De legitimiteit van het algoritmisch bestuur

Een systematisch overzicht van bedreigingen en oplossingsrichtingen

Albert Meijer, Stephan Grimmelikhuijsen & Mark Bovens[1]

Het gebruik van algoritmen door overheden leidt tot rechtsstatelijke zorgen die variëren van het risico op discriminatie tot aantasting van de trias politica. Het juridisch, ethisch en sociaalwetenschappelijk onderzoek naar deze zorgen heeft geleid tot een rijk – maar ook gefragmenteerd – palet aan relevante inzichten. Dit artikel biedt een systematisch overzicht van wat het gebruik van algoritmen door de overheid betekent voor de legitimiteit van het staatsbestuur. Het analyseert wat de belangrijkste bedreigingen zijn en welke juridische, bestuurlijke en organisatorische maatregelen nodig zijn voor een legitiem gebruik van algoritmen door de overheid.

1. Zorgen over de inzet van algoritmen

De belofte van algoritmen is een efficiëntere en effectievere overheid. Deze belofte kan alleen worden waargemaakt als de inzet van algoritmen ook als legitiem wordt ervaren. Daaraan schort het op dit moment. Er leven in het veld veel rechtsstatelijke zorgen. Twee voorbeelden illustreren deze zorgen.

Voorbeeld 1. In het *Sensing Project* in Roermond experimenteert de politie met algoritmen gekoppeld aan sensoren die 'verdachte bewegingspatronen' registreren. Deze sensoren kunnen verdachte situaties blootleggen op basis van informatie over type personen, auto's, en land van herkomst. Als een voertuig door het algoritme als 'hoog risico' wordt aangemerkt komt er een melding bij de politie binnen en kan het voertuig staande worden gehouden. In een rapport uitte Amnesty International (2020) hier stevige kritiek op omdat dit systeem vooral tot discriminatie van mensen uit Oost-Europese landen zou leiden en er onvoldoende wettelijke bescherming is tegen dit soort praktijken.

Voorbeeld 2. Het *OxRec algoritme* wordt door de Reclassering gebruikt om rechters te adviseren over het recidiverisico van een verdachte. Een reclasseringsambtenaar vult gegevens in over geslacht, leeftijd, postcode, detentieduur, relatiestatus, inkomen, psychische aandoeningen en andere zaken. Gebaseerd op een bestaande dataset wordt aan de hand van de ingevulde gegevens een risicotaxatie gegeven: laag, middel of hoog risico op recidive. Deze *tool* wordt gebruikt in de rechtspraak, maar er worden grote vraagtekens geplaatst bij de kwaliteit van het voorspellingsmodel (Van Dijck, 2020) en de mogelijke ongelijke behandeling op basis van bijvoorbeeld ras of sociale klasse door het combineren van verschillende variabelen in het model (Braverman et al., 2016).

Ook vanuit de rechtswetenschap is er veel kritiek op het gebruik van dergelijke systemen. Bijlsma et al. (2019) geven aan dat voorspellende algoritmen niet waardenvrij zijn en dat daarom normatieve inbreng van juristen bij het ontwikkelen van dergelijke systemen vereist is (zie voor vergelijkbare zorgen: Van Eck, 2018). Vanuit staatsrechtelijk perspectief laten Passchier (2021) en Goossens et al. (2021) zien dat constitutionele beginselen, zoals democratie, de trias politica en rechtsgelijkheid, door geautomatiseerde besluiten onder druk komen te staan. De Poorter & Goossens (2019) stellen zelfs dat 'Het black box-karakter van algoritmen, en zeker van zelflerende, kan [...] leiden tot het ontstaan van een rechterlijk vacuüm.' Dit zou kunnen leiden tot een disbalans binnen de trias, waarbij de uitvoerende macht gebruik maakt van de 'kracht en macht' van technologie, maar waarbij de twee andere machten achterblijven in kennis en expertise (Prins, 2016; Passchier, 2020).

Doordat het gebruik van algoritmen vele rechtsgebieden en disciplines raakt, ontstaat echter ook steeds meer fragmentatie in het debat. Het vernieuwende van dit artikel schuilt niet zozeer in de rechtsstatelijke vragen die worden geïdentificeerd, maar veeleer in het systematische kader

Doordat het gebruik van algoritmen vele rechtsgebieden en disciplines raakt ontstaat steeds meer fragmentatie in het debat

waarin deze worden geplaatst. Dit artikel plaatst het begrip legitimiteit centraal als normatief kader voor het denken over gebruik van algoritmen door de overheid en laat zien wat de kritische analyses van juristen, maar ook van ethici en bestuurskundigen, betekenen voor die legitimiteit.

Legitimiteit is een kernbegrip in de sociale en juridische wetenschappen en wordt op vele manieren gedefinieerd (Bokhorst, 2014). In brede zin is legitimiteit het geïnternaliseerd gevoel van het willen gehoorzamen aan autoriteit zonder dat er daadwerkelijke dwang wordt uitgeoefend (Suchman, 1995). Dit 'geïnternaliseerde gevoel' komt op verschillende wijzen tot stand. Politicologen leggen het accent op de onderbouwing van macht via democratische processen (Easton, 1953), juristen benadrukken het belang van de legaliteit van autoriteit (Gribnau, 2001) en bestuurskundigen kijken steeds sterker naar de publieke waarde die autoriteiten creëren (Moore, 1995). Juist dit brede begrip van legitimiteit helpt om de diverse zorgen over algoritmen te ordenen en oplossingsrichtingen te identificeren. In dit artikel betogen wij dat niet alleen juridische kaders maar een veelheid aan nieuwe institutionele mechanismen nodig zijn om de legitimiteit van het algoritmisch bestuur te waarborgen.

