



Universiteit Utrecht



MONTAIGNE
CENTRUM

VOOR RECHTSSTAAT
EN RECHTSPLEGING

S. Kulk & S. van Deursen, *Juridische aspecten van algoritmen die besluiten nemen. Een verkennend onderzoek*, Den Haag: WODC 2020

Samenvatting

1. Aanleiding en onderzoeksvraag

Iedereen heeft in het dagelijks leven te maken met beslissingen die worden genomen door of met behulp van algoritmen. De inzet van algoritmen kan kansen opleveren voor het verwezenlijken van publieke waarden en belangen. Zo kunnen algoritmen besluitvormingsprocessen efficiënter maken en bijdragen aan het vinden van oplossingen voor verschillende soorten maatschappelijke uitdagingen. Tegelijkertijd kan de inzet van algoritmen risico's met zich brengen en vragen oproepen over de bestendigheid van de juridische kaders die beschikbaar zijn om publieke waarden en belangen te beschermen. De onderzoeksvraag van dit onderzoek is in dit verband als volgt gedefinieerd:

Welke kansen en risico's doen zich voor bij algoritmische besluitvorming met betrekking tot de bescherming en realisering van publieke waarden en belangen, en zijn de bestaande juridische kaders voldoende bestendig om kansen te verwezenlijken en het intreden van geïdentificeerde risico's te voorkomen of de gevolgen daarvan te mitigeren?

Centraal in het onderzoek staan de huidige toepassingen van algoritmen in besluitvormingsprocessen en de ontwikkelingen die in de komende vijf tot tien jaar op dat gebied te verwachten zijn. Voor de beantwoording van de onderzoeksvraag is onder meer gebruikgemaakt van casestudy's naar de inzet van algoritmen in vier, door het WODC en de Directie Wetgeving en Juridische Zaken van het Ministerie van Justitie en Veiligheid geselecteerde domeinen: contentmoderatie, zelfrijdende auto's, rechtspraak en overheidsincasso bij verkeersboetes.

2. Algoritmische besluitvorming en publieke belangen

Algoritmische besluitvorming in de zin van dit onderzoek omvat alle processen waarin een algoritme wordt ingezet om beslissingen te nemen die raken aan de rechtspositie van rechtssubjecten of die hen anderszins in hun belangen treffen. Het kan daarbij gaan om gevallen waarin een algoritme zelf een besluit neemt, of gevallen waarin de uitvoer van een algoritme wordt meegenomen in een menselijk besluitvormingsproces. Vormen van algoritmische besluitvorming kunnen verschillen afhankelijk van het doel waarmee algoritmen worden ingezet. Ook kan er op technisch vlak een onderscheid worden gemaakt tussen regelgebaseerde algoritmen en zelflerende algoritmen. Tot slot is ook de organisatorische en maatschappelijke context waarin algoritmen worden ingezet van belang.

Zowel voor het identificeren van kansen en risico's, als voor de beoordeling van de bestendigheid van de juridisch kaders, fungeren publieke waarden en belangen als normatief kader. De publieke waarden en belangen zijn vanwege de juridische invalshoek van het onderzoek geconcretiseerd aan de hand van drie grondrechten die vrijwel steeds aan de orde zijn bij de verschillende toepassingen van algoritmen: het recht op gegevensbescherming, het recht op non-discriminatie en het recht op rechtsbescherming. In de verschillende casestudy's spelen daarnaast meer specifieke waarden als duurzaamheid of de vrijheid van meningsuiting een rol.

3. Kansen en risico's van algoritmische besluitvorming

Kansen zijn in dit onderzoek gedefinieerd als de mogelijkheid om publieke waarden en belangen te verwezenlijken, terwijl het bij risico's gaat om de mogelijkheid dat publieke waarden en belangen niet verwezenlijkt of zelfs geschaad worden. Bij de bestendigheid van het juridisch kader gaat het om de vraag in hoeverre het juridisch kader het mogelijk maakt dat kansen worden verwezenlijkt en dat het intreden van risico's wordt vermeden of gemitigeerd.

