

onderzoeksopzet

vervolgonderzoek algoritmes

december 2022



1 inleiding

1-1 achtergrond

In april 2021 heeft de rekenkamer het rapport 'Gekleurde technologie' gepubliceerd.¹ Dit onderzoek liet zien dat de gemeente Rotterdam op dat moment onvoldoende zicht had op de aard en het gebruik van algoritmes binnen de organisatie. Daarnaast werd geconcludeerd dat de gemeente te weinig aandacht besteedde aan het beheersen van ethische risico's bij de ontwikkeling en uitvoering van algoritmes. Zo bleek onder meer dat het cluster Werk en Inkomen (W&I) gebruik maakte van een algoritme dat mogelijke proxy-variabelen voor etniciteit bevatte. In het rapport deed de rekenkamer vervolgens een aantal aanbevelingen aan het gemeentebestuur rondom de toepassing van algoritmes en het beheersen van ethische risico's.

In navolging van het rekenkamerrapport werd besloten om het algoritme van W&I buiten werking te stellen en pas weer in te zetten op het moment dat het aan de kwaliteitseisen voldoet die door de raad zijn gesteld.² Ook werden diverse moties rondom het thema algoritmes aangenomen. Eén van deze moties was 'Democracy is watching Big Brother' waarin de rekenkamer werd verzocht om een vervolgonderzoek naar algoritmes uit te voeren.

Motie: Democracy is watching Big Brother

Constaterende dat:

- Uit het Rekenkamerrapport 'Gekleurde technologie; verkenning ethisch gebruik algoritmes' blijkt dat er nog een hoop moet gebeuren op het terrein van publieke controle op algoritmes die gebruikt worden door de gemeente;
- De casus van het algoritme bij W&I aantoonde dat de gemeente 4 jaar lang met een algoritme heeft gewerkt dat de ethische toetsing niet zou doorstaan.

Verzoekt het college:

- De Rekenkamer te verzoeken om in het begin van de volgende collegeperiode een vervolgonderzoek te doen naar de toepassing van algoritmes bij de gemeente Rotterdam.³

Het afgelopen jaar heeft de gemeente een governancestructuur uitgewerkt om een verantwoorde toepassing van algoritmes binnen de organisatie te borgen.⁴ Onderdeel hiervan is onder andere een inventarisatie van hoog-risico algoritmes, het ontwikkelen van een algoritmeregister, het beleggen van de functie van algoritme-expert en het samenstellen van een Algorithm Advisory Board (AAB).

In juni 2022 concludeerde het college dat de governance op hoofdlijnen staat, maar

¹ Rekenkamer Rotterdam, 'Gekleurde technologie. Verkenning ethisch gebruik algoritmes', 14 april 2021.

² Wethouder Werk, Inkomen en Nationaal Programma Rotterdam-Zuid, 'Rechtmatige vestrekking uitkering', 23 juni 2021 en Werk en Inkomen, Monitor Werk & Inkomen, 2e tertaal 2022', 18 oktober 2022.

³ Relevante delen uit de tekst motie Gemeenteraad Rotterdam, 'Democracy is watching Big Brother', 30 september 2021.

⁴ Burgemeester en Wethouders van Rotterdam, 'Governance rondom algoritmes', 7 september 2021 en Burgemeester en Wethouders van Rotterdam, 'Stand van zaken algoritmes', 15 februari 2022 en Burgemeester en Wethouders van Rotterdam, 'Algoritmegovernance', 7 juni 2022.

onder de aandacht blijft in het licht van mogelijke nieuwe ontwikkelingen en richtlijnen.⁵

Niet alleen binnen de gemeente Rotterdam, maar ook op Europees en landelijk niveau en in andere grote steden is er in toenemende mate aandacht voor de risico's van de inzet van algoritmes door overheden en het beleid hieromtrent. Dit gegeven in combinatie met de implementatie van het governancekader maakt dat de rekenkamer het van belang acht om, zoals verzocht door de raad, vervolgonderzoek te doen naar de inzet van algoritmes binnen de gemeente Rotterdam.

Met dit onderzoek beoogt de rekenkamer antwoord te geven op de vraag in hoeverre de governance van de gemeente Rotterdam toereikend is voor een verantwoorde toepassing van algoritmes.

1-2 leeswijzer

In hoofdstuk 2 wordt de context van het onderzoek beschreven en wordt ingegaan op de definitie van algoritmes, de risico's rondom het gebruik van algoritmes en de beleidscontext. Vervolgens worden in hoofdstuk 3 de doel- en vraagstelling en afbakening van het onderzoek toegelicht. Hoofdstuk 4 beschrijft de onderzoeks aanpak en in hoofdstuk 5 volgt de organisatie en planning van het onderzoek.

⁵ Burgemeester en Wethouders van Rotterdam, 'Algoritmegovernance', 7 juni 2022.