De centrale vraag in dit artikel is hoe het gebruik van algoritmen de legitimiteit van het bestuur kan ondermijnen (*bedreigingen*). Op basis hiervan wordt bekeken welke institutionele mechanismen noodzakelijk zijn om de legitimiteit van het gebruik van algoritmes door overheidsorganisaties te versterken (*oplossingsrichtingen*). Onze ambitie is daarbij niet zozeer om oplossingen in detail uit te werken en specifieke juridische maatregelen te benoemen maar juist te laten zien dat een zeer diverse set van zowel juridische als bestuurlijke en organisatorische maatregelen nodig is om de legitimiteit te behouden van een bestuur dat steeds meer gebruik maakt van algoritmen.

2. Moderne algoritmen

In basale zin zijn algoritmen rekenregels en in dat opzicht zijn ze niet veel anders dan 'gewone' vormen van menselijke of geautomatiseerde besluitvormingsprocessen (Frissen, Van Eck & Drouen, 2019). De huidige, 'moderne', algoritmen verdienen echter wel bijzondere aandacht omdat ze werken op basis van *machine learning*. Deze algoritmen passen hun regels aan op basis van leerervaringen en dat betekent dat de programmeurs dus ook niet meer weten volgens welke regels het algoritme functioneert. Vetzo, Gerards & Nehmelman (2018, p. 48) karakteriseren *machine learning*-algoritmes als ondoorzichtige en niet-neutrale menselijke constructen. Overigens is het belangrijk te realiseren dat niet alle *machine learning*-algoritmen per definitie ondoorzichtig zijn. Voor gedetailleerdere discussie over de achterliggende technologie verwijzen we naar werk van Guidotti et al. (2018).

Machine learning-algoritmen worden geïntroduceerd om kennisintensieve overheidstaken te ondersteunen of zelfs over te nemen. Het gebruik van deze algoritmen is momenteel nog beperkt maar groeit snel. De verwachting is dat deze technologieën in de komende jaren een steeds belangrijkere rol gaan spelen in overheidsorganisaties (Vogl et al., 2020). Vier kenmerken zijn van belang voor de legitimiteit van het gebruik van deze algoritmen door de overheid: *complexiteit*, *onkenbaarheid*, *afhankelijkheid* en *waarschijnlijkheid*:

- De complexiteit heeft te maken met zowel de ingewikkelde technologische structuur – de duizenden regels programmeercode – als met de koppeling met specifieke juridische contexten van gebruik, zoals mogelijke discriminatie bij de politie en de berekening van de WOZ (Vetzo, Gerards & Nehmelman, 2018, p. 49).
- De onkenbaarheid heeft te maken met het *machine learning* waarmee de regels van het algoritme worden aangepast. Complexere en dynamischere voorspelmodellen die worden gevoed door verschillende grote datasets, zorgen voor gebrekkige transparantie. Hierdoor is het lastig om te weten wat de doorslaggevende variabelen zijn voor een bepaalde uitkomst (Burrell, 2016).
- De afhankelijkheid heeft te maken met het feit dat moderne algoritmen gebruik maken van verschillende datasets van eigen en andere organisaties (Van Eck, 2018). Kulk & Van Deursen (2020) beschrijven bijvoorbeeld hoe algoritmen leren op basis van allerlei databases en sensoren zoals camera's. Zie als voorbeeld het project Sensing in Roermond).
- De waarschijnlijkheid heeft er mee te maken dat een deel van de gebruikte algoritmen uitspraken doet over hoe waarschijnlijk een bepaalde uitkomst is op basis van een voorspellend model en bestaande data. Op basis hiervan wordt classificatie gegeven van bijvoorbeeld een recidiverisico. Denk hierbij aan het hiervoor genoemde OxRec-voorbeeld. Omdat het gaat om waarschijnlijkheden is er altijd een mate van onzekerheid over een voorspelde uitkomst.

Door het gebruik van *machine learning*-algoritmen met deze kenmerken verandert het functioneren van overheidsorganisaties. En deze veranderingen kunnen de legitimiteit van deze organisaties bedreigen. Voordat we ingaan op de aard van de bedreigingen – en ook mogelijke oplossingsrichtingen – karakteriseren we legitimiteit aan de hand van de drie dimensies.

3. *Input-, throughput-* en *output*-legitimiteit

In de politieke wetenschap wordt legitimiteit begrepen aan de hand van de *input*, *throughput* en *output* van een

Auteurs

1. Alle auteurs zijn verbonden aan het Departement Bestuurs- en Organisatiewetenschap van de Universiteit Utrecht. Prof. dr. A.J. Meijer is hoogleraar publieke innovatie, dr. S.G. Grimmelikhuijsen is universitair hoofddocent en prof. mr. dr. M.A.P. Bovens is hoogleraar bestuurskunde.

Een eerdere versie van dit artikel is gepresenteerd op de Staatsrechtconferentie 2020. Verder danken de auteurs Janneke Gerards, Floris Bex, Viola Bex-Reimert, Marlies van Eck en Stefan Kulk voor hun feedback en commentaar.

Programmering codes van onderdelen van algoritmen © Shutterstock

politiek-bestuurlijk systeem (Easton, 1953; Scharpf, 1999; Schmidt & Wood, 2019). De gedachte is dat burgers willen gehoorzamen aan de autoriteit van dit systeem wanneer (1) de preferenties van burgers adequaat worden vertaald naar de werking van dit systeem via verkiezingen en inspraak (*input*-legitimiteit), (2) het systeem functioneert volgens heldere juridische regels en er *checks & balances* bestaan om dit te controleren (*throughput*-legitimiteit) en (3) het politiek-bestuurlijk systeem wel de gewenste en niet de ongewenste uitkomsten oplevert voor burgers (*output*-legitimiteit). Deze drie dimensies van legitimiteit zullen we nader uitwerken omdat zij dienen als basis voor een analyse van de bedreigingen van algoritmen.

