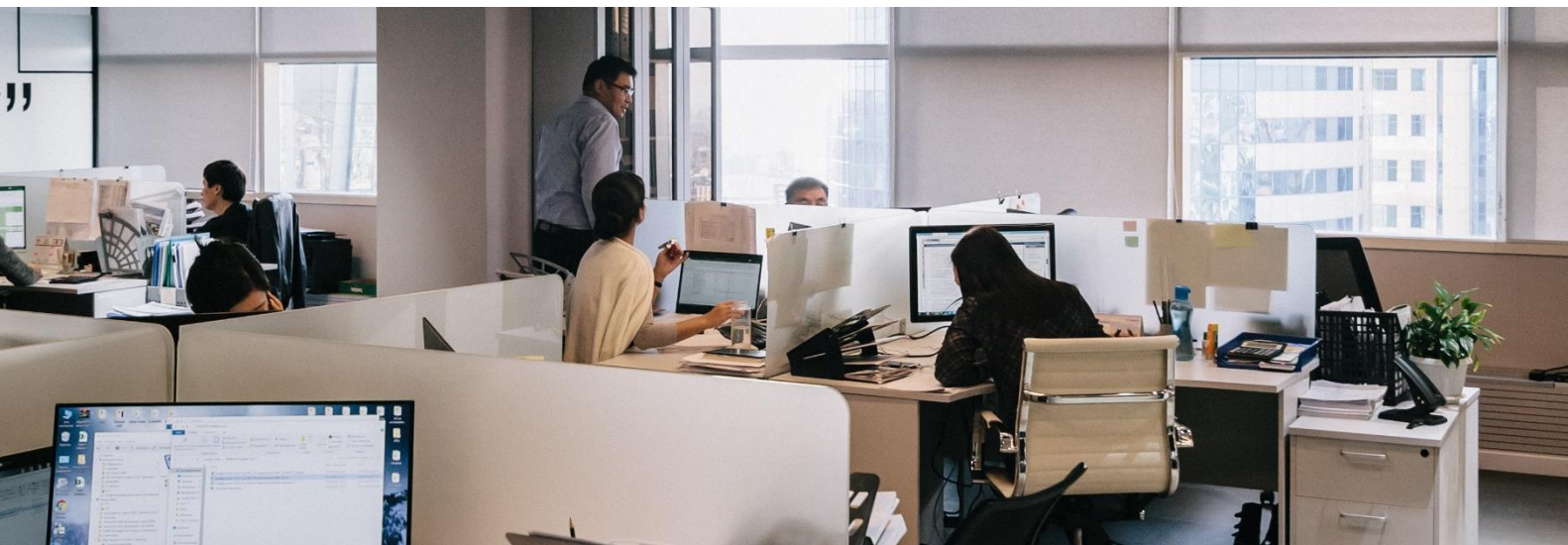


Data-ethiek: moeilijke keuzes en meer informatievoorziening



Bericht aan het parlement

Op 15 februari 2023 spreekt de commissie Digitale Zaken over de inzet van algoritmes en data-ethiek binnen de rijksoverheid. Er is door het kabinet al veel in gang gezet rondom het verantwoord inzetten van algoritmes. Maar er kunnen verdere stappen genomen worden. En ook is het tijd om het gesprek aan te gaan over kwesties die niet met maatregelen alleen op te lossen zijn. Met name stelt de inzet van algoritmes overheid en politiek voor lastige keuzes op het terrein van non-discriminatie, privacy en veiligheid. Het Rathenau Instituut geeft hiervoor drie handvatten.

De risico's van algoritmes zijn zelden geheel te elimineren

Het onderwerp van algoritmes en data-ethiek staat al een aantal jaar op de bestuurlijke en politieke agenda, en de rijksoverheid neemt tal van maatregelen om ethische en mensenrechtelijke risico's van algoritmes te mitigeren. Zo zijn onder andere de *Code Goed Digitaal Openbaar Bestuur*, de *Handreiking non-discriminatie by design* en de *Impact Assessment Mensenrechten en Algoritmes* ontwikkeld, zijn publieke waarden en mensenrechten verankerd in de digitale strategie van de Rijksoverheid, en toetsen tal van overheidsorganisaties hun data en algoritmes op *bias*, vooringenomenheid.

In reactie op de aanbevelingen van onze studie [Algoritmes afwegen](#), beschreef de staatssecretaris van BZK aanvullende maatregelen. Zo is het nationaal algoritmeregister gestart, en werkt het kabinet onder meer aan het opstellen van een generiek implementatiekader 'inzet van algoritmen' – op basis van de al ontwikkelde handreikingen – zodat voor overheidsorganisaties duidelijker wordt welke kaders zij moeten hanteren.