2 context

2-1 typering en werking van algoritmes

Een algoritme is een set instructies die wordt uitgevoerd om een bepaald doel te bereiken. Volgens deze definitie is een autobestuurder die een routebeschrijving volgt in principe ook een algoritme aan het uitvoeren. In het publieke debat wordt er echter meestal aan algoritmes gerefereerd als deze door een computer worden uitgevoerd. Een computer analyseert volgens een systematische procedure verschillende data om uiteindelijk een conclusie te presenteren aan de gebruiker van het algoritme. Een voorbeeld hiervan is een computer die, gebaseerd op de zoekgeschiedenis, locatie en leeftijd van de gebruiker, een bepaalde advertentie laat zien in de webbrowser.

wat is een algoritme?

Een algoritme is, kort gezegd, een set instructies die leidt tot een beoogd doel. Echter, in deze definitie kan bijvoorbeeld een kookrecept of een beslisboom ook gezien worden als een algoritme.

De rekenkamer hanteert de volgende definitie: een algoritme is een set instructies, die *door een computer* wordt uitgevoerd, om te komen tot een beslissing of om te komen tot informatie die gebruikt wordt om een beslissing te nemen.

Er zijn verschillende manieren om algoritmes te typeren. Allereerst is er een verschil tussen rule-based en case-based algoritmes. *Rule-based algoritmes* zijn rekenregels die door mensen zijn bedacht en door een computersysteem worden toegepast. Zo kan bijvoorbeeld uit wetenschappelijk onderzoek blijken dat huishoudens die vaker dan vijf keer verhuizen in tien jaar, een hogere kans hebben om fraude te plegen met hun creditcard. Een programmeur kan een systeem dan zo programmeren, dat er een melding komt wanneer een klant van een bank meer dan vijf keer verhuist in tien jaar. *Case-based algoritmes* zijn zogenaamde zelflerende algoritmes. Het algoritme leert op basis van bekende (historische) informatie patronen te herkennen. Deze patronen past het toe op nieuwe informatie om deze bijvoorbeeld in te delen in een categorie of zoals in bovenstaand voorbeeld: om een kans op creditcardfraude te voorspellen bij nieuwe klanten.⁶

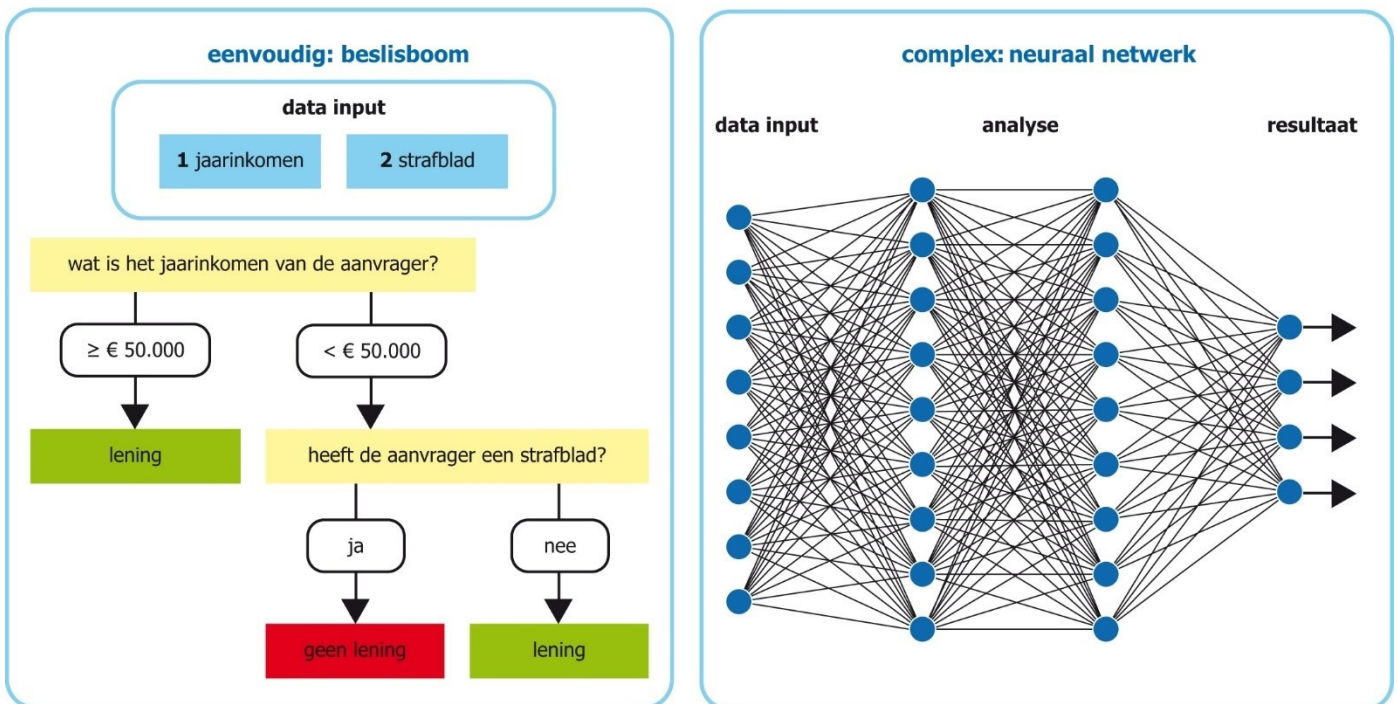
Algoritmes kunnen ook worden ingedeeld naar complexiteit. Een beslisboom heeft een beperkt aantal variabelen en doorloopt voor de mens goed te begrijpen stappen. De beslisboom in figuur 2-1 is een versimpelde weergave van een algoritme dat een bank kan gebruiken om gebaseerd op twee inputvariabelen (jaarinkomen en het hebben van een strafblad) en vooropgezette instructies ($< € 50.000$ of $\geq € 50.000$ en wel of niet een strafblad) te bepalen welke aanvrager een lening krijgt en welke aanvrager geen lening krijgt. Dit is een voorbeeld van een eenvoudig algoritme waarvan het proces en de uitkomst goed te interpreteren zijn.