Kansen en risico's voor publieke waarden en belangen hangen samen met het type algoritme dat wordt gebruikt, maar voor een belangrijk deel ook met het domein en de organisatorische en maatschappelijke context waarin algoritmen worden ingezet. Uiteraard geldt dat bij de inzet van een algoritme tegelijkertijd kansen en risico's kunnen bestaan. De inzet van algoritmen in een besluitvormingsproces vergt dan ook altijd een afweging van de behoeften in het betreffende domein en de waarden en belangen die daarin gelden. De juridische kaders geven de grenzen aan waarbinnen dergelijke afwegingen dienen plaats te vinden.

3.1 Kansen van algoritmische besluitvorming

De inzet van algoritmen kan efficiëntiewinst opleveren doordat zij in staat kunnen zijn om (besluitvormings)processen sneller, beter of nauwkeuriger te doorlopen. Zelflerende algoritmen in het bijzonder kunnen verbanden ontdekken in grote hoeveelheden gegevens. De efficiëntiewinst

die op die manier geboekt kan worden, heeft veelal (bedrijfs)economische waarde, maar kan ook een belangrijke bijdrage leveren aan het verwezenlijken van publieke waarden.

Meer in het bijzonder creëert de inzet van algoritmen concrete kansen ten aanzien van de waarden van rechtsbescherming en non-discriminatie. **De kansen voor rechtsbescherming** hangen met name samen met de efficiëntiewinst die door de inzet van algoritmen geboekt kan worden, zoals vooral blijkt uit de casestudy naar de rechtspraak.

De inzet van algoritmen creëert ook **kansen voor het recht op non-discriminatie**. Doordat algoritmen in staat zijn om veel informatie te verwerken kunnen ze veel individuele kenmerken van personen meenemen in besluitvormingsprocessen. Algoritmen kunnen daarnaast helpen om besluiten beter af te stemmen op betrokken personen, waardoor de inzet van algoritmen kan bijdragen aan het realiseren van materiële gelijkheid. Ook kunnen algoritmen bijdragen aan gelijkheid in de vorm van consistentie van besluitvorming. Bovendien zijn goed-geprogrammeerde en gevalideerde algoritmen in beginsel beter dan mensen in staat om zonder aanzien des persoons een besluit te nemen. Tot slot kunnen algoritmen juist ook worden ingezet om discriminatie in besluitvormingsprocessen te detecteren.

Ten aanzien van de **bescherming van persoonsgegevens** zijn, in het kader van dit onderzoek, **geen mogelijke kansen** vastgesteld.

3.2 Risico's van algoritmische besluitvorming

Risico's ten gevolge van de inzet van algoritmische besluitvorming kunnen bestaan voor alle drie de algemene publieke waarden en belangen. De **bescherming van persoonsgegevens** komt in het gedrang als algoritmen worden ingezet om op grote schaal persoonsgegevens te verzamelen of anderszins te verwerken. Daarnaast maakt de inzet van algoritmen het mogelijk om in bestaande informatie verbanden aan te brengen en zo (nog) meer te weten te komen over personen. Daardoor verliezen individuen niet alleen controle op hun persoonsgegevens, maar vervagen ook de grenzen tussen wat persoonsgegevens zijn en wat niet. Als algoritmen verbanden leggen die niet kloppen, kan dat bovendien raken aan de identiteit en reputatie van individuen.