Hoewel dit pakket van acties zal bijdragen aan het verantwoord toepassen van algoritmes binnen de overheid, kunnen verdere stappen genomen worden. Zo kunnen juridische waarborgen op het terrein van non-discriminatie en het motiveringsbeginsel verder aangescherpt worden – daartoe doen we hieronder een aantal handreikingen.

Maar het is ook tijd om het gesprek aan te gaan over kwesties die niet geheel met maatregelen op te lossen zijn. Met name op het terrein van non-discriminatie, privacy en veiligheid stelt de inzet van algoritmes overheid en politiek voor lastige keuzes. Het zal zelden mogelijk zijn om alle kans op discriminatie te elimineren. Algoritmische systemen kunnen niet volledig bestand zijn tegen hacks. En er is geen garantie dat de privacy van mensen nooit ingeperkt wordt. Afwegingen tussen waarden zijn noodzakelijk, en dat vraagt om bestuurlijk en parlementair debat waarin soms lastige keuzes gemaakt moeten worden.

Om deze keuzes te kunnen maken, dient het parlement inzicht te hebben in de manier waarop overheidsorganisaties concrete toepassingen wegen. Het is daarom zaak dat overheidsorganisaties op gedetailleerde wijze aan het parlement rapporteren over de inzet van algoritmes. Alleen als de informatievoorziening aan het parlement op orde is, kan een inhoudelijk debat over de inzet van data en algoritmes gevoerd worden.

Samengevat geeft het Rathenau Instituut de volgende aanbevelingen mee:

1. Waarborg de informatievoorziening aan het parlement
2. Versterk de juridische waarborgen ten aanzien van non-discriminatie
3. Verhelder de juridische eisen ten aanzien van het motiveringsbeginsel

1. Waarborg informatievoorziening omtrent impactassessments aan het parlement

1.1 Maak impact assessments openbaar

Kabinet en Kamer zijn het eens dat overheidsorganisaties bij de ontwikkeling en inzet van algoritmes mensenrechtelijke risico's moeten onderzoeken en wegen. Daartoe zijn het IAMA en de handreiking non-discriminatie ontwikkeld en ook de AI Act bevat hiertoe bepalingen. Maar het verplicht stellen van een dergelijke impact assessment is slechts de eerste stap. Het parlement moet kunnen inzien hoe overheidsorganisaties moeilijke afwegingen maken, zodat ook op politiek niveau deze keuzes gewogen worden. Een inhoudelijk parlementair debat over de inzet van algoritmes bij de overheid vraagt om de openbaarmaking van impact assessments¹. De Tweede Kamer kan het kabinet vragen hoe zij de openbaarmaking gaat regelen.

1.2 Zorg dat impact assessments een complete proportionaliteitsafweging bevatten

Om te bepalen of een systeem wordt ingezet in lijn met mensenrechten, is het beoordelen van proportionaliteit belangrijk: zijn de risico's in balans met de maatschappelijke opbrengst, en zijn er geen alternatieven die hetzelfde resultaat opleveren met minder risico? [Onderzoek van het Rathenau Instituut](#) laat zien dat in de praktijk bij impact assessments niet alle elementen voldoende worden meegenomen. De Tweede Kamer kan het kabinet vragen het gebruik van impact assessments op drie punten goed te monitoren:

1. Worden de kosten van te nemen beheersmaatregelen meegenomen, zoals het aannemen van data-scientists en het uitvoeren van *bias*-controle?
2. Worden alternatieven meegewogen, inclusief de optie om geen algoritmes in te zetten?
3. Is er voldoende informatie gegeven over de effectiviteit en prestaties van het in te voeren systeem?

Omdat niet alle risico's, opbrengsten en effecten op voorhand duidelijk zijn, is het ook belangrijk dat de Tweede Kamer controleert of impact assessments periodiek worden herhaald en over de inzichten debatteert.

1.3 Bespreek in impact assessments de fairness metric

Algoritmes mogen niet discrimineren, maar dat is lastig 100% te garanderen. Er zijn daarom zogeheten *fairness metrics* ontwikkeld om bias op te sporen. Het punt is echter dat deze *metrics* onderling verschillen: zo kijkt de ene methode of een bepaalde groep oververtegenwoordigd is in de uitkomsten, en houdt de andere juist rekening met verschillen tussen groepen, om de kans op selectie gelijk te maken. Het hanteren van de ene of de andere methode is een normatieve, politieke keuze, die samenhangt met een opvatting over wat eerlijk is. Daarom is het zaak dat de politiek geïnformeerd wordt over welke methode overheidsorganisaties hanteren.