Complexere algoritmes zijn gebaseerd op zeer ingewikkelde rekenregels. Dit zijn in de regel zelflerende algoritmes. Een voorbeeld hiervan is het neurale netwerk in figuur 2-1. Een neuraal netwerk is een type algoritme dat lijkt op de manier waarop hersenen werken. In deze gevallen is het redeneerproces van het algoritme haast onmogelijk te

⁶ CBS, 'Verkenkend onderzoek naar het gebruik van algoritmen binnen overheidsorganisaties', november 2018.

begrijpen voor de mens. Door de hoge technische complexiteit worden deze algoritmes ook wel een 'black box' genoemd. Complexere algoritmes zijn vaak krachtiger. Een complex algoritme zoals een neuraal netwerk zal daarom sneller ingezet worden bij ingewikkelde taken, zoals het herkennen van afbeeldingen. Tijdens het ontwikkelen van een algoritme zal dus een afweging gemaakt moeten worden tussen de interpreteerbaarheid en de rekenkracht en accuraatheid van het algoritme.

figuur 2-1: eenvoudig algoritme en complex algoritme



Algoritmes worden in toenemende mate ingezet, ook bij (lokale) overheden. Door een algoritme te gebruiken, kunnen eenvoudige taken geautomatiseerd worden uitgevoerd zonder tussenkomst van een ambtenaar. Algoritmes kunnen ook worden ingezet om handhavingscapaciteit beter te benutten. Door een algoritme op basis van data risico's te laten inschatten, kunnen controles worden uitgevoerd op de plek of bij de persoon waar het risico het grootst is. Dit laatste was bijvoorbeeld het geval bij het algoritme van W&I waar de rekenkamer in het vorige onderzoek uitgebreid bij stil heeft gestaan.

2-2 risico's van algoritmes

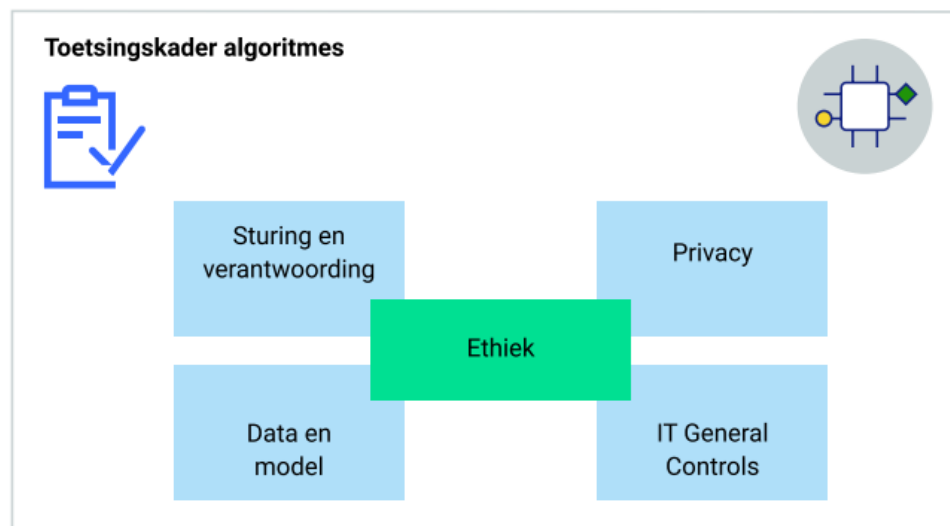
Bij het gebruik van algoritmes kunnen risico's optreden. Er is een kans op onbedoelde effecten (zoals discriminatie) doordat het lastig navolgbaar kan zijn hoe de data naar een uitkomst leidt. Het risico bestaat dat het proces voor betrokkenen intransparant is. De complexiteit van modellen kan ervoor zorgen dat een algoritme moeilijk uitlegbaar is aan betrokkenen. Tevens zijn er bij ontwikkeling en gebruik van algoritmes vaak meerdere stakeholders betrokken in het proces, zoals een programmeur, een data-scientist, een domein-expert en een eindgebruiker. De verantwoordelijkheid voor het beheersen van risico's ligt daardoor in de praktijk veelal verspreid over de organisatie.