Risico's in verband met het **recht op non-discriminatie** zijn er als algoritmen ten onrechte onderscheidingen maken, of die onderscheiding ten onrechte niet maken. Dat kan zich voordoen als er bij het toepassen van regels of het leggen van verbanden sprake is van over- of onderinclusiviteit van bepaalde categorieën. Discriminatie-risico's kunnen zich ook concreet voordoen als vooroordelen of onaanvaardbare stereotypen via de programmeur(s) van een algoritme of via anderen die betrokken zijn bij de ontwikkeling van het algoritme (bewust of onbewust) een weerslag krijgen op het algoritme. Als zelflerende algoritmen worden ingezet

bestaat verder het gevaar dat de data waarmee het algoritme wordt getraind, gevalideerd, of getest niet voldoende representatief zijn voor de groep mensen waarover wordt beslist. Een vergelijkbaar risico van discriminatie bestaat wanneer de gebruikte data een reflectie vormen van problematisch geachte maatschappelijke stigmatisering, stereotypering of vooroordelen. Als er bij de toepassing van het algoritme nieuwe data wordt verzameld die wordt gebruikt om het algoritme te trainen kan er bovendien een *feedback loop* ontstaan waarin het discriminerende effect wordt versterkt.

Het recht op **rechtsbescherming** kan in het gedrang komen doordat de werking van algoritmen vaak lastig uitlegbaar is, wat in de weg kan staan aan de inzichtelijkheid van het besluitvormingsproces en daarmee de mogelijkheid om met argumenten te ageren tegen een besluit. Het probleem van uitlegbaarheid kan komen door de koppeling van verschillende (regelgebaseerde) algoritmische systemen, maar kan ook het gevolg zijn van het gebruik van zelflerende systemen, die inherent lastig uitlegbaar zijn. Van belang is verder dat de verantwoordelijkheidsvragen toenemen en complexer te beantwoorden zijn naarmate de samenleving verder gedigitaliseerd raakt en systemen in toenemende mate met elkaar interacteren. De daaruit resulterende onduidelijkheid over de verantwoordelijkheid in (de keten van) algoritmische besluitvorming, levert eveneens een risico op voor effectieve rechts-bescherming.

4. Bestendigheid juridisch kader

Omdat kansen en risico's van de inzet van algoritmen sterk samenhangen met het domein waarin zij worden ingezet, dient de bestendigheid van de juridische kaders ook plaats te vinden in het licht van de betreffende domeinen. Daarbij geldt dat een concrete toepassing altijd ten dele wordt gereguleerd door domeinspecifieke juridische kaders en ten dele door algemene juridische kaders.

Uit de casestudy's volgt dat de daarin bestudeerde specifieke juridische kaders niet direct in de weg lijken te staan aan het realiseren van kansen ten aanzien van de in dit onderzoek betrokken publieke waarden en belangen. Wel kan het zo zijn dat de juridische kaders onvoldoende voedingsbodem bieden om te kunnen profiteren van de voordelen van de inzet van algoritmen.

Uit de casestudy's volgt daarnaast dat de bestudeerde specifieke juridische kaders veelal voldoende ruimte bieden om geïdentificeerde risico's van algoritmische besluitvorming te vermijden of te mitigeren. Daarvoor is wel vaak vereist dat interpretaties van de bestaande ruim geformuleerde normen worden toegesneden op de specifieke inzet van het algoritme en de daarbij geldende waarden en belangen. Uit het onderzoek blijkt echter ook dat (gedeelten) van specifieke juridische kaders in sommige gevallen tekort schieten om risico's te voorkomen of te mitigeren.

Ten aanzien van de algemene juridische kaders, zoals de AVG, de Awb en het aansprakelijkheidsrecht, kan op basis van het onderzoek worden geconcludeerd dat zij een groot absorberend vermogen hebben ten aanzien van nieuwe technologische ontwikkelingen ten aanzien van algoritmische besluitvorming. De algemene juridische kaders kunnen daardoor ten aanzien van een zich ontwikkelende technologie geleidelijk en flexibel verder vormgegeven worden. Daarvoor is wel vereist dat de nodige rechtsontwikkeling plaatsheeft waarin bijvoorbeeld rechters en toezichthouders een op algoritmen toegesneden uitleg of interpretatie geven van de algemene kaders. De eerste voorzichtige en belangrijke stappen daartoe zijn in de rechtspraak en het toezicht reeds gezet.