De eerste dimensie van legitimiteit is de *input* in het politiek-bestuurlijke systeem (Scharpf, 1999, p. 7-21). Eerlijke verkiezingen en een goed functionerende volksverte-

De legitimiteit van algoritmisch bestuur neemt toe naarmate de wijze wáárop via het algoritme uitkomsten worden gerealiseerd overeenkomt met eisen die democratisch en rechtsstatelijke instituties hieraan stellen

genwoordiging dragen bij aan de legitimiteit van het overheidsbestuur. Daarnaast is van belang dat burgers kunnen participeren in bestuurlijke processen die hen direct aangaan. De legitimiteit van algoritmisch bestuur neemt toe naarmate de preferenties van burgers via democratische processen beter zijn vertaald in het ontwerp en het gebruik van een algoritme door een overheidsorganisatie.

De tweede dimensie van legitimiteit is de *throughput* van het politiek-bestuurlijke systeem. Schmidt & Wood (2019) geven aan dat een behoorlijk en *fair* handelen van de overheid, naast een transparante verantwoording over organisationele processen, cruciaal zijn voor de legitimiteit van het staatsbestuur. Voor deze dimensie geldt dat de legitimiteit van algoritmisch bestuur toeneemt naarmate de wijze waaróp via het algoritme uitkomsten worden gerealiseerd overeenkomt met eisen die democratisch en rechtsstatelijke instituties hieraan stellen.

De derde dimensie van legitimiteit is de *output* van het politiek-bestuurlijke systeem in de vorm van de publieke waarde die overheidsorganisaties leveren voor de samenleving. Voor deze dimensie geldt dat de legitimiteit van algoritmisch bestuur toeneemt naarmate de uitkomsten van het gebruik van een algoritme meer bijdragen aan de realisatie van waarden die door burgers als belangrijk worden beschouwd. Scharpf (1999) noemt als voorbeeld dat de Europese Unie legitimiteit heeft door haar bijdrage aan vrede en veiligheid in Europa.

Deze drie dimensies van legitimiteit kunnen worden gebruikt om systematisch te analyseren welke bedreigingen er zijn voor de legitimiteit van algoritmisch bestuur en welke oplossingsrichtingen mogelijk zijn.

Deze drie dimensies lopen niet altijd synchroon. Algoritmisch bestuur kan heel effectief zijn, en dus hoge *output*-legitimiteit hebben, maar toch niet voldoen aan basale eisen van behoorlijkheid en daarmee laag scoren op *throughput*-legitimiteit.

4. Algoritmen en bedreigingen voor *input*-legitimiteit

Een bedreiging van de *input*-legitimiteit ontstaat als de verbinding tussen democratie en de inzet van algoritmen onduidelijk is. Een fundament van democratisch bestuur is dat in de verbinding tussen politiek en bestuur de vertaling plaatsvindt van de preferenties van burgers naar bestuurlijke actie via democratisch gekozen volksvertegenwoordigers of via directe participatie (Demir & Nyhan, 2008). Door het complexe karakter van algoritmen bestaat het risico dat algoritmen alleen op basis van technische overwegingen vorm krijgen en los staan van het democratische proces. Deze bedreiging geldt in algemene zin voor complexe overheidsdossiers, maar door het gebruik van complexe, zelflerende algoritmen kan de disconnectie tussen democratische besluitvorming en uitvoering versterkt worden.

Een eerste bedreiging betreft *de uitholling van democratische toezicht op besluitvorming door gekozen volksvertegenwoordigers*. Door de technocratische complexiteit is er beperkt democratisch toezicht op het ontwerp van algoritmen (Meijer, Ruijer & Dekker, 2020; Van Est et al., 2020). Het beperkte inzicht in technologie van zowel de ministers die uiteindelijk verantwoordelijk zijn voor de algoritmen als van de volksvertegenwoordigers die hen daarop kunnen aanspreken maakt het voor hen zeer lastig om de relevant politieke vragen te stellen over deze algoritmen. Aanvullend geldt dat veel politiek relevante beslissingen impliciet worden genomen door de ontwikkelaars van algoritmen.

In zijn klassieke werk over informatiesystemen in de publieke sector geeft Lessig (1999) aan dat in computerprogramma's juridische eisen impliciet vorm krijgen. 'Code is Law,' noemt hij dit en hij benadrukt dat de ontwikkelaar van programmeercodes onbewust de rol van regelgever aanneemt. Ook zijn deze ontwikkelaars zich vaak weinig bewust van het feit dat deze technische regels een politiek-bestuurlijke betekenis hebben (Van Eck et al., 2018). En ook wordt het democratisch toezicht ondermijnd doordat gebruik wordt gemaakt van commerciële algoritmen waarvan de werking niet bekend is omdat deze valt onder het bedrijfsgeheim (bijvoorbeeld Brayne, 2021, p. 135). Verder geldt dat vanwege *machine-learning*-algoritmen zich steeds verder ontwikkelen. Hierdoor kan volgens Goossens, Hirsch Ballin & Van Vught (2021) de relatie vertroebelen tussen enerzijds een concreet besluit op basis van een algoritmische beslisregel en

Veel politiek relevante beslissingen worden impliciet genomen door ontwikkelaars van algoritmen

Tabel 1. Dimensies van legitimiteit van algoritmisch bestuur

Dimensie	Legitimiteit van algoritmisch bestuur
Input	De legitimiteit van algoritmisch bestuur neemt toe naarmate de preferenties van burgers via democratische processen beter zijn vertaald in het ontwerp en het gebruik van het algoritme.
Throughput	De legitimiteit van algoritmisch bestuur neemt toe naarmate de wijze waaróp via het algoritme uitkomsten worden gerealiseerd overeenkomt met eisen die democratische en rechtsstatelijke instituties hieraan stellen.
Output	De legitimiteit van algoritmisch bestuur neemt toe naarmate het gebruik van het algoritme meer bijdraagt aan de realisatie van waarden die door burgers als belangrijk worden beschouwd.

anderzijds de algemene regels die op basis van democratische processen tot stand zijn gekomen. Het risico bestaat dat het algoritme nieuwe patronen ontwikkelt zonder dat hierop democratische controle plaatsvindt (De Poorter & Goossens, 2019).