Eenvoudige algoritmes zijn redelijk goed door mensen te controleren, omdat het redelijk goed te herleiden is waarom het algoritme bepaalde keuzes maakt. Hierdoor kunnen fouten in het algoritme eerder worden opgemerkt. Wanneer een algoritme complexer wordt, zijn de risico's doorgaans groter. Dat eenvoudige algoritmes relatief lage risico's hebben, is echter niet per definitie waar. Dit illustreert het voorbeeld van de 'fraudescorekaart'. Dit is een eenvoudig rule-based model dat door verschillende Nederlandse gemeenten werd toegepast en met enkele kenmerken bepaalde of uitkeringsontvangers een grotere kans hadden om uitkeringsfraude te plegen. Dit systeem bleek na onderzoek gebaseerd te zijn op twintig jaar oude statistische verbanden waarvan de oorsprong door de bedenkers niet meer te achterhalen was. In 2022 bleek dit systeem nog steeds gebruikt te worden in enkele Nederlandse gemeenten.⁷

De rekenkamer heeft in haar verkenning uit 2021 gekeken naar drie ethische principes en de bijbehorende risico's, namelijk verantwoordelijkheid, eerlijkheid en transparantie. In de tussentijd is er meer onderzoek gedaan naar de risico's van algoritmes, onder andere door de Algemene Rekenkamer.

De Algemene Rekenkamer heeft in 2021 een toetsingskader ontwikkeld waarmee de risicobeheersing van algoritmes in kaart kan worden gebracht. Zij onderscheidt in dit toetsingskader vijf perspectieven waarbinnen risico's aanwezig zijn (figuur 2-2). Ethiek is hier een onderdeel van.

figuur 2-2: perspectieven toetsingskader Algemene Rekenkamer



bron: Algemene Rekenkamer, 'Aandacht voor algoritmes', 26 januari 2021, p.13.

De Algemene Rekenkamer heeft vanuit wet- en regelgeving, richtlijnen en bestaande kaders de risico's van algoritmes geformuleerd. Risico's zijn aanwezig op het gebied van sturing en verantwoording, privacy, data en model en rondom IT (Information Technology). De ethische risico's van algoritmes zijn gekoppeld aan deze vier aspecten. De grootte van de risico's neemt doorgaans toe met de complexiteit van het

⁷ NRC, 'Ondanks waarschuwingen gingen gemeenten door met omstreden fraudedetectiemethode', 24 juni 2022.

algoritme en de impact van het algoritme op de burger. In 2022 publiceerde de Algemene Rekenkamer het onderzoek 'Algoritmes getoetst', waarin zij concludeerde dat van de negen getoetste algoritmes er zes niet helemaal voldoen aan de voorwaarden uit het toetsingskader. De belangrijkste knelpunten bij deze onvoldoende beheerste algoritmes waren het IT-beheer, het risico op vooringenomenheid en goede afspraken over werking en evaluatie.⁸

Het belang van evaluatie van (de risico's van) algoritmes wordt ook door Berenschot onderstreept in een rapport over de verantwoorde toepassing van artificiële intelligentie (AI). Zij zien in de levenscyclus van een algoritme verschillende 'feedbackloops' waarin op basis van ervaringen de algoritmetoepassing wordt doorontwikkeld of juist beëindigd.⁹ Deze feedbackloops dienen niet alleen tijdens de ontwikkeling van een algoritme plaats te vinden, maar ook op het moment dat een algoritme in de zogenaamde 'beheerfase' zit en dus ingezet wordt in lopende processen.



bron: Berenschot, 'Voor een verantwoorde toepassing van AI', 25 januari 2022, p.8.

2-3 Europees en landelijk beleid algoritmes

Op zowel Europees als landelijk niveau is beleid en wetgeving om te komen tot een verantwoorde inzet van algoritmes volop in ontwikkeling. In 2021 diende de Europese Commissie (EC) een voorstel in voor een verordening voor AI om risico's te beperken.¹⁰ Deze verordening stelt eisen aan publieke en private organisaties die met algoritmes werken op basis van vier verschillende risicocategorieën van algoritmes. Op dit moment wordt aan de definitieve AI-verordening gewerkt die waarschijnlijk in 2024 in werking treedt.

Vooruitlopend hierop zijn de afgelopen jaren op landelijk niveau diverse kaders, richtlijnen en instrumenten ontwikkeld voor overheden en bedrijven die algoritmes toepassen, waaronder het 'Impact Assessment voor Mensenrechten en Algoritmes'.¹¹ Ook kondigde het kabinet-Rutte IV in het coalitieakkoord aan wettelijk te willen regelen dat algoritmes worden gecontroleerd op discriminatie, transparantie en

⁸ Algemene Rekenkamer, 'Algoritmes getoetst', 13 mei 2022.

⁹ Berenschot, 'Voor een verantwoorde toepassing van AI', 25 januari 2022.

¹⁰ Europese Commissie, 'Verordening van het Europees Parlement en de Raad tot vaststelling van geharmoniseerde regels betreffende artificiële intelligentie (wet op de Artificiële Intelligentie) en tot wijziging van bepaalde wetgevingshandelingen van de Unie', 21 april 2021.