¹ De [motie Bouchallikht - Dekker-Abdulaziz](#) riep daartoe al op

2. Versterk juridische waarborgen ten aanzien van non-discriminatie

2.1 Versterk de rechtspositie van burgers

Het is nu voor burgers erg lastig om aan te tonen dat ze door een algoritme zijn gediscrimineerd. De rechtspositie van burgers kan op een aantal vlakken versterkt worden. Allereerst door burgers te laten weten dat een algoritme is toegepast. Geef aan wanneer een overheidsbesluit mede gebaseerd is op een algoritme of risicoprofilering. Onlangs is hiertoe de motie- Van Baarle aangenomen. Meer inzicht in achterliggende profielen, het versimpelen en deels omdraaien van de bewijslast, maken het ook makkelijker om discriminatie aan te tonen. De Tweede Kamer kan het kabinet vragen deze opties te onderzoeken.

2.2 Onderzoek of meer beschermde gronden nodig zijn

Het non-discriminatierecht verbiedt het gebruik van bepaalde discriminatiegronden, waaronder geslacht, religie en etniciteit. De verwachting is dat lerende algoritmes mensen steeds opnieuw in bepaalde groepen zullen sorteren, op manieren die we nu niet kunnen voorspellen. Dat kunnen gevoelige gronden zijn, zoals het inkomen van mensen, en kan dus mensen nadelig raken. De Tweede Kamer kan het kabinet vragen te laten onderzoeken of meer beschermde gronden in de wet dienen te worden opgenomen.

3. Verhelder de juridische eisen ten aanzien van het motiveringsbeginsel

3.1 Verhelder motiveringseisen voor beleidsondersteunende algoritmes

Lerende algoritmes worden binnen de overheid momenteel vooral ingezet als 'beleidsondersteuning'. Deze algoritmes nemen niet autonoom overheidsbesluiten, maar produceren analyses en selecties op basis waarvan ambtenaren besluiten nemen. Een ambtenaar kan bijvoorbeeld op basis van het systeem bekijken of het nodig is een fraude-onderzoek in te stellen. Een ambtenaar voert het onderzoek uit. Via deze 'menselijke tussenkomst' kan de uitkomst worden gemotiveerd. Maar dat biedt geen uitleg over de werking van het systeem, of motivering waarom een burger werd geselecteerd. De bovengenoemde motie -Van Baarle verzoekt de regering te onderzoeken hoe de motiveringsplicht bij overheidsbesluiten zodanig kan worden aangepast dat bij een besluit altijd wordt aangegeven of deze gebaseerd is op een algoritme of risicoprofilering. De volgende stap is uitleg over de werking van het algoritme. De AI Act dient hiertoe bij te dragen, net als het algoritmeregister. Het algoritmeregister zal echter geen motivering geven over een concreet besluit. Op dit moment is onduidelijk welke motiveringseisen precies gelden voor deze beleidsondersteunende algoritmes. De Tweede Kamer kan kabinet vragen dit te verhelderen.

3.2 Verhelder eisen toelaatbaarheid *unsupervised learning*

Ten tweede is er een spanningsveld bij uitvoeringsorganisaties bij de inzet van unsupervised learning algoritmes. Diverse overheidsorganisaties kiezen er bewust voor

deze techniek niet in te zetten, zelfs niet als deze systemen betere resultaten boeken, omdat ze de technologie niet verenigbaar achten met het motiveringsbeginsel. Dit lijkt ook aan te sluiten bij de jurisprudentie over dit beginsel. Tegelijkertijd is duidelijk dat de ontwikkelaars bij deze organisaties wel experimenteren met deze technieken en dat er in de praktijk druk kan ontstaan om ze toch in te zetten. De Tweede Kamer kan het kabinet vragen om heldere normstelling.

Gerelateerde publicaties:

<u>Algoritmes afwegen</u> (2022)	Rapport
<u>Algoritmes afwegen</u> (2022)	Bericht aan het parlement
<u>Non-discriminatie bij algoritmes</u> (2022)	Bericht aan het parlement
<u>Handreiking AI Eerste Kamer</u> (2022)	Notitie
<u>Digitalisering in dienst van de samenleving</u> (2022)	Bericht aan het parlement