stadsdeel). De verantwoordelijkheden van de directeur zijn niet-limitatief omschreven en zijn niet altijd even scherp geformuleerd (wat betekent bijvoorbeeld ‘het toetsen van de impact van het beoogde algoritme bij inwoners en ondernemers van de stad’?). Ook de verantwoordelijkheden van het ontwikkelteam worden uitgebreid beschreven, hoewel tegelijk wordt opgemerkt dat de taken van het ontwikkelteam onderdeel zijn van de opdracht van de directeur (of dagelijks bestuur van het stadsdeel) en daarmee onder de verantwoordelijkheid vallen van die directeur (of dat dagelijks bestuur van het stadsdeel).

[Hoog risico] Wat opvalt aan de *governance*-beschrijving is dat algoritmen met een ‘hoog risico’ een bijzondere status hebben. Blijkens de *governance*-beschrijving is alleen een KI Impact Assessment nodig indien sprake is van een algoritme met hoog risico (terwijl het KI Impact Assessment zelf (zie 3.3.5) dat niet geheel duidelijk maakt en die conclusie meer open laat). Bovendien geldt alleen in geval van hoog risico de eis dat stakeholders ook daadwerkelijk worden toegevoegd aan het ontwikkelteam. Het is op zichzelf positief dat stakeholders niet pas achteraf bij het ontwikkelde algoritme worden betrokken. Tegelijk is niet duidelijk waarom voor het al dan niet betrekken van stakeholders in het ontwikkelteam, het criterium ‘hoog risico’ uit het Commissievoorstel voor een AI-Verordening leidend zou moeten zijn. Daarbij komt dat de *governance*-beschrijving zelf niet duidelijk aangeeft wanneer sprake is van een algoritme met ‘hoog risico’. Sowieso kan de negatieve connotatie van ‘risico’ ertoe leiden dat systemen die beogen te leiden tot ‘positieve beïnvloeding’ in het leven van burgers (bv. algoritmen die zijn gericht op het identificeren van potentieel schrijnende gevallen), minder serieus worden genomen, terwijl de impact daarvan op burgers eveneens groot kan zijn, zoals bij algoritmische signaleringen gericht op schooluitval en (kinder)verwaarlozing, problematische schulden of (jeugd)criminaliteit. Het Top 400/600 algoritme laat bijvoorbeeld zien hoe interventies gestoeld op zulke ‘voorspellingen’ stigmatiserend werken en ook direct inbreken in de individuele persoonlijke levenssfeer van alle gezinsleden. Bij sommige groepen burgers zal het effect van een dergelijk algoritme kunnen zijn dat burgers zich zelfs niet meer aanmelden voor (psycho-medische, financiële, sociale) hulp waarbij ze wél baat bij zouden hebben, een risico dat ook bestaat bij al te gemakkelijk gerealiseerde datakoppelingen.

[Stakeholders] Hoewel verschillende stakeholders worden benoemd, is een belangrijke ontbrekende stakeholder de vertegenwoordiging vanuit *affected groups* (belanghebbende burgers en groepen in de stad). Wanneer het onmogelijk blijkt om die stakeholders bij de ontwikkeling van een algoritme te betrekken, zou dat kunnen worden opgevat als een serieuze *red flag* voor het gebruiken van dit algoritme. Andere experts (zoals een ethisch expert) kunnen wellicht wijzen op

het belang van burgerparticipatie of inspraak voor *affected groups* (belangengroep participatie), maar kunnen deze vertegenwoordiging niet vervangen. In het verlengde hiervan zou de verantwoordelijkheid om deze groepen bij het ontwikkelproces te betrekken, ook expliciet moeten worden benoemd en belegd.

[Advies en toezicht] Ten aanzien van advies en toezicht lijkt een belangrijke rol te zijn weggelegd voor de *privacy officer* (eerste lijn) en voor de Commissie Persoonsgegevens Amsterdam (CPA) en de Functionaris Gegevensbescherming (FG) in de tweede lijn. Zo adviseert de CPA omtrent de gegevensverwerking in breedste zin en wordt daaronder ook begrepen 'de ethische toets' bij de toepassing van algoritmen. De vraag is of het toezicht op algoritmen door het beleggen van deze taken bij deze toezichtsverantwoordelijkheden niet (onbedoeld) grotendeels wordt bepaald door het perspectief van privacy en gegevensbescherming en of deze instanties voldoende zijn toegerust om ook andere perspectieven te betrekken in hun beoordeling. Het toezien op ethische principes en mensenrechten is immers een nieuwe taak voor deze twee partijen, die blijkens hun naamgeving van oudsher met name gericht zijn op privacy en dataprotectie.⁴² In elk geval zal nadrukkelijk moeten worden geborgd dat ook die andere perspectieven in het toezicht betrokken worden.

[Inhoudelijke sturing] In het algemeen is de *governance*-beschrijving zeer procedureel van aard. Zij beschrijft de stappen die door verschillende personen moeten worden genomen en heeft zodoende het karakter van een uitgebreide, maar niet-limitatieve checklist. De *governance*-beschrijving biedt echter weinig inhoudelijke houvast. Zo is de directeur verantwoordelijk voor de beslissing om een algoritme zelf te ontwikkelen of uit te besteden (*buy-or-make decision*), maar maakt het document niet duidelijk welke overwegingen leidend zijn in de keuze om wel of niet te outsourcen. Datzelfde kan beargumenteerd worden voor de beslissing om de ontwikkeling van een algoritme te staken: het is onduidelijk door wie en op welke gronden dit kan worden gedaan, terwijl wel in een 'definition of done' is voorzien. Ook blijkt uit de beschrijving niet duidelijk of er inhoudelijke kaders zijn als het gaat om het al dan niet realiseren van menselijke tussenkomst. Mogelijk zijn de kaders voor die inhoudelijke keuzes elders vastgesteld, maar de *governance*-beschrijving verwijst niet daarnaar.

[Assessments] Ook de relatie tot andere instrumenten uit het beheerskader is niet altijd duidelijk. Hoewel op enkele punten wordt verwezen naar het Kunstmatige Intelligentie Impact Assessment of naar vergelijkbare instrumenten, is niet duidelijk of hiermee wordt verwezen naar de andere instrumenten uit het beheerskader. Ook wanneer wordt gesproken over het vastleggen van de 'risicoanalyse' van het algoritme voordat een beslissing wordt genomen over het ontwikkelen of inkopen ervan, is niet geheel duidelijk welke risicoanalyse hiermee wordt bedoeld. De *governance*-beschrijving zou als verbindend document duidelijk moeten maken wanneer en door wie de verschillende assessment tools zouden moeten worden toegepast, maar is op dit punt niet geheel duidelijk. Ook blijft onduidelijk wat 'de ethische kaders' zijn.

[Levenscyclus] Hoewel het gevisualiseerde *governance*-model wordt gepresenteerd als een iteratief model, blijkt het in de gesuggereerde toepassing nog tamelijk lineair: stappen 1 tot en met 5 worden sequentieel doorlopen. Dat roept de vraag op welke tussentijdse feedbackloops zijn

⁴² Zie bijvoorbeeld de website van de CPA: "De CPA adviseert over het privacybeleid van de gemeente en de uitvoering daarvan. Daarnaast adviseert de commissie de ambtelijke organisatie bij complexe en/of politiek gevoelige kwesties rondom persoonsgegevens. De CPA signaleert wat er binnen de gemeente speelt op het gebied van persoonsgegevens en brengt dit onder de aandacht van het gemeentebestuur en het management van de ambtelijke organisatie." <https://www.amsterdam.nl/bestuur-organisatie/organisatie/overige/adviesraden/commissie-persoonsgegevens-amsterdam/>.

gerealiseerd (bijvoorbeeld als tijdens het ontwikkelen van het algoritme blijkt dat er een nieuwe dataset nodig is of als in praktijk blijkt dat bepaalde principiële keuzes of een beleidsdoel moeten worden herzien). Verder valt op dat de verantwoordelijkheden in de ontwikkelfase uitgebreider worden beschreven dan in de fase van toepassing en evaluatie. Dat roept de vraag op hoe die evaluatie eruit ziet en wanneer bijvoorbeeld wordt geoordeeld dat een algoritme 'effectief' of 'succesvol' is.

Beoordeling

De *governance*-beschrijving bevat een uitgebreid overzicht van de taken en verantwoordelijkheden van diverse actoren bij het gebruik van algoritmen. Daarmee doet de beschrijving recht aan een praktijk waarin vele facetten van de ambtelijke organisatie worden geraakt en wordt richting gegeven aan de wijze waarop de ontwikkelpraktijk zou behoren te verlopen. Het zou goed zijn om een speerpunt te maken van het verhelderen en versimpelen van de huidige opzet. Op dit moment leidt het uitgebreide overzicht ook tot complexiteit en onduidelijkheid, met name in de verhouding tussen opdrachtgever (directeur) en opdrachtnemer (ontwikkelteam). Bovendien dreigt het risico van versnippering binnen de ambtelijke organisatie.