¹¹ Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, 'Impact Assessment Mensenrechten en Algoritmes', 31 juli 2021.

willekeur.¹² Hoe aan de ambities van het kabinet wat betreft algoritmes en digitalisering invulling wordt gegeven, is recent beschreven in de ‘Werkagenda Waardengedreven Digitaliseren’.¹³ Hierin staat onder meer dat er eisen worden gesteld aan overheidsorganisaties die algoritmes inzetten, er een landelijk algoritmeregister ingesteld wordt en dat een algoritmetoezichthouder wordt ingericht.

2-4 gemeentelijk beleid algoritmes

De gemeente Rotterdam heeft al enige tijd aandacht voor de monitoring en kwaliteit van de toepassing van algoritmes vanuit het programma Datagedreven Werken (DGW). DGW is een gemeentelijke stuurgroep die jaarlijks een actieplan opstelt om te komen tot een datagedreven organisatie. Onderdeel van het actieplan DGW 2020 was de kwaliteitsborging van analysemodellen en algoritmes.¹⁴ In het kader hiervan is de gemeente Rotterdam een consortium aangegaan met enkele gemeenten en andere overheden om gezamenlijk beleidsinstrumenten voor algoritmes te ontwikkelen. De instrumenten die in 2020 op de planning stonden waren onder andere een algoritmekader om processen met algoritmes te kunnen beheersen en een algoritmeregister.

Na het verschijnen van het verkennende onderzoek van de rekenkamer naar algoritmes is de ontwikkeling van een governancekader voor algoritmes in een stroomversnelling geraakt. Een belangrijk uitgangspunt hierbij was dat de gemeente besloot dat het beleid en de governance gericht moesten zijn op hoog-risico algoritmes.¹⁵ Deze classificatie komt voort uit de AI-verordening van de EC waarin dit onderscheid eveneens wordt gemaakt. De gemeente beschrijft dat hoog-risico algoritmes ‘aanzienlijke risico’s kunnen opleveren voor de gezondheid, veiligheid of grondrechten van personen’. In 2021 en 2022 heeft het college de gemeenteraad een aantal keer geïnformeerd over de stand van zaken met betrekking tot de governance van algoritmes.¹⁶ Op basis van deze documenten worden de belangrijkste onderdelen van het governancekader hieronder kort uiteengezet.

- 1 Algorithm Risk Assessment (ARA). De ARA is een instrument om te bepalen of er sprake is van een hoog-risico algoritme. Indien op basis van de ARA de conclusie wordt getrokken dat sprake is van een hoog-risico algoritme is de algoritme governance van toepassing.
- 2 Algoritmeregister en inventarisatie algoritmes. In 2022 heeft de gemeente geïnventariseerd welke algoritmes binnen de gemeente Rotterdam worden gebruikt. Vervolgens zijn deze algoritmes aan de hand van de ARA geclassificeerd. Alle algoritmes, dus zowel de laag- als hoog-risico algoritmes, worden samen met de technische documentatie opgenomen in een intern algoritmeregister. De hoog-risico algoritmes worden daarnaast opgenomen in een openbaar algoritmeregister.
- 3 Impact Assessment Mensenrechten en Algoritmes (IAMA). Het IAMA is een hulpmiddel om te komen tot een antwoord op de vraag of een algoritmetoepassing verantwoord is, wat de risico’s zijn en welke eventuele beheersmaatregelen nodig

¹² Coalitieakkoord 2021-2025, ‘Omzien naar elkaar, vooruitkijken naar de toekomst’, 15 december 2021.

¹³ Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, ‘Werkagenda Waardengedreven Digitaliseren’, 31 juli 2022.

¹⁴ Gemeente Rotterdam, ‘Actieplan DGW 2020’, 28 januari 2020, p.14.

¹⁵ Burgemeester en Wethouders van Rotterdam, ‘Governance rondom algoritmes’, 7 september 2021.

¹⁶ Burgemeester en Wethouders van Rotterdam, ‘Governance rondom algoritmes’, 7 september 2021 en Burgemeester en Wethouders van Rotterdam, ‘Stand van zaken algoritmes’, 15 februari 2022 en Burgemeester en Wethouders van Rotterdam, ‘Algoritmegovernance’, 7 juni 2022.

zijn. Het IAMA wordt uitgevoerd voor hoog-risico algoritmes voordat deze in gebruik worden genomen. Omdat een deel van de hoog-risico algoritmes al in gebruik zijn worden IAMA's met terugwerkende kracht uitgevoerd.