Om de democratische controle op algoritmische besluitvorming te versterken en de legitimiteit van de overheid te behouden dient het politieke toezicht op algoritmen versterkt te worden (Frissen et al., 2019, p. 3-4). Belangrijk is dat bestuurders en volksvertegenwoordigers beter in staat zijn om politiek gevoelige issues zoals privacy, discriminatie of eigenaarschap over data te identificeren (Van Est et al., 2020). Dit vergt een deskundige ambtelijke ondersteuning van de volksvertegenwoordiging. Ook dient de politieke gevoeligheid van de ambtelijke aansturing van algoritmen te worden versterkt. Voor de aansturing van ontwerpprocessen is het belangrijk dat ambtenaren kunnen identificeren welke aspecten van algoritmen dusdanig gevoelig zijn dat deze moeten worden voorgelegd aan politieke besluitvormers. Verder geldt dat commerciële algoritmen alleen gebruikt mogen worden wanneer volstrekte helderheid bestaat over de werking. En als reactie op het doorontwikkelen van algoritmen door *machine-learning* kan vereist worden dat algoritmen zelf uitleggen op welke wijze zij de regels hebben doorontwikkeld: uitleggende algoritmen ('Explainable Artificial Intelligence' – XAI). De term XAI wordt vaak gebruikt om aan te geven hoe algoritmen individuele besluiten uitleggen maar kan ook verwijzen naar een uitlegbaarheid op het niveau van modellen (Samek et al., 2019). Het algoritme geeft dan aan op welke wijze regels op basis van ervaringen zijn aangepast opdat deze aanpassingen vervolgens democratisch kunnen worden gecontroleerd.

Een tweede bedreiging van de *input*-legitimiteit is *het ontbreken van directe responsiviteit van de besluitvorming*. Bij ontwikkeling en toepassing van algoritmen in publieke organisaties ontbreekt directe participatie van burgers. Op beleidsterreinen zoals de ruimtelijke ordening is de participatie van burgers formeel geregeld door de zienswijzenprocedure waarbij iedereen mag inspreken. Ook in andere publieke sectoren bestaan medezeggenschaps- en cliëntenraden opdat burgers kunnen participeren in besluitvorming die hen treft (Michels & De Graaf, 2010). Het gebruik van algoritmen kan burgers direct treffen maar door het technische karakter van deze systemen komt vaak de mogelijkheid van participatie door burgers niet eens ter sprake omdat dit wordt beschouwd als een kwestie van technische uitvoering.

Voor het behoud van de *input*-legitimiteit van de overheid is versterking nodig van directe participatie van burgers in het ontwerp van algoritmen (Brayne, 2021, p. 147). Via participatie kunnen burgers zelf of via actiegroepen aangeven waar gevoelige issues liggen en welke aandachtspunten van belang zijn zoals bijvoorbeeld potentiële stigmatisering van burgers bij gebruik van algoritmen in de sociale zekerheid. Het delen van codes via platforms zoals Github, zoals recentelijk is gebeurd bij het ontwikkelen van de CoronaMelder, kan hieraan bijdragen wanneer dit burgers met voldoende technische kennis in staat stelt te participeren.

5. Algoritmen en bedreigingen voor *throughput*-legitimiteit

De *throughput*-legitimiteit van de overheid kan worden bedreigd als de besluitvorming met behulp van algoritmen niet voldoet aan de eisen van behoorlijkheid en wanneer er onvoldoende *checks & balances* aanwezig zijn. De vraag is of kernbeginselen van behoorlijk bestuur, zoals het motiveringsbeginsel of het zorgvuldigheidsbeginsel, voldoende getoetst kunnen worden wanneer een besluit grotendeels is genomen door een *machine learning*-algoritme. Voor veel standaardgevallen levert de motivering waarschijnlijk geen problemen op, maar juist waar er bijzondere omstandigheden zijn is maatwerk – en dus een goede motivering – noodzakelijk (Van Eck, Bovens en Zouridis 2018). Ook zijn er grote vragen over de mogelijkheden om algoritmische besluitvorming te controleren wanneer de werking hiervan niet transparant is.

Een eerste bedreiging voor de *throughput*-legitimiteit is de *aantasting van een correcte rechtsgang* (Widlak, Van Eck & Peeters, 2021). In algemene zin gaat het om een gebrekkige vertaling van algemene en specifieke juridische eisen in het ontwerp en het gebruik van algoritmen waardoor het procedurele verloop van besluitvorming wordt bedreigd (Goossens et al., 2021). Rechtsstatelijke eisen, zoals het recht op non-discriminatie en op rechtsbescherming, zijn hierbij van groot belang. Algoritmen wor-

Tabel 2. ***Input*-legitimiteit: bedreigingen en oplossingsrichtingen**

Bedreigingen	Oplossingsrichtingen

• Uitholling van sturing door de volksvertegenwoordiging op algoritmische besluitvorming
  • Politieke besluiten door ontwikkelaars van algoritmen zonder politiek toezicht
  • Privatisering van besluitvorming bij gebruik van commerciële algoritmen
  • Algoritmen ontwikkelen door, buiten het bereik van politiek toezicht
• Beperkte directe responsiviteit van de algoritmische besluitvorming
  • Geen burgerparticipatie in het ontwerp van algoritmen
	• Versterken van sturing door de volksvertegenwoordiging op algoritmische besluitvorming
	  • Versterken van politieke gevoeligheid van ambtenaren die algoritmen aansturen en versterken politiek toezicht hierop
	  • Versterken democratische sturing op aanschaf van commerciële algoritmen
	  • Uitleggende algoritmen (XAI) geven aan hoe regels evolueren
	• Versterken directe responsiviteit van algoritmische besluitvorming
	  • Participatie in het ontwerp en het monitoren van algoritmen

