

Vergaderjaar 2018–2019

26 643

Informatie- en communicatietechnologie (ICT)

32 761

Verwerking en bescherming persoonsgegevens

Nr. 570

BRIEF VAN DE MINISTER VOOR RECHTSBESCHERMING

Aan de Voorzitter van de Tweede Kamer der Staten-Generaal

Den Haag, 9 oktober 2018

Tijdens het VAO van 6 juni jl. over «Big Data en de bescherming van persoonsgegevens» heeft het lid Verhoeven van uw Kamer een motie ingediend waarin de regering in de eerste plaats wordt verzocht na te streven dat de werking en de broncode van door overheden gebruikte algoritmes en analysemethoden die een aanmerkelijke impact op burgers hebben openbaar zijn, inclusief informatie over gebruikte databestanden. In de motie wordt de regering voorts verzocht bij algoritmes en analysemethoden die om zwaarwegende redenen niet openbaar gemaakt kunnen worden een technische audit uit te voeren om te controleren of de algoritmes en analysemethoden niet onbedoeld discrimineren of andere negatieve effecten hebben die al dan niet wet- en regelgeving schenden, en het resultaat van die audit openbaar te maken.¹

Op mijn verzoek is deze motie aangehouden onder de toezegging dat ik uw Kamer in het derde kwartaal van dit jaar een brief zou schrijven over wat het effect van uitvoering van deze motie zou zijn.² Daarbij heb ik aangegeven te willen kijken of tegemoet kan worden gekomen aan het principe dat in de motie is verwoord, en, voor zover dat niet kan, of op de desbetreffende algoritmes dan een kwaliteitstoets kan worden uitgevoerd. Dit spoort met het door mij al verwoorde uitgangspunt dat algoritmes bij de overheid zo transparant mogelijk dienen te zijn.³ Met deze brief doe ik deze toezegging gestand.

In deze brief geef ik, mede namens de Staatssecretaris van Binnenlandse Zaken en Koninkrijksrelaties, allereerst een beeld van typen algoritmes die er zijn, en van het gebruik daarvan in de praktijk. Aansluitend zet ik uiteen wat onder transparantie van algoritmes kan worden verstaan. Daarna beschrijf ik welke verplichtingen op dit punt uit de huidige wet- en regelgeving voortvloeien. Tegen de achtergrond van deze verplichtingen

¹ Kamerstuk 32 761, nr. 117.

² Handelingen II 2017/18, nr. 90, item 3, p. 4–5.

³ Kamerstukken 26 643 en 32 761, nr. 543, p. 28.

benoem ik vervolgens een aantal niet-juridische aspecten waarmee met betrekking tot transparantie rond algoritmes ook rekening dient te worden gehouden. Tot slot trek ik uit een en ander een aantal conclusies voor de uitvoering van de motie.

Deze brief staat niet op zichzelf. Over transparantie van algoritmes wordt ook in ander verband nagedacht. Ik noem in de eerste plaats de activiteiten van de Europese Commissie op het gebied van artificiële intelligentie (AI). Zij zal begin 2019 met een Charter over AI en ethiek komen, waarin naar verwachting ook op het aspect van transparantie zal worden ingegaan. Verder wijs ik op een aantal trajecten in nationaal verband, die mede kunnen dienen om een standpunt over het Charter van de Europese Commissie te bepalen:

- een verkennend onderzoek van JenV naar de juridische aspecten van algoritmes die zelfstandig besluiten nemen,
- een onderzoek van BZK naar de impact van digitalisering op de effectiviteit van de bescherming van de grondrechten,
- een verkennend onderzoek van BZK naar algoritmes die door overheden worden ingezet,
- de voorbereiding van een beleidsvisie van BZK op het vrijgeven van overheidssoftware als open source software,
- de adviesaanvraag van het kabinet aan de Wetenschappelijke Raad voor het Regeringsbeleid (WRR) over kunstmatige intelligentie⁴ en
- het advies van de Raad van State van 31 augustus jl. over de effecten van digitalisering voor de rechtsstatelijke verhoudingen⁵, waarop van kabinetszijde binnenkort een reactie aan de Tweede Kamer zal worden gezonden.