3.2.3 Contractvoorwaarden

Beschrijving

In aanvulling op de Gemeentelijke inkoopvoorwaarden bij IT van de gemeente Amsterdam (GIBIT) heeft de gemeente Amsterdam contractvoorwaarden ontwikkeld voor 'het rechtvaardig gebruik van Algoritmische toepassingen'. Deze contractvoorwaarden zijn zowel van toepassing wanneer een algoritmische toepassing op zichzelf wordt geleverd als wanneer die algoritmische toepassing deel uitmaakt van een meeromvattende ICT-prestatie. De contractvoorwaarden hebben betrekking op de kwaliteit van zowel de data als de algoritmische toepassing. Bijzondere aandacht wordt besteed aan transparantie van de algoritmische toepassing, de te volgen risicomanagementstrategie, het beheer en het auditen van de algoritmische toepassing. De reden voor de gemeente Amsterdam om deze contractvoorwaarden op te stellen, is dat dergelijke voorwaarden tot nog toe ontbraken.

In de terminologie van de contractvoorwaarden wordt een onderscheid gemaakt tussen enerzijds de 'algoritmische toepassing' en anderzijds het gebruik hiervan bij het nemen of voorbereiden van 'besluiten'. Een algoritmische toepassing wordt daarbij ruim gedefinieerd als 'software waarmee op geautomatiseerde wijze voorspellingen worden gedaan, beslissingen worden genomen en/of adviezen worden gegeven door gebruik te maken van data-analyse, statistiek en/of zelflerende logica', terwijl 'Besluiten' zien op 'besluiten van de gemeente die bestuursrechtelijk, privaatrechtelijk en/of feitelijk van aard zijn en die direct of indirect één of meer burgers van de gemeente, bezoekers van de gemeente of bedrijven of andersoortige instellingen die in de gemeente zijn gevestigd in aanmerkelijke mate treffen'.

Bevindingen

[Make-or-buy] Het instrument van contractvoorwaarden heeft betrekking op de situatie waarin de gemeente ervoor heeft gekozen de ontwikkeling van een algoritmische toepassing aan een externe partij over te laten.⁴³ Dit betekent dat de *make-or-buy* beslissing al is genomen en dat de contractvoorwaarden hierop geen betrekking hebben. Evenmin geven andere instrumenten uit het Amsterdamse beheerskader voor algoritmen inhoudelijk richting aan deze voorafgaande beslissing. Daarnaast is de vraag of aan *in-house* ontwikkeling van algoritmische toepassingen

⁴³ "De Contractbepalingen zijn niet van toepassing indien de Opdrachtgever onder eigen regie een Algoritmische toepassing ontwikkelt en daarbij gebruikmaakt van personeel van een externe partij."

dezelfde eisen worden gesteld. Zo dragen de contractvoorwaarden de leverancier op om de kwaliteit van de algoritmische toepassingen en de risicomanagementstrategie voortdurend te monitoren en waar nodig te actualiseren. Het beheerskader is via het *governance*-model veel minder expliciet en concreet als het gaat om een vergelijkbare verplichting voor *in-house* ontwikkelde algoritmische toepassingen.

[Bereik: 'in aanmerkelijke mate'] In de contractvoorwaarden is ervoor gekozen om het bereik van de contractvoorwaarden te beperken tot 'besluiten', waarvan enkel sprake is indien beslissingen burgers 'in aanmerkelijke mate' treffen. De beperking tot beslissing met een *significant effect* ('in aanmerkelijke mate') is blijkens de toelichting op de contractvoorwaarden ontleend aan artikel 22 van de AVG. Op basis van deze bepaling is volledig geautomatiseerde besluitvorming uitgesloten indien een beslissing rechtsgevolg heeft of de betrokkene 'in aanmerkelijke mate' treft. Het is echter onduidelijk waarom bij de contractvoorwaarden wordt gekozen voor deze zelfde beperking ('in aanmerkelijke mate'). De beperking lijkt namelijk te zijn ontleend aan een andere context (namelijk de noodzaak om al dan niet menselijke tussenkomst te garanderen),⁴⁴ terwijl de contractvoorwaarden op een ruimer palet aan maatregelen zien (waaronder transparantie, risicomanagementstrategie, etc.). Bovendien kan door de toevoeging van dit element aan de definitie discussie ontstaan over de interpretatie van het criterium 'in aanmerkelijke mate', die bovendien zeer casuïstisch zal zijn en daarmee (rechts)onzekerheid over de toepasselijkheid van de contractvoorwaarden in de hand kan werken.

[Data en algoritmen] De contractvoorwaarden maken op het eerste gezicht een strikt onderscheid tussen enerzijds de algoritmische toepassing en anderzijds de data (als input en output van de algoritmische toepassing). Juist omdat in de toelichting ook wordt erkend dat data en algoritmische toepassing overlap kunnen vertonen, is de vraag hoe dit onderscheid tussen data en algoritmische toepassingen doorwerkt in de contractvoorwaarden. Zo wordt expliciet bepaald dat de rechten op de data (in beginsel) toekomen aan de opdrachtgever, maar laten de contractvoorwaarden in het midden welke rechten aan de opdrachtgever toekomen ten aanzien van de algoritmische toepassing, dus de software. Wordt er bewust niet over de rechten op de 'algoritmische toepassing' gecontracteerd (bijvoorbeeld aan wie de broncode toekomt)? Hoe zit het met het hergebruik van modellen door de leverancier (bijvoorbeeld bij *transfer learning*)? Het lijkt erop dat de contractvoorwaarden impliciet ervan uitgaan dat de broncode bij de leverancier blijft berusten (en dat hij zelfs ervoor kan kiezen de opdrachtgever zelf geen inzage hierin te verlenen), maar de reden daarvoor (bijvoorbeeld dat de verplichting tot overdragen van de broncode een onredelijke of zelfs verboden contractvoorwaarde zou zijn) blijkt niet uit de toelichting op de contractbepalingen.

[Transparantie] Een belangrijk onderwerp binnen de contractvoorwaarden betreft transparantie van algoritmische toepassingen. De contractvoorwaarden maken in dit verband een onderscheid tussen technische transparantie, procedurele transparantie en uitlegbaarheid. Positief is dat in de contractvoorwaarden bij de bespreking van deze vormen van transparantie een koppeling wordt gelegd met andere instrumenten uit het beheerskader. Procedurele transparantie hangt bijvoorbeeld samen met de informatie die verstrekt wordt in het algoritmeregister, terwijl uitlegbaarheid ook raakt aan de bestuursrechtelijke motiveringsplicht bij individuele besluiten. Voor de concrete invulling van deze transparantiestandaarden wordt in de contractvoorwaarden mede aangesloten bij de uitleg die de bestuursrechter in de zogeheten AERIUS-jurisprudentie heeft gegeven aan het beginsel van 'equality of arms'. Het is positief dat de contractvoorwaarden de eis

⁴⁴ Zo zijn de transparantieverplichtingen die de artikelen 13, 14 en 15 van de AVG bevatten ten aanzien van geautomatiseerde besluitvorming niet beperkt tot beslissingen die de betrokkene in aanmerkelijke mate treffen.

van uitlegbaarheid concreet maken, namelijk het inzichtelijk maken van de belangrijkste factoren die ten grondslag liggen aan een bepaalde uitkomst en het benoemen welke wijzigingen in de input zouden moeten worden doorgevoerd om tot een andere uitkomst te komen (zogenoeten 'counterfactuals'). Deze concrete invulling is nog niet met zoveel woorden geëxpliciteerd in de jurisprudentie en kan daarmee worden opgevat als een verscherpte interpretatie van het motiveringsbeginsel. Die verscherpte interpretatie is toe te juichen en bovendien in lijn met huidige voornemens om het bestuursrechtelijke motiveringsbeginsel sowieso aan te scherpen met de eis dat een motivering voor de belanghebbende redelijkerwijs begrijpelijk moet zijn.⁴⁵ Tegelijk wordt in de contractvoorwaarden ruimte geboden om van deze transparantiestandaarden af te wijken, hetgeen de vraag oproept welke grenzen aan transparantie door de gemeente worden geaccepteerd.

[Toepassing] Ook al is geen sprake van onredelijke contractvoorwaarden die *juridisch* niet door de beugel kunnen, kan toch de vraag gesteld worden welke 'prijs' deze contractvoorwaarden hebben. Enerzijds kunnen zij betekenen dat een leverancier wel een bepaalde algoritmische toepassing wil leveren, maar enkel tegen een hogere prijs. Is de gemeente bereid die dan te betalen? Anderzijds kunnen zij betekenen dat de leverancier niet akkoord wil gaan met een of meerdere modelbepalingen. De vraag is in hoeverre de gemeente in een dergelijk geval bereid is om te contracteren in afwijking van de modelbepalingen. Het beheerskader bevat momenteel geen waarborgen voor het proces waarin die afwijkende contractvoorwaarden zouden worden opgesteld.