Juridische eisen zijn vaak niet eenduidig waardoor omzetting in een algoritmische regel complex of misschien in bepaalde gevallen fundamenteel onmogelijk is

den ontworpen door technische experts die vaak beperkte kennis hebben van juridische regels en daarmee ontstaat het risico dat een algoritme wordt ontwikkeld dat niet voldoet aan de bredere juridische eisen, bijvoorbeeld op het gebied van mensenrechten (Vetzo et al., 2018). Een complicerende factor is daarbij dat juridische eisen vaak niet eenduidig zijn waardoor omzetting in een algoritmische regel complex of misschien in bepaalde gevallen fundamenteel onmogelijk is. Meer specifiek geldt hier ook de dreiging dat overheden met het gebruik van algoritmen inbreuk maken op de privacy van burgers, in het bijzonder hun recht op persoonsgegevensbescherming (Young, Katell & Krafft, 2019). Hier vormt de Algemene Verordening Gegevensbescherming (AVG) een belangrijk juridisch ankerpunt. De toepassing van algoritmen vergt dat er veel data worden gecombineerd en verwerkt uit verschillende datasets, die nooit zijn verzameld met dit doel. De combinaties van verschillende datasets kan bovendien leiden tot nieuwe inzichten over individuele personen (Kulk & Van Deursen, 2020, p. 5). Ook is voor individuen vaak onbekend welke data nu precies op welke manier worden verwerkt. De privacy van burgers is bijvoorbeeld in het geding wanneer er geanonimiseerde data worden gebruikt, omdat de combinatie van data uit verschillende sets kan leiden tot de identificatie van unieke individuen.

Om deze bedreiging van de legitimiteit te voorkomen kan een algemene juridische toets op algoritmisch bestuur worden ingevoerd, eventueel gecombineerd met extern toezicht. De juridische toets wordt uitgevoerd voorafgaand aan het gebruik van een algoritme door de overheidsorganisatie. De toets betreft een serie specifieke vragen, bijvoorbeeld over bewaring en openbaarheid van gegevens, beveiliging, etc. (zie ook Kulk & Van Deursen, 2020), maar ook bredere vragen gerelateerd aan mensenrechten. Een externe toezichthouder (Tutt, 2017) of eventueel de bestaande toezichthouders (Kulk & Van Deursen, 2020) kunnen de uitvoering van dergelijke toetsen controleren. Dit externe toezicht kan ook reageren op specifieke klachten van burgers over algoritmen.

Voor de privacybescherming bestaat reeds een Data Protection Impact Assessment (DPIA). Een DPIA is een organisatorisch instrument om vooraf de privacy-risico's van gegevensverwerking in kaart te brengen zodat betere besluiten kunnen worden genomen over de inzet van een algoritme (Bu-Pasha, 2020). Het uitvoeren van een DPIA is overigens vaak geen vrije keuze: onder de AVG zijn overheidsorganisaties verplicht een DPIA uit te voeren wanneer de organisaties systematisch en uitgebreid persoonlijke aspecten evalueren gebaseerd op geautomatiseerde verwerking en daarop besluiten baseren die gevolgen hebben voor mensen.

Een tweede bedreiging voor de *throughout*-legitimiteit betreft *het ontbreken van checks en balances rond de besluitvorming* (Passchier, 2021). Daarbij wordt veel gewezen op de beperkte transparantie van algoritmen. Van Eck (2018) laat zien dat beslisregels in algoritmen vaak onzichtbaar zijn voor zowel de persoon die met een algoritme een besluit neemt als voor controlerende instanties. Kulk & Van Deursen (2020, p. 5) benadrukken dat hierdoor het recht op rechtsbescherming in het gedrang kan komen. Ook geldt dat algoritmen vaak zijn ontwikkeld door commerciële partijen die over intellectuele eigendomsrechten beschikken die in de weg kunnen staan aan de openbaarmaking van de code (Mittelstadt et al., 2016). De Poorter & Goossens (2019) spreken zelfs de zorg uit dat er een rechterlijk vacuüm kan ontstaan door het *black box*-karakter van algoritmen. Deze risico's worden versterkt door het verdampen van verantwoordelijkheid in de complexe relaties rondom het gebruik van algoritmen (De Fine Licht & De Fine Licht, 2020). Men zou zeggen dat degene die de algoritmen gebruikt verantwoordelijk blijft maar dat is lastig als deze het algoritme niet volledig kan doorgronden. De vraag ontstaat dan of niet de ontwikkelaar van het systeem verantwoordelijk blijft. Tegelijkertijd geldt dat het systeem zich verder ontwikkelt op basis van processen van *machine-learning*. Dan geldt misschien dat degenen die het algoritme 'trainen' verantwoordelijk blijven, of degenen die de datasets beheren waarmee het algoritme wordt getraind, of een combinatie daarvan. Zeker wanneer steeds meer data worden gedeeld en gebruikt voor andere doeleinden dan oorspronkelijk bedacht bestaat het risico dat in deze complexe verhoudingen uiteindelijk niemand meer echt verantwoordelijk kan worden gehouden (Bovens, 1998).

Een logische reactie op deze bedreiging van de legitimiteit betreft het vergroten van de transparantie van algoritmisch bestuur (Meijer & Grimmelikhuijsen, 2020). In de literatuur over algoritmen is de transparantie ervan wellicht het meest besproken punt: benadrukt wordt dat de uitlegbaarheid en toegankelijkheid moeten worden vergroot (Lepri et al, 2018). Uitlegbaarheid betekent dat duidelijk gemaakt moet worden dat de substantiële redenen voor een besluit helder en correct zijn (Tutt, 2017). Toegankelijkheid betekent dat zowel de beslisregels als de gebruikte data vrij worden gegeven (Mittelstadt et al., 2016). De Fine Licht & De Fine Licht (2020) benadrukken dat het erom gaat dat duidelijk is welk besluit is genomen, op basis van welke argumenten en wie ervoor verantwoordelijk is. Young, Katell & Krafft (2019, p. 2) geven aan dat er manieren moeten worden gevonden om algoritmische systemen 'leesbaar' te maken voor beleidsmakers en stakeholders.