Hoewel er geen structurele knelpunten zijn aan te wijzen, zijn er wel verschillende knelpunten geïdentificeerd die raken aan concrete normen in de algemene kaders. Zo kunnen de strenge normen met betrekking tot het gebruik van bijzondere persoonsgegevens, zoals gegevens over etnische afkomst en seksualiteit, juist in de weg staan aan het detecteren van discriminerende effecten in algoritmen. Ook is ten aanzien van art. 22 AVG niet duidelijk welke mate van het ontbreken van menselijke betrokkenheid nodig is om beschermd te zijn tegen volledig geautomatiseerde individuele besluitvorming en om een beroep te kunnen doen op bijbehorende informatierechten.

5. Conclusie

De eventuele regulering van algoritmen vraagt om een integrale (beleids)afweging, waarbij wordt geïdentificeerd welke van de relevante kansen en risico's voor publieke waarden en belangen zich voordoen en of het – in het licht van de behoeften, normen, waarden, belangen en context in een specifiek domein – mogelijk is om de risico's te mitigeren of te vermijden, terwijl de kansen wel kunnen worden gerealiseerd. De weging van kansen en risico's en het vinden van de balans daartussen is uiteindelijk een politiek en beleidsmatig proces. Het juridisch kader dat daarvan het resultaat is normeert en stuurt of en hoe algoritmen worden ingezet en heeft zo ook invloed op de mate waarin kansen en risico's worden gerealiseerd of vermeden.

Het is daarom van belang dat regelgeving zo wordt geformuleerd dat verantwoorde innovatie op het gebied van algoritmische besluitvorming mogelijk is. Voorkomen moet in het bijzonder worden dat normen te veel worden toegespitst op reeds bestaande technologieën en geen ruimte laten voor nieuwe ontwikkelingen. Dergelijke regelgeving biedt namelijk enkel rechtszekerheid en bescherming zolang de specifieke technologie ook daadwerkelijk gereguleerd wordt door de opgestelde regels. Als de technologieën zich ontwikkelen bestaat de mogelijkheid dat de die regelgeving niet meer actueel is, waardoor deze geen houvast meer biedt als nieuwe technologische ontwikkelingen zich voordoen. Bovendien kan het gebruik van technologiespecifieke regelgeving het zicht op de onderliggende uitgangspunten en beginselen

ontnemen, wat uiteindelijk de rechtsontwikkeling niet ten goede komt. De publieke waarden die de wetgever probeerde te waarborgen, komen dan door de snelle ontwikkeling van de technologie weer op het spel te staan.

Tot slot lijkt het in het licht van de bevindingen van dit onderzoek niet zinvol om het juridisch kader in te richten op algoritmische besluitvorming in algemene zin. De algemene kaders zoals die momenteel beschikbaar zijn, vertonen geen grote tekortkomingen of structurele problemen. Integendeel: hiervoor werd al vermeld dat ze een aanzienlijk absorberend vermogen hebben. Bij de ontwikkeling en inzet van nieuwe technologieën kunnen deze algemene kaders in belangrijke mate richting bieden. De casestudy's laten bovendien zien dat de kansen en risico's voor alle onderzochte publieke waarden en belangen sterk afhankelijk zijn van het domein en de organisatorische context waarin een algoritme wordt ingezet. Nadere algemene regelgeving heeft voor het bestrijden van risico's voor de publieke waarden dan ook nauwelijks meerwaarde. De kansen en risico's voor publieke waarden en belangen moeten vooral in het licht van de dynamiek in een domein en de verhoudingen tussen betrokken partijen in kaart worden gebracht en gewogen. Het voorgaande pleit er dan ook voor om, daar waar knelpunten worden ervaren, deze zoveel mogelijk domeinspecifiek aan te pakken. Alleen op die manier kan voldoende recht worden gedaan aan de specifieke kansen en risico's voor de specifieke publieke waarden en belangen die op het spel staan.