[Volledige levenscyclus] De contractvoorwaarden voorzien niet alleen in het als zodanig leveren van een algoritmische toepassing, maar ook in het eventueel onderhouden en monitoren daarvan gedurende het gebruik. Die benadering doet recht aan de levenscyclusbenadering die centraal staat in het Amsterdamse beheerskader voor algoritmen. Tegelijk erkent de toelichting op de contractvoorwaarden dat die 'onderhoudsverplichting' niet van toepassing is indien de leverancier uitsluitend een algoritmische toepassing in opdracht van de gemeente ontwikkelt en daarna de contractuele relatie tussen de gemeente en de leverancier eindigt. Hoewel in dit situatie blijken de contractvoorwaarden nog wel een 'inspanningsverplichting' geldt om transparantie te bieden, laat de toelichting op de contractvoorwaarden tevens doorschemeren dat juist in die situatie aanvullende contractvoorwaarden nodig kunnen zijn zonder richting te geven aan de aard en de inhoud van die voorwaarden.

Beoordeling

Het is positief dat het Amsterdamse beheerskader expliciet aandacht besteedt aan de binding van private partijen aan publiekrechtelijke normen. Dit instrument laat zien dat het 'uitbesteden' van de ontwikkeling van algoritmen niet hoeft te betekenen dat niet langer kan worden voldaan aan elementaire normen op het terrein van transparantie en risicomanagement. Daarbij is gekozen voor een fijnmazige benadering die recht beoogt te doen aan de verschillende belangen van leverancier, gemeente en burger. De modelbepalingen zijn fijnmazig uitgewerkt en beogen daarmee recht te doen aan de verschillende eisen die in wet- en regelgeving worden gesteld. Tegelijk is de vraag in hoeverre die bepalingen in de praktijk voldoende werkbaar zijn en daadwerkelijk sturend richting de leverancier kunnen zijn gelet op het risico van *vendor lock-in*.

⁴⁵ Zie de preconsultatie van het wetsvoorstel tot versterking van de waarborgfunctie van de Algemene wet bestuursrecht: [wetshttps://www.rijksoverheid.nl/documenten/kamerstukken/2023/01/20/aanbiedingsbrief-tk-toezending-ter-kennisname-teksten-pre-consultatie](https://www.rijksoverheid.nl/documenten/kamerstukken/2023/01/20/aanbiedingsbrief-tk-toezending-ter-kennisname-teksten-pre-consultatie).

3.2.4 Handreiking bezwaar

Beschrijving

Het Amsterdamse beheerskader bevat een handreiking voor bezwaarbehandelaars. Deze handreiking bevat aanbevelingen voor behandelaars van een bezwaar waarbij een algoritme ten grondslag ligt aan het genomen primaire besluit. De handreiking bevat voor verschillende fases in de behandeling van het bezwaar (indiening van bezwaar, opstellen verweerschrift, hoorzitting en samenstelling adviescommissie, beslissing op bezwaar) aanbevelingen die zijn toegespitst op algoritmen en bevat ook enkele bijlagen die meer achtergrondinformatie geven over de geldende juridische verplichtingen (AVG, SyRI-uitspraak, AERIUS-jurisprudentie).

Bevindingen

[Bestuursrechtelijke oriëntatie] De aanbevelingen die de handreiking (inclusief bijlagen) bevat, leunen sterk op de jurisprudentie en de literatuur die hierover zijn verschenen. Zo bouwt de handreiking voort op de AERIUS-jurisprudentie en de 'verfijning' die in die jurisprudentie is aangebracht tussen standaard(invoer)gegevens en maatwerkgegevens. Die verfijning in de jurisprudentie houdt qua terminologie echter sterk verband met het specifieke systeem AERIUS dat in die (stikstof)jurisprudentie centraal stond,⁴⁶ waardoor het letterlijk herhalen van die terminologie de bezwaarbehandelaar mogelijk onvoldoende houvast biedt. Terecht vraagt de handreiking daarnaast aandacht voor het zorgvuldigheidsbeginsel en het motiveringsbeginsel, maar dat roept uiteraard ook weer nieuwe vragen op: wat zijn onder een verscherpte interpretatie van die beginselen in een algoritmische context 'relevante feiten' en 'rechtstreeks betrokken belangen' voor een zorgvuldige voorbereiding van besluiten (artikel 3:2 Awb) en wat is een 'deugdelijke motivering' (artikel 3:46 Awb)? Het is daarom verstandig dat de handreiking op een aantal punten verder gaat dan de bestuursrechtelijke stand van zaken op basis van de huidige jurisprudentie, die immers nog duidelijk in ontwikkeling is.⁴⁷ Zo geeft de handreiking een voorzet als het gaat om een 'deugdelijke motivering', namelijk het benoemen van de belangrijkste factoren en het benoemen van *counterfactuals* (zoals ook in de contractvoorwaarden het geval is). Tegelijk biedt de handreiking minder houvast als het bijvoorbeeld gaat om het afbakenen van de 'relevante feiten' en 'rechtstreeks betrokken belangen', juist nu een verscherpte interpretatie in een algoritmische context zou kunnen meebrengen dat op voorhand geen enkel feit of belang kan worden uitgesloten. De handreiking zou in dit verband ook expliciet kunnen benoemen waarin ze zelf minder richtinggevend is.

Daarnaast moet in de handreiking niet uit het oog worden verloren dat ook andere algemene beginselen van behoorlijk bestuur, zoals het evenredigheidsbeginsel waaraan sinds kort intensiever kan worden getoetst,⁴⁸ aan belang kunnen winnen in een algoritmische context. Bovendien kan de bezwaarbehandelaar worden geconfronteerd met kwesties die nog niet in de jurisprudentie zijn geadresseerd, zoals de vraag in hoeverre het bestuursorgaan inzage moet bieden in de broncode van algoritmische toepassingen die door derde partijen zijn ontwikkeld of de vraag in hoeverre de bezwaarbehandelaar de rechtmatigheid van de inzet van

⁴⁶ Zie de annotatie in ABRvS 17 mei 2017, ECLI:NL:RVS:2017:1259 (*AERIUS I*), *AB Klassiek* 2022/42, m.nt. C.J. Wolswinkel.

⁴⁷ Zie in dit verband ook de preconsultatie van het wetsvoorstel tot versterking van de waarborgfunctie van de Algemene wet bestuursrecht: [wetshttps://www.rijksoverheid.nl/documenten/kamerstukken/2023/01/20/aanbiedingsbrief-tk-toezending-ter-kennisname-teksten-pre-consultatie](https://www.rijksoverheid.nl/documenten/kamerstukken/2023/01/20/aanbiedingsbrief-tk-toezending-ter-kennisname-teksten-pre-consultatie).

⁴⁸ K.J. de Graaf and A.T. Marseille, 'Exit willekeurstoets. Bestuursrechtelijke toetsing aan het evenredigheidsbeginsel na 2-2-'22', *Ars Aequi* 2022(4), 306-314.

geautomatiseerde beslisregels als zodanig moet beoordelen. Ook hier zou de handreiking mogelijk meer richting kunnen bieden.

[Ondergeschikte rol van AVG] De handreiking leunt uiteindelijk sterker op de algemene beginselen van behoorlijk bestuur dan op de AVG. Weliswaar verwijst de handreiking naar de Uitvoeringswet AVG en de noodzaak om 'passende maatregelen' te nemen bij volledig geautomatiseerde besluitvorming, maar de Nederlandse wetgever heeft ervan afgezien die passende maatregelen verder te specificeren bij besluiten van bestuursorganen, omdat in die gevallen al 'de gebruikelijke en zeer vergelijkbare bestuursrechtelijke waarborgen' gelden zoals neergelegd in de Awb, waaronder het zorgvuldigheidsbeginsel, het motiveringsbeginsel en het evenredigheidsbeginsel, en omdat de betrokkene, zij het achteraf, altijd tegen een besluit kan opkomen (in bezwaar).⁴⁹ Deze oriëntatie in de handreiking op de algemene beginselen van behoorlijk bestuur is op zich te begrijpen, zeker nu deze beginselen ook van toepassing zijn in zaken waarin niet de verwerking van persoonsgegevens aan de orde is. Tegelijk zou ook de eis uit de AVG om de betrokkene «nuttige informatie» te verstrekken over «de onderliggende logica» van de verwerking alsmede over het belang en de verwachte gevolgen van de verwerking voor de betrokkene, invulling kunnen geven aan het motiveringsbeginsel, temeer nu de Raad van State de AVG als 'nuttige proeftuin' beschouwt voor de invulling van de algemene beginselen van behoorlijk bestuur.⁵⁰ Juist waar de handreiking nadrukkelijk ingaat op de juridische inbedding, zou hier ook aandacht moeten worden besteed aan de transparantieplichting uit de AVG. Nu gaat de handreiking echter uitsluitend in op het verbod op volledig geautomatiseerde besluitvorming zoals neergelegd in artikel 22 van de AVG, terwijl het richtinggevende karakter van die bepaling in Nederland juist beperkt is.⁵¹ De transparantieplichting uit de AVG kan echter een belangrijke rol bij de rechtsontwikkeling vervullen, zeker nu die verplichting mogelijk binnenkort nader wordt geïnterpreteerd door het Hof van Justitie van de Europese Unie wat betreft het concept 'nuttige informatie'. Ook het juridische kader van de AVG is nog nadrukkelijk in ontwikkeling, maar dat zou juist een reden kunnen zijn voor het beheerskader om proactief en ruimhartig in te zetten op uitgebreidere rechtsbescherming en begrijpelijke motivering van besluiten, waarmee de gemeente de rechtsontwikkeling positief kan beïnvloeden.