Om deze bedreiging van de legitimiteit te voorkomen dienen de verantwoordelijkheden voor algoritmisch bestuur te worden geëxpliciteerd (Busuioc, 2020). Waar liggen de verantwoordelijkheden van het bestuursorgaan, van de ontwikkelaar van het algoritme, van de toeleveranciers van datasets, van degene die het algoritme onderhoudt, etc.? Dit vergroot de aanspreekbaarheid en verkleint de kans dat verschillende betrokkenen naar elkaar

Tabel 3. *Throughput*-legitimiteit: bedreigingen en oplossingsrichtingen

Bedreigingen	Oplossingsrichtingen

• Algoritmische besluitvorming voldoet niet aan eisen van behoorlijkheid (due process/good governance):
  • Inbreuk op de privacy van burgers
  • Gebrekkige vertaling van juridische eisen in algoritmen
• Checks en balances rond de besluitvorming zijn onvoldoende:
  • Beperkte transparantie van algoritmen
  • Verdampen van verantwoordelijkheden rondom algoritmen
	• Waarborgen van de behoorlijkheid van algoritmische besluitvorming
	  • Data Protection Impact Assessment
	  • Juridische toets op algoritmisch bestuur en onafhankelijke toezichthouder
	• Waarborgen van checks and balances voor algoritmische besluitvorming
	  • Transparantie van algoritmisch bestuur
	  • Verhelderen van verantwoordelijkheden bij algoritmisch bestuur

kijken en verwachten dat eisen door anderen worden verwerkt.

6. Algoritmen en bedreigingen voor de *output*-legitimiteit

Een eerste bedreiging voor de output-legitimiteit is dat de *besluitvorming niet effectief en efficiënt verloopt*. Zo gaat de inzet van algoritmen gepaard met hoge kosten maar leidt dit niet altijd tot de beloofde verbeteringen. De verwachtingen van de inzet van algoritmen zijn vaak hoog, maar het harde bewijs voor hun werkzaamheid wordt niet altijd duidelijk geleverd. Vaak lijkt er eerder sprake van het volgen van een trend en een groot geloof in technologische mogelijkheden dan dat er sprake is van een heldere *business case*. Ook kunnen de effecten van een algoritme in de loop van de tijd verschuiven. Het kan zijn dat er een *bias* ontstaat doordat het algoritme leert op basis van bepaalde datasets. Hierdoor kan een algoritme rekenregels ontwikkelen die een *bias* bevatten of om andere redenen resulteren in onwenselijke uitkomsten.

Het verplichten van een kosten/baten-analyse van algoritmen is een strategie om met deze dreiging om te gaan. Bij de analyse van de financiële en niet-financiële kosten en baten dienen verschillende expertises en stakeholders te worden betrokken, zodat vooral ook de mogelijke onwenselijke bijeffecten goed in beeld worden gebracht. Aanvullend is een periodieke audit van belang om te checken of het gebruik van het algoritme nog leidt tot de gewenste uitkomsten en of ongewenste uitkomsten worden vermeden. Frissen et al. (2019, p. 4) pleiten voor een periodieke audit door een externe partij op onder andere de werking van de algoritmen, het dienen van het doel waarvoor ze worden ingezet, en het in acht nemen van de menselijke maat. Deze audit kan de vorm krijgen van de eerdergenoemde DPIA of de juridische toets, maar kan ook breder gaan over de wenselijkheid van de uitkomsten van het gebruik van het algoritme.

Een tweede bedreiging voor de *output*-legitimiteit is dat *besluitvorming leidt tot ongewenste uitkomsten*. Een veel genoemde ongewenste uitkomst is dat een *bias* optreedt in de besluitvorming wanneer deze wordt ondersteund door algoritmische systemen. In de literatuur worden onder andere de volgende soorten van *bias* genoemd (Jackson, 2018): focus op specifieke doelgroepen, focus op specifieke gebieden en focus op '*past performance*' in plaats van op de huidige en toekomstige situatie. Deze *bias* kan op verschillende wijzen ontstaan: door gebruik van vertekende data waarmee het algoritme wordt getraind; het gebruik van selectieve data, waarbij dus de *bias* optreedt bij het selecteren van data; of door incorrecte analyses dan wel interpretaties, waarbij de *bias* dus pas na de analyse optreedt.

Voor deze bedreiging worden in de literatuur verschillende oplossingen genoemd. Om de *bias* in algoritmen te minimaliseren is het van belang om kritische tegendenkers vanuit diverse achtergronden in projectteams te betrekken die steeds aan kunnen geven welke vormen van *bias* kunnen ontstaan. Ook is het bij het gebruik van algoritmen van groot belang om allerlei mogelijke vormen van *bias* te meten om te controleren dat het algoritme niet vertekent. Kulk & Van Deursen (2020, p. 4-5) merken op dat, mits de *bias* in data en algoritme kan worden voorkomen, het gebruik van algoritmen op zichzelf de mogelijkheid biedt om non-discriminatie te versterken omdat goed-geprogrammeerde en gevalideerde algoritmen in beginsel beter dan mensen in staat zijn om zonder aanziens des persoons een besluit te nemen. Daarom wordt aanbevolen om voor de introductie van een algoritme een '*algorithmic impact assessment*' uit te voeren (Brayne, 2021, p. 146).

Een ander ongewenst effect is dat het algoritme de kwaliteit van menselijke contacten ondermijnt (WRR, 2016, p. 131). In de meest vergaande vorm heeft een burger alleen met een algoritme te maken, zoals nu bijvoorbeeld al bij beslissingen die massaal worden genomen over studiefinanciering of verkeersboetes (Van Eck, 2018). Ook bij het gebruik van algoritmes op de achtergrond kan gelden dat de rol van de menselijke besluitvormer wordt gereduceerd tot het toepassen van het advies van het algoritme.