- 4 Aanstellen algoritme-expert. De functie van algoritme-expert is nieuw binnen de gemeentelijke organisatie en sinds maart 2022 vervuld. De algoritme-expert ondersteunt de organisatie concernbreed met advies over de ethisch verantwoorde toepassing van algoritmes. Daarnaast ondersteunt de algoritme-expert desgevraagd bij het uitvoeren van een IAMA.
- 5 Algorithm Advisory Board (AAB). In juni 2022 is door het college een externe adviescommissie ingesteld bestaande uit vijf leden die expert zijn op het gebied van algoritmes. De adviesraad geeft gevraagd en ongevraagd advies over de inzet van technologische innovaties en digitale middelen, maakt een jaarverslag en analyseert algoritmetoepassingen binnen de gemeente op kwaliteit.

afwegingskader voor sensoren, data-toepassingen en AI

In 2021 heeft de gemeenteraad de motie 'Wie smart is moet ook slim zijn' aangenomen.¹⁷ Hierin werd het college onder meer opgedragen om tot een breed beleidskader te komen voor 'smart city oplossingen die voortkomen uit (toekomstige) sensor, data-toepassingen en AI-inzet'. Hier is invulling aan gegeven door het opstellen van een afwegingskader voor sensoren, data-toepassingen en AI.¹⁸ Dit kader biedt algemene richtlijnen voor onder andere de toepassing van algoritmes. De gemeente beschrijft dat dit afwegingskader, in combinatie met het hiervoor beschreven governancekader, ertoe dient een zorgvuldige en transparante toepassing van algoritmes binnen de gemeente te borgen. Het college stelt voor dat het kader nog verder wordt uitgewerkt door per principe aan te geven welk beleid en instrumenten hiervoor nodig zijn.¹⁹

privacy en informatiebeveiliging

Het governancekader voor algoritmes is een nieuw ontwikkeld kader, maar gegevensverwerkingen bij ontwikkeling en uitvoering van algoritmes vallen tevens onder bestaande kaders voor privacybescherming. De Algemene Verordening Gegevensbescherming (AVG) stelt namelijk eisen aan processen waarin (privacygevoelige) gegevens worden verwerkt. Algoritmes waarin persoonsgegevens worden verwerkt, vallen onder de AVG. Dit betekent dat bij dit type verwerkingen van data verschillende stappen dienen te worden doorlopen.

- 1 Bij alle verwerkingen wordt eerst een Privacy Risico Assessment (PRA) uitgevoerd. De PRA is een intake bij nieuwe projecten en zowel simpele als complexe processen, waarin aan de hand van een vragenlijst een overzicht wordt gemaakt van welke gegevens de gemeente gaat gebruiken en daaraan verwante risico's.
- 2 De 'privacy officer' van het cluster bepaalt aan de hand van de PRA of een Data Protection Impact Assessment (DPIA) wettelijk noodzakelijk, wenselijk of niet nodig is. Indien van toepassing voert de privacy officer of een functionaris gegevensbescherming van het cluster een DPIA uit.
- 3 In een DPIA worden alle vereisten die de AVG stelt aan een verwerking, gecontroleerd.²⁰ Bijvoorbeeld of het doel waarvoor de gebruikte informatie is

¹⁷ Gemeenteraad Rotterdam, 'Wie smart is moet ook slim zijn', 18 maart 2021

¹⁸ Gemeente Rotterdam, 'Afwegingskader voor sensoren, data-toepassingen en AI', 21 maart 2022

¹⁹ Burgemeester en Wethouders van Rotterdam, 'Afwegingskader voor sensoren, data-toepassingen en AI', 2 september 2022

²⁰ Dit is de zogenaamde gegevensbeschermingseffectbeoordeling zoals genoemd in artikel 35 van de AVG.

verzameld verenigbaar is met het doel van de verwerking in het algoritme en de databeveiliging.

- 4 Elke drie jaar wordt een DPIA herzien (of direct wanneer de verwerking van gegevens significant wijzigt).²¹

Naast de hierboven genoemde kaders rondom privacybescherming zijn er ook algemene werkwijzen rondom informatiebeveiliging (zoals autorisaties voor toegang tot bepaalde stukken). Deze werkwijzen rondom informatiebeveiliging gelden ook voor gegevensverwerking en algoritmes.

2-5 bestuurlijke en ambtelijke verantwoordelijkheid

De governance van algoritmes valt onder de bestuurlijke verantwoordelijkheid van de wethouder Financiën, Organisatie, Dienstverlening en Grote Projecten. De bestuurlijke verantwoordelijkheid voor de toepassing en keten van specifieke algoritmes valt onder de wethouder die de bestuurlijk opdrachtgever is van het betreffende algoritme.

De ambtelijke verantwoordelijkheid voor de governance van algoritmes is belegd bij de directeur Innovatie, Informatievoorziening, Facilitair en Onderzoek (IIFO). Ambtelijke verantwoordelijkheid voor specifieke algoritmes ligt bij de desbetreffende clusterdirecteur.

²¹ Gemeente Rotterdam, 'Proces voor de tijdige uitvoering en borging van de resultaten van een Privacy Impact Assessment', 18 januari 2019.