[Mens versus machine] Een ander risico is dat de aandacht in de handreiking voor motivering van het algoritme (onbedoeld) de aandacht afleidt van het belang van een deugdelijke motivering als zodanig van het besluit waartegen bezwaar is ingesteld, dus ook van de menselijke elementen in de besluitvorming (en het risico van bijvoorbeeld individuele of institutionele vooringenomenheid daarbinnen). Dit is mede van belang omdat het in steeds grotere mate illusoir is om de vele geautomatiseerde en 'handmatige' stappen van besluitvorming strikt te scheiden. Ook uitleg van beleid dat ten grondslag ligt aan een bepaald algoritme is belangrijk, alsmede uitleg over de interactie tussen mens en machine ('menselijke tussenkomst'). De handreiking moet dus geen afbreuk doen aan de andere eisen die bij de behandeling van bezwaar in acht moeten worden genomen. Integendeel, de behandelaar in bezwaar moet juist ook alert zijn op bezwaren die mogelijk niet specifiek verband houden met het algoritme, maar zich eigenlijk richten op het achterliggende, menselijk vastgestelde beleid.

[Bovenindividuele effecten van bezwaar] De handreiking is sterk gericht op de behandeling van een individueel bezwaar. Juist als het gaat om algoritmen, waarvan de (schadelijke) gevolgen soms pas zichtbaar worden door de veelvuldige toepassing hiervan, is een belangrijk ontbrekend

⁴⁹ *Kamerstukken II 2017/18*, 34 851, nr. 3, p. 121-122.

⁵⁰ *Digitalisering. Wetgeving en bestuursrechtspraak*, p. 66.

⁵¹ Zie paragraaf 2.2.1.

element in de handreiking de feedback die een individueel bezwaar oplevert voor het ontwikkelde algoritme: wat gebeurt er als bezwaar gegrond is verklaard? Wordt de uitkomst teruggekoppeld aan de ontwikkelaars van het algoritme? Wordt het aantal gegronde bezwaren bijgehouden? De handreiking geeft geen antwoord op dat proces van terugkoppeling.

[Toepassing van handreiking] De handreiking sluit nadrukkelijk aan op de terminologie die in bepaalde jurisprudentie is gehanteerd (zoals ‘maatwerkgegevens’ en ‘standaardinvoergegevens’ in het kader van stikstofberekeningen). Dat geeft de behandelaar in bezwaar mogelijk te weinig houvast in de zaken waarmee die te maken krijgt, omdat in die zaken andere terminologie wordt gehanteerd. Hoewel het goed is dat in de context van bezwaar aandacht wordt besteed aan algoritmische besluitvorming, is wel de vraag in hoeverre de behandelaar in bezwaar voldoende geëquipeerd is om vragen rond geautomatiseerde besluitvorming adequaat te behandelen, juist nu moet worden voorkomen dat de behandelaar enkel de handreiking volgt. De handelende ambtenaar krijgt weliswaar een aantal aanbevelingen mee, maar kan hij daadwerkelijk hiermee uit de voeten? Kan hij bijvoorbeeld de vraag beantwoorden of de uitkomst anders was geweest indien geen algoritme was gebruikt?

[Primaire besluitvorming] Naar aanleiding van deze handreiking rijst verder de vraag waarom niet al in de fase van primaire besluitvorming algoritmische transparantie wordt verzekerd. De handreiking zet in op dergelijke transparantie in het verweerschrift van het bestuursorgaan en in de beslissing op bezwaar, maar die transparantie zou eerder kunnen en moeten worden gerealiseerd. Nu wordt transparantie afhankelijk gemaakt van het instellen van bezwaar, hetgeen een grote drempel kan zijn voor burgers, met name voor burgers die zich in kwetsbare posities bevinden en zich bovendien geen rechtshulp kunnen veroorloven. Bovendien wordt van burgers verwacht dat ze voldoende inzicht hebben in het genomen besluit om bezwaargronden aan te voeren, maar daarvoor is juist een adequate motivering van het primaire besluit nodig. De nadruk op het zorgvuldigheidsbeginsel en het motiveringsbeginsel brengt mee dat die aandacht voor algoritmen al in de fase van de primaire besluitvorming zou moeten bestaan. Daarbij komt dat de AERIUS-jurisprudentie op bestuursorganen de verplichting legt om informatie over keuzes, gegevens en aannames, tijdig uit eigen beweging beschikbaar te stellen. Dit betekent dat ook eisen moeten worden gesteld aan de wijze waarop primaire besluiten worden genomen (bv. realiseren van menselijke tussenkomst,⁵² verstrekken van informatie, etcetera). Illustratief is in dit verband een recente uitspraak van de rechtbank Amsterdam waarbij een geautomatiseerde stap in een traject naar een aanvraag als primair besluit werd beschouwd. Als gevolg hiervan is deze geautomatiseerde beslissing onderworpen aan de eisen die de Algemene wet bestuursrecht aan het nemen van besluiten stelt, waaronder een deugdelijke motivering.⁵³

[Andere beslissingen dan bestuursrechtelijke besluiten] In het verlengde van de vorige bevinding rijst zelfs de vraag of deze waarborgen beperkt zouden moeten worden tot besluiten in bestuursrechtelijke zin (primair besluit of beslissing op bezwaar). De *SyRI*-uitspraak maakt duidelijk dat zelfs wanneer (nog) geen sprake is van een besluit in bestuursrechtelijke zin, maar van een feitelijke handeling (bv. de plaatsing van sensoren, de verzameling van bepaalde data, de opname van bepaalde data in een dataset), aanmerkelijke gevolgen voor burgers kunnen worden gecreëerd. Daarmee komt de AVG in beeld, alsmede het daaraan ten grondslag liggende transparantiebeginsel. Dat zou betekenen dat informatieverplichtingen van de overheid die primair in een bestuursrechtelijke context van bezwaar tegen een besluit in de zin van de Algemene wet

⁵² De impliciete veronderstelling is nu dat in bezwaar in elk geval menselijke tussenkomst is verzekerd.

⁵³ Rb. Amsterdam 17 mei 2022, ECLI:NL:RBAMS:2022:3066, AB 2023/12, m.nt. T. Barkhuysen en L.C. van Boven.

bestuursrecht zijn ontwikkeld, ook daarbuiten kunnen worden toegepast. Het Amsterdamse beheerskader waarborgt momenteel echter niet dat de punten die in bezwaar in acht moeten worden genomen, ook worden toegepast in situaties waarin geen sprake is van besluitvorming in bestuursrechtelijke zin. In die gevallen kan het algoritmeregister weliswaar enige transparantie bieden, maar zal daarnaast ook behoefte kunnen bestaan aan een meer geïndividualiseerde transparantie die uitlegt waarom in een concreet geval een bepaalde handeling is verricht.

Beoordeling

Het is positief dat in de Handreiking bezwaar het belang van ambtshalve aandacht in bezwaar voor geautomatiseerde besluitvorming wordt benadrukt. Bovendien wordt hiermee concreet invulling gegeven aan de menselijke tussenkomst die in bezwaar plaatsvindt. Tegelijk doet de beperking tot de fase van bezwaar de vraag rijzen welke beheersinstrumenten beschikbaar zijn voor de primaire besluitvorming en voor feitelijk handelen. Bij afwezigheid van andere instrumenten in het Amsterdamse beheerskader lijken die andere situaties een blinde vlek te vormen. Dat is geen tekortkoming van de handreiking zelf, maar wel van het beheerskader als geheel. Daarbij is het belangrijk dat de gemeente zich realiseert dat de juridische eisen die aan de kwaliteit van de motivering en aan andere vormen van transparantie worden gesteld, momenteel in beweging zijn, zodat eerder een ruimere dan een minimale bescherming op zijn plaats is.

3.2.5 Mensenrechtenimpactanalysemodel

Beschrijving

Een vijfde beheersinstrument betreft het mensenrechten impact assessment model (hierna: MRIA). De mensenrechtenimpactanalyse (inclusief bias-analyse)⁵⁴ wordt gepresenteerd als onderdeel van het 'Kunstmatige Intelligentie Impact Assessment' (KIIA). In dit impact assessment worden in vijftien categorieën vragen gesteld, onder meer over 'wettige en ethische toepassing' (onderdeel 7). In die specifieke categorie van 'wettige en ethische toepassing' wordt expliciet verwezen naar '(grond)rechten' en een bijbehorend template.

Bevindingen

[Beoordeling ex ante] Met het KIIA bevat het beheerskader een instrument waarmee niet alleen achteraf, maar juist ook vooraf de impact van de inzet van algoritmen kan worden beoordeeld. Zoals de Raad van State in haar publicatie *Digitalisering. Wetgeving en bestuursrechtspraak* opmerkte, vormt een impact assessment een mogelijke bron van informatie voor de bestuursrechter bij de toetsing van de zorgvuldigheid van algoritmische besluiten. De Raad van State houdt nadrukkelijk de mogelijkheid open dat het nalaten om een dergelijk impact assessment uit te voeren, nadelig uitpakt voor het bestuursorgaan dat zijn besluit voor de rechter komt verdedigen.⁵⁵

[Meer dan mensenrechten?] Hoewel positief aan het MRIA is dat aandacht bestaat voor de impact van het gebruik van algoritmen op mensenrechten en dat hierbij verder wordt gekeken dan enkel privacy,⁵⁶ is de MRIA momenteel zodanig ruim opgezet dat ze eigenlijk niet kan worden opgevat als een assessment tool enkel voor mensenrechten. De gehanteerde benaming 'Kunstmatige Intelligentie Impact Assessment' doet wat dit betreft meer recht aan de tool, hoewel de terminologie van 'kunstmatige intelligentie' (KI) afwijkt van de terminologie van 'algoritmen' die

⁵⁴ Zie hierna onder 3.2.6.