Wettelijk dient te zijn vastgelegd dat een menselijke besluitvormer het besluit van het algoritme moet kunnen heroverwegen: algoritmen moeten *aanvechtbaar* zijn

Tabel 4. *Output*-legitimiteit: bedreigingen en oplossingsrichtingen

Bedreigingen	Oplossingsrichtingen

• Algoritmische besluitvorming is niet effectief en efficiënt
  • Algoritmische systemen zijn kostbaar maar leveren weinig op
  • Baten en lasten van algoritme verschuiven in de tijd
• Algoritmische besluitvorming leidt tot ongewenste uitkomsten
  • Bias in algoritmische systemen
  • Algoritmen beperken het menselijk contact
	• Meer aandacht voor effectiviteit en efficiëntie van algoritmische besluitvorming
	  • Verplichte kosten/baten-analyses van het algoritme
	  • Periodieke audits op het gebruik van algoritmen
	• Ongewenste effecten van algoritmische besluitvorming voorkomen
	  • Kritische tegendenkers in ontwerpteams en algoritmic impact assessment
	  • Richtlijnen voor recht op menselijk contact en aanvechtbaarheid

Hiervoor kan het nuttig zijn om – in aanvulling op het reeds bestaande kader van de AVG – regels voor het recht op menselijk contact nader uit te werken. Vooral in het geval van het gebruik van algoritmen die beslissingen nemen over individuele burgers is het van groot belang dat burgers contact kunnen krijgen met een menselijke besluitvormer (Frissen et al., 2019, p. 4). Daarnaast dient wettelijk te zijn vastgelegd dat een menselijke besluitvormer ook het besluit van het algoritme moet kunnen heroverwegen: algoritmen moeten *aanvechtbaar* zijn (Almadal, 2019). Hoewel burgers gebruik kunnen maken van een standaard bezwaar- en beroepsprocedure om besluiten aan te vechten, garandeert dit nog onvoldoende een rechtvaardige uitkomst van een algoritme in individuele gevallen (zie ook Van Eck et al. 2018; Goossens et al., 2021).

7. Conclusies

Een zorgvuldige inzet van algoritmen kan op termijn de legitimiteit van het staatsbestuur verhogen (Wolswinkel, 2020, p. 53-55). Voor we zover zijn moet een fors aantal hobbels worden genomen. Een eerste moeilijkheid betreft de vele verbindingen tussen algoritmen. In de praktijk gaat het vaak om een veelheid van verbonden algoritmen en datasets die elkaar op allerlei wijzen beïnvloeden. Hier speelt het probleem van de vele handen. Deze datasets en algoritmen worden gebruikt door verschillende organisaties – soms zelfs buiten de overheid – en het uiteindelijk functioneren hangt af van deze interacties. Een tweede moeilijkheid betreft het internationale karakter van de technologie-ontwikkeling. Veel technologische componenten zijn niet in Nederland maar in andere landen ontwikkeld en specifiek inzicht in het functioneren van deze technologische componenten is soms lastig te verkrijgen.

De vormgeving van legitimiteit van algoritmisch bestuur is daarom een opgave van institutioneel ontwerp. Hoewel Kulk & Van Deursen (2020, p. 6-7) terecht opmerken dat algemene juridische kaders een voldoende basis bieden is er ook een nadere invulling nodig. Er zal een veelheid aan nieuwe institutionele mechanismen moeten worden ontwikkeld die waarborgen dat het gebruik van algoritmen niet leidt tot een uitholling van de legitimiteit van het bestuur. Deze institutionele mechanismen betreffen soms regelgeving (zoals de eis tot een juridische toets op algoritmen), soms het uitbreiden van organisationele capaciteiten (zoals het vermogen tot politiek toezicht op algoritmisch bestuur) en soms ook aandacht voor nieuwe manieren van werken (bijvoorbeeld de politieke sensitiviteit van ambtenaren). Dit geheel aan mechanismen is nodig om de legitimiteit van algoritmische bestuur te behouden en te vergroten. •

Referenties

M. Almadal, 'Human intervention in automated decision-making: Toward the construction of contestable systems', *Seventeenth International Conference on Artificial Intelligence and Law* (ICAIL 2019), 17-21 juni 2019, Montreal, QC, Canada, New York: ACM 2019, 10 p.

Amnesty International, *We sense trouble. Automated discrimination and mass surveillance in predictive policing in the Netherlands*, London: Amnesty International 2020. Beschikbaar op: www.amnesty.nl/content/uploads/2020/09/Report-Predictive-Policing-RM-7.0-FINAL-TEXT_CK-2.pdf?x30221

A.M. Bokhorst, *Bronnen van legitimiteit: over de zoektocht van de wetgever naar zeggenschap en gezag*, Den Haag: Boom Juridische uitgevers 2014.

M.A.P. Bovens, *The quest for responsibility: Accountability and citizenship in complex organisations*, Cambridge: Cambridge University Press 1998.

S. Brayne, *Predict and surveil: Data, discretion, and the future of policing*, New York: Oxford University Press 2021.

S. Bu-Pasha, 'The controller's role in determining 'high risk' and data protection impact assessment (DPIA) in developing digital smart city', *Information & Communications Technology Law* 2020, 29, 3, p. 391-402.

J. Burrell, 'How the machine 'thinks': Understanding opacity in machine learning algorithms', *Big Data & Society*, 2016, 3, 1, 2053951715622512.

D.W. Braverman, S.N. Doernberg, C.P. Runge & D.S. Howard, 'OxRec model for assessing risk of recidivism: ethics', *The Lancet Psychiatry*, 2016, 3, 9, p. 808-809.