3 onderzoeksvragen en afbakening

3-1 doel- en vraagstelling

Binnen de gemeente Rotterdam is er in toenemende mate aandacht voor de risico's van de inzet van algoritmes en zijn het afgelopen jaar diverse stappen gezet in de implementatie van een governancekader. De toepassing is op dit moment nog in ontwikkeling. Dit betekent dat de rekenkamer zich richt op een praktijk die nog niet volledig is uitgekristalliseerd. Zo kunnen gedurende de looptijd van het onderzoek bijstellingen plaatsvinden. Niettemin acht de rekenkamer het juist van belang in deze fase onderzoek te doen. Door vroegtijdig inzicht te geven in de mate waarin de risico's worden beheerst, kan het onderzoek de raad ondersteunen bij de kaderstellende rol. Bij de uitvoering van het onderzoek heeft de rekenkamer oog voor de fase waarin de implementatie zich bevindt. De focus ligt dan ook niet op eventuele kinderziekten die de gemeente al zelf oppakt.

De centrale onderzoeksvraag luidt als volgt:

In hoeverre is de algoritme-governance van de gemeente Rotterdam toereikend voor een verantwoorde toepassing van algoritmes en in hoeverre leidt dit tot het beheersen van risico's in de praktijk?

De centrale onderzoeksvraag is uitgewerkt in de volgende deelvragen:

- 1 Wat is de aard en omvang van de inzet van algoritmes binnen de gemeente?
- 2 Hoe is de governancestructuur en het bijbehorende instrumentarium voor de inzet van algoritmes vormgegeven?
- 3 Is het instrumentarium om de risico's van algoritmes te beheersen in lijn met de heersende normen hieromtrent?
- 4 In hoeverre worden de risico's beheerst bij algoritmes die de gemeente in de praktijk inzet?

3-2 afbakening onderzoek

In het onderzoek wordt de definitie van een algoritme gebruikt, zoals geformuleerd in paragraaf 2-1. Het betreft dus een set instructies die door een computer wordt uitgevoerd om te komen tot een beslissing of tot informatie die bijdraagt aan besluitvorming. In het onderzoek wordt gefocust op algoritmes die leiden tot of bijdragen aan besluiten die impact hebben op burgers. Verder worden zowel algoritmes meegenomen die gedurende het onderzoek nog in ontwikkeling zijn, als algoritmes die worden toegepast. Dit laatste duidt de gemeente zelf aan als algoritmes die 'in productie' zijn.

In het onderzoek worden aan de hand van een toetsingskader de risico's inzichtelijk gemaakt vanuit vijf perspectieven: sturing en verantwoording, privacy, data en model, IT-beheer en ethiek. In het onderzoek worden mogelijke risico's geïdentificeerd, maar wordt niet onderzocht in hoeverre deze risico's daadwerkelijk optreden in de praktijk. De rekenkamer kijkt dus bijvoorbeeld wel of er bij een algoritme een risico bestaat op discriminatie, maar kijkt niet of met dat algoritme daadwerkelijk een aantal mensen is gediscrimineerd. Wel wordt onderzocht in hoeverre de gemeente op deze risico's heeft getoetst en daar actief toezicht op houdt.



Als laatste is het van belang te benoemen dat het een vervolgonderzoek is en niet alleen een opvolgsonderzoek. Dit betekent dat er breder wordt gekeken dan alleen de vraag of de aanbevelingen uit het rapport 'Gekleurde technologie' zijn opgevolgd.

4 onderzoeksaanpak

Voor de beantwoording van de onderzoeksvragen wordt een combinatie van methoden gebruikt.

literatuuronderzoek

Om inzicht te krijgen in recente ontwikkelingen op het gebied van ethiek en algoritmes worden wetenschappelijke artikelen bestudeerd. Daarnaast worden onderzoeken van kennisinstituten naar algoritmes en het beleid hieromtrent geraadpleegd.

documentstudie

Om zicht te krijgen op het beleid rondom de toepassing van algoritmes wordt een documentstudie gedaan. Hiervoor worden relevante (beleids)documenten, rapportages, kaders, en collegestukken verzameld en bestudeerd. De rekenkamer kijkt specifiek naar hoe de governancestructuur is vormgegeven en tot stand is gekomen.

interviews

Door middel van interviews met betrokkenen binnen de gemeente wordt inzicht verkregen in de governancestructuur en de implementatie hiervan. Er worden interviews gehouden met:

- Ambtenaren die zich bezighouden met de ontwikkeling en toepassing van algoritmes, waaronder ontwikkelaars in het data-science team van de afdeling Onderzoek en Business Intelligence (OBI) en gebruikers van algoritmes binnen de verschillende clusters.
- Personen die betrokken zijn bij de ontwikkeling of implementatie van de governancestructuur, bijvoorbeeld ambtenaren van het programma DGW, afdelingshoofden van OBI en privacy-officers.
- Functionarissen in rollen die recent zijn belegd als onderdeel van de governancestructuur, zoals de algoritme-expert, de secretaris en leden van de AAB.

participerende observatie

De rekenkamer kijkt gedurende het onderzoek naar mogelijkheden om participerende observatie toe te passen, om zo een beeld te krijgen van de wijze waarop het governancekader in de praktijk werkt. Daarbij kan bijvoorbeeld gedacht worden aan het bijwonen van een bijeenkomst van de AAB.