⁵⁵ *Digitalisering. Wetgeving en bestuursrechtspraak*, p. 79.

⁵⁶ Op grond van artikel 35 AVG kan bij de verwerking van persoonsgegevens reeds een *Data Protection Impact Assessment* (DPIA) verplicht zijn.

in andere beleidsdocumenten centraal staat. De impact assessment lijkt vooral neer te komen op een algemene *Algorithm Impact Assessment (AIA)*, waarvan een mensenrechtenanalyse en een biasanalyse deel uitmaken.

[Opzet] Mede gelet op de zeer diverse inhoud die aan bod komt in het KIIA, lijkt de gehanteerde vragenlijst een ad hoc verzameling waarvan de logische volgorde niet altijd duidelijk is. Het KIIA biedt bovendien weinig context en toelichting, waardoor begrippen uit de vragenlijst verkeerd of niet eenduidig kunnen worden geïnterpreteerd.⁵⁷ Door de grote verzameling aan vragen waarvan sommige een technisch detail betreffen en andere juist een hoofdzaak benoemen, is het instrument als geheel niet heel evenwichtig opgezet.⁵⁸ Ook bestaat specifiek onduidelijkheid op het onderdeel van de mensenrechten zelf. De evenredigheid van een beperking van grondrechten zal een belangrijk onderdeel van de analyse vormen, maar is in paragraaf 7 afwezig. Wel komt deze evenredigheid terug in paragraaf 15 onder de vragen 'Is de toepassing proportioneel?' en 'Is met minder ingrijpende middelen hetzelfde doel te verwezenlijken?', maar op dat moment staan die vragen buiten de context van mensenrechten.

[Toepassing] De MRIA zelf maakt niet duidelijk wanneer een MRIA wordt toegepast en onder wiens verantwoordelijkheid. Is het inderdaad bedoeld als een ex ante instrument dat wordt toegepast voorafgaand aan de inzet van algoritmen? Of heeft het ook betekenis gedurende de inzet van algoritmen? Verder is van belang voor welke gebruikersgroepen dit MRIA is bedoeld en hoe werkzaam dit MRIA voor deze personen is. Ook is onduidelijk waartoe de gegeven antwoorden moeten leiden: zijn die antwoorden enkel bedoeld als input voor een discussie of verplichten zij tot het nemen van bepaalde maatregelen bij het ontwikkelen of toepassen van het betreffende algoritme?⁵⁹ Tot slot is niet duidelijk of de uitkomsten van een MRIA ook worden gepubliceerd, bijvoorbeeld als onderdeel van het algoritmeregister.

[Andere mensenrechtenkaders] Inmiddels zijn ook andere mensenrechten impact assessment tools ontwikkeld, zoals het *Impact Assessment Mensenrechten en Algoritmes (IAMA)* dat in opdracht van het ministerie van BZK is ontwikkeld en een sterkere focus heeft op de mensenrechten dan de Amsterdamse tegenhanger. De vraag is daarom waarom de gemeente Amsterdam zou willen vasthouden aan een eigen MRIA. Daarvoor kunnen wellicht goede redenen zijn, maar die worden niet gegeven in het beheerskader zelf. Bovendien is met de huidige vormgeving van het MRIA (dat eigenlijk een veel ruimer KIIA is) eigenlijk geen sprake van een assessment dat specifiek betrekking heeft op mensenrechten. Bovendien laat het IAMA - anders dan het KIIA - nadrukkelijk zien waarvoor het uitvoeren van het assessment is bedoeld, namelijk het gezamenlijk in discussie gaan over te maken keuzes, het vastleggen van gemaakte keuzes en het functioneren als naslagwerk wanneer op basis van het algoritme beslissingen worden genomen. Tot slot beoogt het IAMA nadrukkelijk aan te sluiten bij andere documenten die in het kader van de omgang met algoritmen zijn ontwikkeld, zoals de Handreiking non-discriminatie by design en het toetsingskader van de Algemene Rekenkamer, en bovendien bij niet-technologisch ingestoken waarborgkaders. Die oriëntatie op andere richtsnoeren is afwezig in het KIIA van de gemeente Amsterdam.

⁵⁷ Zie bijvoorbeeld diverse begrippen bij de afweging om wel of geen assessment te doen, zoals 2.6 'gevoelig werkproces'; 2.8 'impact'; 2.11 'aanzienlijke mate treffen'.

⁵⁸ Zoals de procedurele vraag of een grondrechtenanalyse is uitgevoerd (7.6) versus de inhoudelijke vraag of door keuze-architectuur of de gebruikersinterface van de KI menselijk gedrag, meningen of beslissingen worden beïnvloed (7.9). Het antwoord op laatstgenoemde vraag zal ogenschijnlijk altijd 'ja' zijn. Het is daarom van belang om duidelijk te maken wat dit soort vragen beoogt te (laten) onderzoeken.

⁵⁹ Het bijgevoegde template maakt evenmin voldoende duidelijk hoe de weging van betrokken (grond)rechten werkt.

Beoordeling

De 'mensenrechtenimpactanalyse' heeft toegevoegde waarde binnen het Amsterdamse beheerskader voor algoritmen omdat zij een inhoudelijke tool biedt voor het *vooraf* beoordelen van algoritmen. In de huidige opzet is echter geen sprake van een assessment tool voor mensenrechten, maar eerder van een alomvattend *algorithm impact assessment*. De vraag is of het instrument in zijn huidige vorm voldoende richting en houvast biedt. Hoewel toepassing van dit instrument ertoe leidt dat vragen over de inzet van algoritmen worden beantwoord (en gedocumenteerd), is onvoldoende duidelijk waartoe de antwoorden vervolgens kunnen of moeten leiden (*follow-up*).

3.2.6 Biasanalysemodel

Beschrijving

Het Biasanalysemodel (BAM) is als onderdeel 8 opgenomen in de mensenrechtenimpactanalyse (ofwel het kunstmatige intelligentie impact assessment). Het doel van dit biasanalysemodel is om een aantal stappen te formuleren die moeten worden doorlopen om ontoelaatbare bias zoveel mogelijk uit het algoritme te weren. Het doorlopen van de verschillende stappen mondt uit in een zogeheten 'Rapport Bias Analyse', dat de oorspronkelijk gevonden en uiteindelijke biases omvat, alsmede eventuele maatregelen die genomen zijn en de overwegingen om een bias al dan niet aanvaardbaar te achten.

Bevindingen

[Ruim begrip van bias] Het KIIA besteedt afzonderlijk aandacht aan de analyse van bias. Het Biasanalysemodel vangt daartoe aan met een beschrijving van kenmerken waarop de gemeente geen ongerechtvaardigd onderscheid wil maken. Zelfs deze uitgebreide lijst is nog niet volledig (bv. opleiding, woonsituatie) en voor kritiek vatbaar (nationaliteit is niet altijd geschikt als proxy voor raciale kenmerken), terwijl de notie van intersectionaliteit (inhoudende dat juist de combinatie van bepaalde kenmerken problematisch kan zijn) ontbreekt. Tevens maakt het BAM niet duidelijk welke kenmerken in een concreet algoritme uiteindelijk worden geanalyseerd en aan de hand van welke maatstaf (*metrics*) de groepen met en zonder een bepaald kenmerk worden vergeleken. Het model maakt weliswaar expliciet dat die keuzes moeten worden gemaakt, maar biedt weinig houvast bij de keuze van de vergelijkingsmaatstaf (*metrics*) voor een specifiek algoritme. Verder maakt het model expliciet dat naast directe bias (ofwel bias op grond van een gevoelig kenmerk, zoals nationaliteit of geslacht) ook aandacht moet worden besteed aan indirecte bias, waarmee wordt bedoeld op bias die wordt veroorzaakt door het meenemen van een op het eerste gezicht onschadelijk kenmerk, maar die samenhangt met een wel gevoelig kenmerk. Tot slot maakt het model terecht duidelijk dat bias in zichzelf nog niet problematisch hoeft te zijn, maar dat uiteindelijk de vraag moet worden beantwoord of een bepaalde vorm van bias al dan niet aanvaardbaar is. Bij die uiteindelijke keuze wordt een aantal relevante perspectieven genoemd (business, juridisch, ethisch, tech). Een ontbrekend, maar voor een sociaal robuuste inzet belangrijk aanvullend perspectief is hier het perspectief van betrokken doelgroepen (dat zal moeten worden vastgelegd op basis van hun participatie).