M. Busuioc, 'Accountable Artificial Intelligence: Holding Algorithms to Account', *Public Administration Review* 2000, https://doi.org/10.1111/puar.13293.
K. de Fine Licht & J. de Fine Licht, 'Artificial intelligence, transparency, and public decision-making', *AI & Society* 2020, p. 1-10.
T. Demir & R.C. Nyhan, 'The politics-administration dichotomy: An empirical search for correspondence between theory and practice', *Public Administration Review*, 2008, 68, 1, p. 81-96.
G. van Dijck (2020), 'Algoritmische risicotaxatie van recidive. Over de Oxford Risk of Recidivism tool (OXREC), ongelijke behandeling en discriminatie in strafzaken', *NJB* 2020/1558, afl. 25, p. 1784-1790.
J. van Dijck, T. Poell & M. de Waal, *The platform society: Public values in a connective world*, Oxford: Oxford University Press 2018.
D. Easton, *The Political System: An Inquiry into the State of Political Science*, New York: Knopf 1953.
M. van Eck, *Geautomatiseerde ketenbesluiten & rechtsbescherming: Een onderzoek naar de praktijk van geautomatiseerde ketenbesluiten over een financieel belang in relatie tot rechtsbescherming* (diss. Tilburg), Tilburg: Tilburg University 2018.
M. van Eck, S. Zouridis & M. Bovens, 'Algoritmische rechtstoepassing in de democratische rechtsstaat', *NJB* 2018/2101, afl. 40, p. 3008-3017.
R. van Est, E. Bakker, J. van den Broek, J. Deuten, P. Diederen, I. van Keulen, I. Korthagen & H. Voncken, *Waardevol digitaliseren. Hoe lokale bestuurders vanuit publiek perspectief mee kunnen doen aan het 'technologiespel'*, Den Haag: Rathenau Instituut 2000.
V. Frissen, M. van Eck & T. Drouen, *Toezicht op het gebruik van algoritmen door de overheid*, Den Haag: Hooghiemstra & Partners 2019.
J. Goossens, E.H. Ballin & E. van Vugt, 'Algoritmische beslisregels vanuit constitutioneel oogpunt: Tweedeling tussen algemene regels en concrete toepassing onder druk', *Tijdschrift voor constitutioneel recht* 2021, 12, 1, p. 4-19.
J.L.M. Gribnau, Legaliteit en legitimiteit, *NTB* 2001, 1, p. 9-19.
R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti & D. Pedreschi, 'A survey of methods for explaining black box models', *ACM computing surveys* (CSUR) 2018, 51, 5, p. 93.
J.R. Jackson, 'Algorithmic bias. Journal of Leadership', *Accountability and Ethics* 2018, 15, 4, p. 55-65.
S. Kulk & S. van Deursen, *Juridische aspecten van algoritmen die besluiten nemen. Een verkennend onderzoek*, Den Haag: WODC 2020.
A. Meijer & S. Grimmelikhuijsen, 'Responsible and Accountable Algorithmization: How to Generate Citizen Trust in Governmental Usage of Algorithms', in: M. Schuilenburg & R. Peeters (eds.), *The Algorithmic Society: Technology, Power and Knowledge*, London: Routledge 2020.
A. Meijer E. Ruijer & R. Dekkers, *Navigatiestrategie. Lessen uit drie casusstudies over de kennispositie van de Tweede Kamer op het gebied van digitalisering*. Rapport voor de Tijdelijke Commissie Digitale Toekomst van de Tweede Kamer, Utrecht: Universiteit Utrecht 2020.
A. Michels & L. de Graaf, 'Examining citizen participation: Local participatory policy making and democracy', *Local Government Studies* 2010, 36, 4, p. 477-491.
M.H. Moore, *Creating public value: Strategic management in government*, Cambridge MA: Harvard University Press 1995.
R. Passchier, 'Digitalisering en de (dis) balans binnen de trias politica', *Ars Aequi* 2020, 69, 10, p. 916-927.
R. Passchier, *Artificiële intelligentie en de rechtsstaat: Over verschuivende overheidsmacht, Big Tech en de noodzaak van constitutioneel onderhoud*, Den Haag: Boom Juridisch 2021.
J. Poorter & J. Goossens, 'Effectieve rechtsbescherming bij algoritmische besluitvorming in het bestuursrecht', *NJB* 2019/2777, afl. 44, p. 3303-3312.
J.E.J. Prins, 'Digitale (dis) balans binnen de TRIAS', *NJB* 2016/682, afl. 14, p. 909.
W. Samek, G. Montavon, A. Vedaldi, L.K. Hansen & K.R. Müller, (eds.), *Explainable AI: interpreting, explaining and visualizing deep learning*, Cham, Zwitserland: Springer Nature 2019.
F.W. Scharpf, *Governing in Europe: Effective and democratic?*, Oxford: Oxford University Press 1999.
V. Schmidt & M. Wood, 'Conceptualizing throughput legitimacy: Procedural mechanisms of accountability, transparency, inclusiveness and openness in EU governance', *Public Administration* 2019, 97, 4, p. 727-740.
M.C. Suchman, 'Managing legitimacy: Strategic and institutional approaches', *Academy of management review* 1995, 20, 3, p. 571-610.
A. Tutt, 'An FDA for algorithms', *Administrative Law Review* 2017, 69, 1, p. 83-124.
M.J. Vetzo, J.H. Gerards & R. Nehmelman, *Algoritmes en grondrechten*, Den Haag: Boom juridisch 2018.
T.M. Vogl, C. Seidelin, B. Ganesh, J. & Bright, 'Smart Technology and the Emergence of Algorithmic Bureaucracy: Artificial Intelligence in UK Local Authorities', *Public Administration Review* 2020 (online), DOI: https://doi-org.proxy.library.uu.nl/10.1111/puar.13286.
A. Widlak, M. van Eck & R. Peeters, 'Towards principles of good digital administration: Fairness, Accountability and Proportionality in Automated Decision Making', in: Marc Schuilenburg & Rik Peeters (eds.), *The Algorithmic Society. Technology, power and knowledge,* London: Routledge 2020, p. 67-83.
J. Wolswinkel, *Willekeur of algoritme? Laveren tussen analoog en digitaal bestuursrecht* (oratie Tilburg), Tilburg: Tilburg University 2020.
WRR, *Big Data in een vrije en veilige samenleving*, Amsterdam: AUP 2016.
M. Young, M. Katell & P.M. Krafft, 'Municipal surveillance regulation and algorithmic accountability', *Big Data & Society* 2019, 6, 2, 2053951719868492.