betrekken experts

Het werkveld rondom algoritmes bestaat uit verschillende disciplines en is op dit moment nog sterk in ontwikkeling. De rekenkamer acht het daarom van belang om experts uit de wetenschap en van kennisinstituten bij het onderzoek te betrekken. Zij kunnen aanvullende inzichten bieden op het gebied van ethiek en algoritmes en de resultaten van het onderzoek plaatsen in een breder kader.

casestudy

Om zicht te krijgen op mogelijke risico's bij algoritmes binnen de gemeente Rotterdam wordt casestudy onderzoek gedaan. Het casestudy onderzoek wordt dus met name gedaan voor de beantwoording van deelvraag 4. In totaal worden drie tot vijf algoritmes geselecteerd. Om tot deze selectie te komen wordt eerst een inventarisatie gemaakt van de algoritmes die worden toegepast binnen de gemeente. Het interne algoritmeregister dient daarvoor als basis. Vervolgens wordt aan de ambtelijke

organisatie gevraagd één van deze algoritmes aan te dragen voor de casestudy. De overige algoritmes worden door de rekenkamer geselecteerd. Bij de selectie wordt gestreefd naar voldoende variatie wat betreft type algoritmes. Kenmerken waarnaar wordt gekeken zijn:

- 1 Ontwikkelingsfase: er kan onderscheid worden gemaakt in algoritmes die in ontwikkeling zijn en algoritmes die worden toegepast en in productie zijn. Daarnaast zijn sommige algoritmes recent ontwikkeld en worden anderen al enkele jaren toegepast.
- 2 Technische complexiteit: zoals beschreven in paragraaf 2-1 variëren algoritmes wat betreft complexiteit. Dit kan samenhangen met de mate van risico's.
- 3 Clusters van de gemeente: de gemeente heeft diverse toepassingsgebieden voor algoritmes die binnen verschillende clusters zijn ondergebracht.

samenwerking Rekenkamer Metropool Amsterdam

De Rekenkamer Metropool Amsterdam voert in dezelfde periode eveneens een onderzoek uit naar de toepassing van algoritmes.²² Tijdens het onderzoek worden ervaringen en kennis uitgewisseld met het onderzoeksteam van Amsterdam. Daarnaast wordt verkend of het mogelijk is om eenzelfde type algoritme te bestuderen in het casestudy onderzoek.

toetsingskader

Voor het beoordelen van de geselecteerde algoritmes worden documenten en verantwoordingsstukken verzameld. De beschikbare documentatie wordt bestudeerd en aanvullend worden personen geïnterviewd die betrokken zijn bij de ontwikkeling, implementatie en toepassing van de algoritmes. Op basis van deze informatie worden de algoritmes aan de hand van een toetsingskader beoordeeld vanuit vijf perspectieven, namelijk sturing en verantwoording, model en data, privacy, IT-beheer en ethiek.

Omdat er nog geen algemeen vastgesteld toetsingskader bestaat voor de toepassing van algoritmes binnen de gemeentelijke praktijk, wordt het toetsingskader van de Algemene Rekenkamer zoals beschreven in paragraaf 2-2 als uitgangspunt genomen. Indien nodig wordt het toetsingskader aan de hand van literatuuronderzoek, documentstudie en eventuele input van experts op punten aangepast zodat het beter aansluit op de gemeentelijke praktijk. Het ontwikkelen van het toetsingskader kan hiermee ook gezien worden als een opbrengst van het onderzoek.

²² Rekenkamer Metropool Amsterdam, 'Onderzoeksopzet Algoritmen', 28 juli 2022.

5 organisatie en planning

5-1 organisatie

Het onderzoek wordt uitgevoerd door een onderzoeksteam bestaande uit:

- Samantha Langendoen;
- Simone Luijk;
- Jaap Wils (projectleider).

5-2 planning

De uitvoering van het onderzoek start in januari 2023. De onderzoeksbevindingen worden vastgelegd in een concept nota van bevindingen die voor ambtelijk wederhoor aan de gemeente wordt aangeboden. Na verwerking van de ambtelijke reactie wordt een bestuurlijke nota opgesteld met daarin de conclusies en aanbevelingen die volgen uit het onderzoek. De bestuurlijke nota wordt, samen met de nota van bevindingen als bijlage, voorgelegd aan het college van B en W voor wederhoor. De bestuurlijke reactie wordt samen met het nawoord van de rekenkamer opgenomen in het definitieve rapport. Het definitieve rapport wordt naar verwachting in het najaar van 2023 aangeboden aan de raad.

5-3 privacy

Voor het onderzoek interviewt de rekenkamer verschillende functionarissen van de gemeente Rotterdam en samenwerkingspartners. De persoonsgegevens die in dit kader worden verwerkt worden zoveel mogelijk beperkt, maar betreffen in elk geval naam, functie en zakelijke contactgegevens. Deze gegevens zullen conform de AVG en het daarop gebaseerde privacybeleid van de rekenkamer worden verwerkt (www.rekenkamer.rotterdam.nl). De rekenkamer zal geen bijzondere persoonsgegevens verwerken.