[Opzet BAM] De opzet van het biasanalysemodel wijkt af van die van andere onderdelen van de KIIA. Waar die andere onderdelen worden gekenmerkt door uitgebreide vragenlijsten, biedt het onderdeel over bias meer context. Wel blijft een belangrijke vraag hoe, door wie en wanneer de bias-analyse wordt getoetst. Is dat enkel eenmalig aan de voorkant bij de ontwikkeling van het algoritme (waarbij nog verschillende fases als dataselectie en 'model training' kunnen worden

onderscheiden) of wordt die analyse gedurende de levenscyclus van het algoritme onderhouden, bijvoorbeeld om 'model drift' in het vizier te houden? Ook is de vraag of de uitkomst van de biasanalyse op enig moment wordt gepubliceerd.⁶⁰

De biasanalyse lijkt niet te zijn gekoppeld aan mensenrechtelijke noties van non-discriminatie, hetgeen op zich goed aansluit bij het fenomeen van bias: het gaat er niet om dat alle bias als zodanig wordt uitgesloten, maar enkel dat de verkeerde bias wordt geweerd. Tegelijk zal de biasanalyse dienstbaar moeten zijn aan het kunnen hanteren van het juridische non-discriminatiekader en idealiter zelfs aan een breder en inclusiever mensenrechtelijk beleid. Dit laatste is van belang, juist waar het geldende mensenrechtenkader tekortschiet in het licht van hoe algoritmen werken. Hier is bijvoorbeeld van belang dat de situatie van groepen die door een algoritme worden geraakt niet alleen voor, maar ook mét hen wordt onderzocht, ondersteund door vertegenwoordigers en experts.

[Negatieve, smalle benadering] Hierop aansluitend is het wellicht niet raadzaam om te werken met een 'negatieve' lijst van kenmerken als uitgangspunt voor de beoordeling van bias. Die benadering kan afleiden van de meer positieve insteek die een algoritme zou moeten hebben om gelijkwaardigheid en gelijke kansen te willen bevorderen. Dat raakt wederom aan de vraag of het BAM een minimum aan (juridische) bescherming wil bieden of juist verder wil gaan dan dat.

Beoordeling

Het biasanalysemodel vormt binnen het Amsterdamse beheerskader voor algoritmen een instrument dat zich specifiek richt op bias. Daarbij lijkt een ruime benadering van bias te worden gehanteerd die bias niet zonder meer vereenzelvigt met non-discriminatie. Dat is op zichzelf positief, maar dan is wel noodzakelijk dat dit instrument wordt ingebed in een bredere benadering waarin juist ook in gesprek met de betrokken doelgroepen wordt onderzocht of bias wordt ervaren. De belangrijkste vraag is hoe dit instrument wordt ingezet gedurende de volledige levenscyclus van het algoritme, omdat een toepassing enkel in de fase van ontwikkeling van het algoritme geen recht doet aan de (onvoorziene) effecten ten aanzien van bias die het gebruik van een algoritme in de praktijk kan hebben.

3.2.7 Risicoanalysemodel

Beschrijving

Het risicoanalysemodel beoogt aan de hand van vijftien 'risicostellingen' te komen tot een risicoanalyse voor algoritmische toepassingen in Amsterdam. Het model bestaat uit twee onderdelen. Eerst wordt op basis van een inventarisatie van voordelen, mogelijke risico's en andere afwegingen de vraag beantwoord in hoeverre de inzet van een algoritmisch systeem een 'zinvol idee' is. Deze 'afwegingscriteria' moeten worden ingezet in de onderzoeksfase van een project, dus voordat een algoritme wordt ontwikkeld. Daarna wordt een risicoanalyse toegepast op dit algoritmische systeem. Deze moet worden uitgevoerd wanneer daadwerkelijk wordt begonnen met de ontwikkeling van een algoritmisch systeem. Alvorens deze risicoanalyse toe te passen, wordt daartoe eerst nagegaan of sprake is van een algoritmische toepassing. Wanneer hiervan sprake is, presenteert de risicoanalyse op basis van de ja/nee-antwoorden die worden gegeven op de verschillende risicostellingen, risico-beheersmaatregelen die zijn afgestemd op de verschillende risicostellingen. Het model voorziet in een uitgebreide toelichting op de verschillende risicostellingen en bevat een verklarende woordenlijst (onder meer 'algoritmische toepassing',

⁶⁰ In het algoritmeregister is wel een item dat betrekking heeft op non-discriminatie.

‘schade’). Het model verwijst tevens naar het document 'Taken en verantwoordelijkheden bij het gebruik van Algoritmen door de gemeente Amsterdam' voor een overzicht van alle beheersmaatregelen en verantwoordelijkheden.

Bevindingen

[Afwegingscriteria] Het eerste deel van het risicoanalysemodel ('afwegingscriteria') ziet op de vraag of de inzet van een algoritme sowieso een zinvol idee is en gaat daarmee vooraf aan de eigenlijke risicoanalyse. Binnen dit eerste deel wordt gevraagd om een inventarisatie van voordelen van het gebruik van een algoritme. De *quick scan* van risico's richt zich specifiek op de verboden lijst van AI-toepassingen uit het Commissievoorstel voor een AI-Verordening en op de zes principes voor verantwoorde digitalisering van het openbaar bestuur uit de Code Goed Digitaal Openbaar Bestuur (CODIO). Deze zes principes en hun relatie tot risico's zijn nog tamelijk algemeen verwoord (zeker in verhouding tot de risicoanalyse hierna). Zo wordt het belang om belanghebbenden en gebruikers te betrekken bij de ontwikkeling van het algoritme weliswaar genoemd, maar verder niet uitgewerkt. Ook de overige afwegingen zijn tamelijk generiek geformuleerd en dragen het risico in zich dat de principekeuze voor een algoritmisch systeem feitelijk al is gemaakt, namelijk wanneer 'het project' reeds wordt opgevat als een algoritmisch project waaraan het alternatief (geen algoritme) vervolgens niet kan tippen. Wat nadrukkelijk ontbreekt in dit eerste onderdeel is een handvat (kader) om een afweging te maken tussen de verschillende afwegingscriteria die zijn geïdentificeerd. Zo zijn belangrijke vragen als 'wat is de noodzaak?' en 'zijn er ook andere manieren om deze voordelen te bereiken?' afwezig.

[Risicoanalyse: reikwijdte] Wat betreft het toepassingsbereik maakt het risicoanalysemodel in de 'checks vooraf' een aantal bediscussieerbare keuzes. Allereerst lijkt de tool zodanig ingericht dat enkel wanneer sprake is van een 'hoog risico'-toepassing in de zin van het Commissievoorstel voor een AI-Verordening, de risicoanalyse moet worden toegepast. Dat lijkt een behoorlijke nauwe opvatting van de risico's die zich kunnen voordoen bij de inzet van algoritmen. Daarnaast stelt de toelichting op de tool dat wanneer het algoritmische systeem geen vorm van kunstmatige intelligentie bevat, bijvoorbeeld als het systeem alleen een selectie maakt op basis van vooraf door experts vastgestelde criteria, bepaalde risicobeheersingsmaatregelen juist niet nodig zijn. Niet alleen krijgt deze stellingname ogenschijnlijk geen vervolg in de tool, maar ook roept deze stellingname de vraag op wat precies wordt verstaan onder kunstmatige intelligentie (enkel *machine learning* of ook expertsystemen?) en welke maatregel om welke reden bij andere algoritmische systemen niet hoeft te worden toegepast. Het voorbeeld van de door experts vastgestelde criteria roept namelijk de vraag op hoe de risicoanalyse is ingebed in onderliggende beleidsmatige keuzes en in hoeverre risico's bij het maken van die onderliggende keuzes onder ogen zijn gezien.

[Risicostellingen] Wat betreft de risicostellingen is positief dat deze worden gekoppeld aan acties van verantwoordelijke instanties. De risicostellingen zijn echter zeker niet volledig. Zo ontbreken onder meer risicostellingen over afhankelijkheden bij de ontwikkeling van een systeem, het gebrek aan capaciteit bij de ontwikkeling van het systeem, onbedoelde gevolgen (zoals sabotage en *gaming the system*) en de afbreuk die een algoritmisch systeem kan doen aan het onderliggende beleidsdoel.

[Risicoanalyse: beheersmaatregelen] Het risicoanalysemodel is een behoorlijk uitgewerkte tool die beoogt een aantal risico's rond de inzet van algoritmen te identificeren en, waar nodig, risico-beheersmaatregelen aan te dragen. Deze koppeling tussen geconstateerde risico's en te nemen beheersmaatregelen is positief, omdat de risicoanalyse daarmee niet alleen neerkomt op het

beantwoorden van een vragenlijst (zoals het KIIA lijkt te doen), maar ook aanzet tot actie. Daarbij komt dat het risicoanalysemodel ook nadrukkelijk houvast beoogt te bieden aan degene die de tool hanteert, onder meer door de verschillende risicostellingen toe te lichten en van voorbeelden te voorzien en door een aantal kernbegrippen in een afzonderlijke legenda van een toelichting te voorzien. Die legenda zou nog verder kunnen worden uitgewerkt (bijvoorbeeld met schade als gevolg van gebrekkige beveiliging, zoals identiteitsfraude) en mogelijk ook kunnen worden ontsloten via het publiek toegankelijke algoritmeregister.

[Toepassing] Een belangrijke vraag voor de werking van dit instrument is wie het instrument op welk moment toepast. Blijkens de toelichting moet de risicoanalyse worden toegepast bij de aanvang van de ontwikkeling van een algoritmisch systeem. De tool zelf maakt niet duidelijk wie het instrument toepast, hoewel de *governance*-beschrijving (zie 3.3.2) suggereert dat de directeur vooraf een risicoanalyse moet uitvoeren. Evenmin maakt de tool zelf duidelijk hoe de uitkomsten van de risicoanalyse worden vastgelegd en hoe die een vervolg krijgen. Hoewel het model de verantwoordelijkheid voor de te treffen risico-beheersmaatregelen beoogt expliciet te maken, wordt die verantwoordelijkheid steeds slechts in algemene zin bij 'ontwikkelteam en stakeholders' gelegd.

Beoordeling

Het risicoanalysemodel is een concreet instrument dat in vergelijking met de andere instrumenten uit het Amsterdamse beheerskader voor algoritmen veel houvast biedt, onder meer door middel van een toelichting op de belangrijkste begrippen, een verduidelijking van de risicostellingen door middel van voorbeelden en een koppeling van antwoorden op risicostellingen aan te treffen risico-beheersmaatregelen. Het is minder duidelijk waarom dit model enkel betrekking heeft op zogeheten 'hoog risico'-toepassingen, omdat daarmee het risico eigenlijk al is gekwalificeerd voordat de risicostellingen zijn beantwoord.

3.3 Beheersinstrumenten in onderlinge samenhang

Het Amsterdamse beheerskader voor algoritmen beoogt een aantal instrumenten aan te reiken die gezamenlijk leiden tot de beheersing van risico's bij de ontwikkeling en het gebruik van algoritmen. Daarbij wordt uitgegaan van een levenscyclus, die nadrukkelijk centraal staat in de *governance*-beschrijving. Deze aandacht voor de verschillende facetten van de ontwikkeling en het gebruik van algoritmen door de levenscyclus heen is waardevol voor een integrale beheersing van risico's bij het gebruik van algoritmen, aangezien die risico's zich op verschillende momenten in de levenscyclus kunnen verwezenlijken.

[Samenhang] Wanneer naar de onderlinge samenhang van de verschillende beheersinstrumenten wordt gekeken, vervult de *governance*-beschrijving een centrale rol. Deze beschrijving bevat namelijk de verschillende taken en verantwoordelijkheden in de verschillende fasen van de levenscyclus van een algoritme. Ten aanzien van twee onderdelen in die levenscyclus, namelijk het uitbesteden van de ontwikkeling van een algoritme en het behandelen van een bezwaar, bevat het beheerskader instrumenten die meer specifieke en inhoudelijke houvast beogen te bieden. Die focus op bepaalde onderdelen in de levenscyclus roept direct de vraag op waarom voor andere onderdelen minder aandacht lijkt te bestaan. Dat geldt in het bijzonder voor de fase van primaire besluitvorming, omdat juist ook in die fase eisen aan de zorgvuldige voorbereiding en de motivering van de beslissing kunnen worden gesteld. Verder bevat het beheerskader drie assessment tools waarvan de onderlinge verhouding niet duidelijk is. De mensenrechtenanalyse wordt gepresenteerd als een 'kunstmatige intelligentie impact assessment', waarvan tevens een biasanalyse deel uitmaakt. Naast dit kunstmatige intelligentie

impact assessment is er echter ook een risicoanalysemodel dat moet worden toegepast voordat een algoritme wordt ontwikkeld. Het is onduidelijk hoe die verschillende tools op elkaar ingrijpen.

[Algoritme] De verschillende beheersinstrumenten roepen in onderlinge samenhang ook de vraag op wat nu precies wordt verstaan onder een ‘algoritme’ of een ‘algoritmische toepassing’ in de Amsterdamse context. Hoewel over de verschillende instrumenten heen met een vergelijkbare definitie wordt gewerkt, wordt in bepaalde gevallen de toevoeging gehanteerd dat de algoritmische toepassing de Amsterdamse burger ‘in aanmerkelijke mate’ moet treffen (*governance*-beschrijving, contractvoorwaarden, risicoanalysemodel), terwijl die toevoeging bij andere instrumenten ontbreekt (algoritmeregister). Tegelijk wordt in verschillende instrumenten (mensenrechtenimpactanalysemodel, risicoanalysemodel) ook de term KI-systeem gehanteerd, zonder dat duidelijk wordt of hiermee een ander begrip is beoogd dan algoritmisch systeem.

[Oriëntatie op AI-Verordening] Wat verder opvalt in de verschillende beheersinstrumenten, is de sterke oriëntatie op het voorstel van de Europese Commissie voor een AI-Verordening. Hoewel deze verordening nog niet is aangenomen, lijkt het Amsterdamse beheerskader het Commissievoorstel reeds als geldend recht te beschouwen. Die oriëntatie op het Commissievoorstel is begrijpelijk, omdat dit voorstel een van de eerste documenten betreft met voorschriften die specifiek zijn toegesneden op AI-systemen. Tegelijk bevat die benadering ook kwetsbaarheden. Allereerst is het Commissievoorstel nog aan verandering onderhevig (zoals bijvoorbeeld het begrip ‘AI-systeem’), zodat het Amsterdamse beheerskader voor algoritmen voortdurend hierop moet worden aangepast (en evenmin volkomen zeker is dat Commissievoorstel ongewijzigd de eindstreep zal halen).⁶¹ Bovendien is ook vanuit mensenrechtelijke en *civil society* hoek gefundeerde kritiek geleverd op het voorstel.⁶² Tot slot is de vraag of het verstandig is om zo nadrukkelijk aan te sluiten bij het Commissievoorstel voor de doelen die het Amsterdamse beheerskader nastreeft. Zo lijkt het risicoanalysemodel enkel een risicoanalyse nodig te achten bij ‘hoog risico’ AI-systemen in de zin van het Commissievoorstel, terwijl ook AI-systemen zonder dergelijk ‘hoog risico’ (aanzienlijke) impact kunnen hebben op het leven van burgers. Bovendien is de vraag of het begrip ‘AI-systeem’ zoals dat wordt gehanteerd in de voorgestelde AI-Verordening, eenzelfde bereik heeft als het begrip ‘algoritme’ binnen (de verschillende instrumenten van) het Amsterdamse beheerskader.

[Ontbrekende schakels] Sowieso blijven binnen het Amsterdamse beheerskader bepaalde schakels in de levenscyclus momenteel onderbelicht. Dat betreft onder meer de aandacht voor de primaire besluitvorming en voor handelingen die niet (direct) rechtsgevolg hebben. Verder is in de verschillende fasen van de levenscyclus nog weinig concrete aandacht zichtbaar voor de noodzaak van participatie van burgers en specifiek van vertegenwoordigers van verschillende doelgroepen (*‘affected groups’*), terwijl juist de afwezigheid van die burger- of belangengroep participatie voor het gemeentebestuur een *red flag* zou kunnen zijn bij de inzet van algoritmen. Voorts is de aandacht voor aspecten als terugkoppeling, evaluatie en *feedback loops* nog tamelijk betrekkelijk in verhouding tot de aandacht voor de fase van de ontwikkeling van een algoritme. Een meer algemene observatie is dat het Amsterdamse beheerskader als geheel genomen een veelheid aan (op zich belangrijke) vragen en overwegingen bevat die op verschillende plaatsen in verschillende bewoordingen met verschillende nuances staan. Die veelheid aan overwegingen kan ertoe leiden dat het zicht op de belangrijkste aspecten van het

⁶¹ Zo heeft de Raad van Ministers onlangs overeenstemming bereikt over een zogeheten ‘algemene benadering’ voor de onderhandelingen met het Europees Parlement over de voorgestelde AI-Verordening. Zie <https://data.consilium.europa.eu/doc/document/ST-14954-2022-INIT/nl/pdf>.

⁶² Zie bijvoorbeeld <https://algorithmwatch.org/en/response-to-eu-ai-regulation-proposal-2021> en <https://edri.org/our-work/the-eus-artificial-intelligence-act-civil-society-amendments/>

beheerskader vervaagt. In het bijzonder zou de vraag meer centraal moeten komen te staan welke publieke belangen met de inzet van algoritmen worden gediend en hoe het publieke belang van het voorkomen van schade zich verhoudt tot andere publieke belangen (zoals het realiseren van publieke veiligheid). Het al dan niet bijdragen aan die gearticuleerde publieke belangen kan namelijk als rode draad dienen bij de beoordeling van het algoritme vanuit de verschillende beheersinstrumenten.

[Verhouding tot andere beheersinstrumenten] De instrumenten van het Amsterdamse beheerskader zijn ontwikkeld op een moment dat nog weinig vergelijkbare instrumenten voorhanden waren. Inmiddels is die situatie veranderd. Zo is op landelijk niveau een mensenrechtenimpactassessment (IAMA) ontwikkeld, terwijl momenteel ook het nationaal algoritmeregister wordt uitgerold (algoritme.overheid.nl). Dat roept de vraag op hoe de instrumenten uit het Amsterdamse beheerskader functioneren ten opzichte van die andere instrumenten. Het antwoord op die vraag kan aanleiding zijn om aansluiting te zoeken bij die andere instrumenten of juist de toegevoegde waarde van de eigen instrumenten te benadrukken. Dat hangt ook samen met het ethische perspectief en het antwoord op de vraag waar Amsterdam (voor) wil staan als het gaat om de inzet van algoritmen (zoals de Tada-waarden).⁶³

[Toepassing] De toereikendheid van het Amsterdamse beheerskader voor algoritmen als geheel hangt mede af van de wijze waarop de verschillende instrumenten worden ingezet. Daarvoor is onder meer van belang welke kennis, expertise en capaciteit binnen de ambtelijke organisatie beschikbaar zijn om de verschillende instrumenten succesvol in te zetten. Het beheerskader zelf geeft geen antwoord op die vraag, maar de 'werkbaarheid' van het instrumentarium in de praktijk zal mede bepalen of de inzet van algoritmen toereikend kan worden beheerst.

⁶³ <https://www.amsterdam.nl/innovatie/digitalisering-technologie/data/tada-waarden/>

4. Conclusies en aanbevelingen

4.1 Conclusies

Centraal in dit rapport staat de vraag in hoeverre het Amsterdamse beheerskader voor algoritmen toereikend is voor een verantwoorde toepassing van algoritmen. Deze vraag laat zich niet eenvoudig beantwoorden, niet alleen omdat de begrippen ‘toereikend’ en ‘verantwoord’ nadere invulling verdienen, maar ook omdat inzichten over en ervaringen met de beheersing van de inzet van steeds verder evoluerende algoritmen nog voortdurend in ontwikkeling zijn. Niettemin is het op basis van het gecombineerde perspectief van de privacy, de mensenrechten, het bestuursrecht en de ethiek wel mogelijk een aantal belangrijke gezichtspunten te formuleren die sowieso onder ogen moeten worden gezien bij de beheersing van algoritmen door middel van het Amsterdamse beheerskader. Omdat deze beoordeling enkel is gebaseerd op documentanalyse van het Amsterdamse beheerskader, ontbreekt in dit rapport achterliggende informatie waarom bepaalde keuzes zijn gemaakt binnen het beheerskader.

Het Amsterdamse beheerskader voor algoritmen bestaat momenteel uit zeven instrumenten: het algoritmeregister, de *governance*-beschrijving, modelcontractvoorwaarden, de handreiking bezwaar, een mensenrechtenimpactanalysemodel, een biasanalysemodel en een risicoanalysemodel. Gezamenlijk adresseren deze instrumenten verschillende facetten van de ontwikkeling en het gebruik van algoritmen, zowel procedureel als inhoudelijk. Daarmee vormen deze instrumenten nuttige en waardevolle bouwstenen om de risico's van de inzet van algoritmen gedurende hun volledige levenscyclus te beheersen. Tegelijk zijn de verschillende instrumenten nog duidelijk in ontwikkeling. Zo lijkt het algoritmeregister als beheersinstrument nog geen prominente rol te vervullen in de ontwikkelfase van het algoritme, terwijl de *kunstmatige intelligentie impact assessment* (waarvan de mensenrechtenimpactanalyse en de biasanalyse deel uitmaken) meer richting zou kunnen bieden ten aanzien van de condities waaronder een algoritme wel of niet kan worden ingezet.

Het huidige Amsterdamse beheerskader voor algoritmen is momenteel nog duidelijk een lappendeken. Dit betekent dat niet alleen kritisch moet worden gekeken naar de inhoud en werking van de instrumenten afzonderlijk, maar ook naar hun onderlinge samenhang. Zo valt op dat centrale begrippen in de verschillende instrumenten (zoals ‘algoritme’) niet altijd naadloos op elkaar aansluiten. Verder keren de diverse instrumenten van het Amsterdamse beheerskader nog weinig duidelijk terug in de *governance*-beschrijving en is de verhouding tussen de risicoanalyse en de *kunstmatige intelligentie impact assessment* onvoldoende duidelijk. Ook laat het overkoepelende beeld van de lappendeken zien waar nog gaten (lacunes) zitten in het beheerskader. Zo wordt veel aandacht besteed aan de omgang met geautomatiseerde besluitvorming in bezwaar, maar ontbreekt een vergelijkbaar instrument voor de omgang met algoritmen in de fase van de primaire besluitvorming of bij feitelijk handelen. Ook blijkt het element van burger- of belangengroep participatie onvoldoende geborgd in het beheerskader als geheel, terwijl dat cruciaal is voor een sociaal robuuste, maatschappelijk verantwoorde inzet van algoritmen.

De aandacht van het beheerskader voor een verantwoorde inzet van algoritmen lijkt zeker niet beperkt tot privacygerelateerde overwegingen. Tegelijk wordt voor de toepassing van deze instrumenten onder meer gekeken naar *privacy officers* en adviesorganen op het terrein van persoonsgegevens, waardoor in de praktijk het privacy-perspectief toch onbedoeld dominant zou kunnen worden (of de andere perspectieven minder uit de verf komen). Het beheerskader zelf lijkt ogenschijnlijk in te zetten op een breder perspectief van impact van algoritmen op mensenrechten.

Tegelijk komt die mensenrechtenanalyse uiteindelijk neer op een impact assessment die veel verder reikt dan enkel mensenrechten, terwijl cruciale overwegingen rond noodzakelijkheid en proportionaliteit van beperkingen van mensenrechten onderbelicht zijn. Andere beheersinstrumenten (contractvoorwaarden, handreiking bezwaar) nemen nadrukkelijk het bestuursrecht (mede) tot uitgangspunt. Ongeacht het perspectief dat wordt gekozen, geldt steeds als intrinsieke beperking dat zodra op een bepaald perspectief de nadruk wordt gelegd, andere perspectieven uit het oog kunnen worden verloren. Voor het Amsterdamse beheerskader is met name van belang dat het ethische perspectief niet uit het oog wordt verloren. Het is immers niet alleen van belang dat het beheerskader leidt tot overeenstemming met de geldende, maar tegelijk voortdurend in ontwikkeling zijnde wet- en regelgeving (privacy, mensenrechten, bestuursrecht), maar ook dat het Amsterdamse beheerskader uitdrukking geeft aan waar de gemeente Amsterdam (voor) staat bij de inzet van algoritmen, ongeacht waartoe zij juridisch verplicht is. Daarbij is met name van belang dat het beheerskader zich niet beperkt tot het 'intern' organiseren van overeenstemming met geldende regels en richtsnoeren, maar zich ook richt op het 'extern' betrekken van betrokken doelgroepen bij de ontwikkeling en inzet van algoritmen, hoe lastig dit politiek-ethische perspectief soms ook is te verwezenlijken in de praktijk. Zulke investeringen dragen direct bij aan een verantwoorde doorontwikkeling van de juridische en ethische kaders, waarmee die op termijn meer houvast kunnen bieden aan alle betrokkenen.

Uiteindelijk hangt het vermogen van het Amsterdamse beheerskader voor algoritmen om een verantwoorde inzet van algoritmen te bewerkstelligen mede af van de toepassing van dit beheerskader. Zelfs wanneer de bestaande instrumenten worden doorontwikkeld en de afstemming tussen de verschillende instrumenten wordt versterkt, geldt dat het moment en de wijze waarop de instrumenten worden toegepast mede zullen bepalen of het beheerskader daadwerkelijk kan bijdragen aan een verantwoorde inzet van algoritmen. In dat opzicht verdient niet alleen het beheerskader zelf voortdurende aandacht, maar ook de omgang hiermee binnen de gemeentelijke organisatie en de context waarbinnen algoritmen worden ingezet, vaak als schakels in een keten van beslissingen.

4.2 Aanbevelingen

- 1) Stel binnen het beheerskader de publieke belangen waaraan het algoritme zou moeten bijdragen centraal, zodat het kader niet enkel in negatieve zin is gericht op het beperken van risico's of schade, maar ook in positieve zin op het verwezenlijken van publieke waarden.
- 2) Zorg voor meer stroomlijning tussen de verschillende beheersinstrumenten (niet alleen terminologisch, maar ook inhoudelijk) en voor een duidelijke inbedding van de verschillende beheersinstrumenten in de *governance*-beschrijving.
- 3) Kies beargumenteerd ofwel voor een eenduidig begrip van 'algoritme' binnen het beheerskader ofwel voor variatie afhankelijk van het toepasselijke instrument.
- 4) Geef extra aandacht aan de blinde vlekken of onderontwikkelde punten binnen het huidige beheerskader (zoals het realiseren van burgerparticipatie en vertegenwoordiging van betrokken doelgroepen, het gebruik van algoritmen in de fase van primaire besluitvorming of bij feitelijk overheidshandelen, de inrichting van feedbackmechanismen).
- 5) Onderzoek in hoeverre instrumenten die buiten Amsterdam zijn of worden ontwikkeld, zoals het IAMA, zinvol kunnen worden ingezet binnen de Amsterdamse context.
- 6) Blijf aandacht schenken aan de context waarbinnen een algoritme wordt ingezet, zeker ook waar menselijk handelen een belangrijke rol vervult naast het algoritme.
- 7) Monitor de werking van de verschillende beheersinstrumenten in de gemeentelijke praktijk en investeer in de doorontwikkeling van.