Al deze trajecten kunnen ook nieuwe inzichten opleveren met betrekking tot de vraag in hoeverre algoritmes die de overheid gebruikt, transparant kunnen en moeten zijn. Dit is inherent aan de snelheid waarmee zich nieuwe ontwikkelingen voordoen die relevant voor dit vraagstuk zijn. In zoverre vormen de standpunten in deze brief in zekere zin niet meer dan een momentopname.

Typen algoritmes en gebruik daarvan in de praktijk

Om het effect van uitvoering van de motie goed te kunnen beoordelen is het wenselijk eerst stil te staan bij wat we onder een algoritme verstaan, welke typen algoritmes er zijn en hoe deze kunnen worden ingezet.

Een algoritme is een set instructies om een bepaalde taak uit te voeren. Aan de hand van data die in het algoritme worden ingevoerd, wordt in verschillende stappen toegewerkt naar het beoogde eindresultaat, bijvoorbeeld het toekennen van een toeslag. Op basis van dat resultaat kan men dan actie ondernemen (toeslag wel of niet toekennen). De precieze stappen die worden doorlopen verschillen per algoritme. In het algemeen geldt dat deze stappen variabelen bevatten, bijvoorbeeld de hoogte van het inkomen als relevante factor voor het toekennen van de toeslag.

Er bestaan verschillende typen algoritmes. Het algoritme voor het toekennen van een toeslag heeft doorgaans het karakter van een eenvoudige beslisboom (*rule based*) met een beperkt aantal variabelen en drempelwaarden, bijvoorbeeld dat om voor een toeslag in aanmerking te

⁴ Deze trajecten komen uitgebreider aan bod in: Kamerstuk 26 643, nr. 541, bijlage (Nederlandse Digitaliseringsstrategie), p. 44, en Kamerstuk 26 643, nr. 549, bijlage (Agenda Digitale Overheid «NL DIGibeter»), p. 26, 28.

⁵ Stcrt. 2018, nr. 50999.

komen het inkomen lager dient te zijn dan bedrag x. Een algoritme kan ook meer complex zijn en bijvoorbeeld op basis van een aantal casussen voorspellingen doen over nog niet bekende gevallen (*case based*) of gebruikt worden voor gezichts- of objectherkenning. We betreden dan de wereld van *machine learning*, *deep learning*, zelflerende algoritmes en kunstmatige intelligentie.

Als we *deep learning* daar uit lichten, zien we een technologie die gebruik maakt van neurale netwerken. Deze bestaan uit verschillende lagen kunstmatige neuronen die onderling waarden uitwisselen. Hoe meer lagen, hoe complexere vraagstukken het netwerk kan oplossen. Deze netwerken worden op dezelfde manier getraind als andere zelflerende algoritmes, maar zijn in staat complexe verbanden te vinden die een mens zelf moeilijk of niet kan vinden. Bij *deep learning* is het redeneerproces inherent ondoorzichtig en voor een mens lastig of niet te begrijpen.

De Staatssecretaris van Binnenlandse Zaken en Koninkrijksrelaties is op dit moment bezig een verkennend onderzoek te laten uitvoeren naar het gebruik van algoritmes door de overheid. Dit onderzoek wordt naar verwachting voor het eind van dit jaar afgerond. De bevindingen uit het onderzoek zullen naar de Kamer worden gezonden.⁶

Daarop vooruitlopend valt nu wel al op te merken dat algoritmes in de vorm van eenvoudige beslisbomen door de overheid al enige tientallen jaren worden gebruikt. Denk hierbij aan algoritmes die grote uitvoeringsorganisaties van de rijksoverheid gebruiken om beslissingen te nemen over verlening van vergunningen of uitkeringen, of voor toezicht op de naleving van daaraan gestelde voorwaarden. Bij dergelijke eenvoudige algoritmes kan het per organisatie al gauw om enige honderden tot zelfs duizenden algoritmes gaan. Het lijkt hier veelal om simpele algoritmes te gaan die een één-op-één vertaling vormen van wetgeving en beleidsregels, die beide gepubliceerd zijn. Openbaarmaking van de beslisregels van het algoritme zou in dergelijke gevallen weinig tot niets aan de gewenste transparantie toevoegen, maar wel extra uitvoeringslasten meebrengen. Het gebruik van meer complexe algoritmes door de overheid komt ook voor, met dien verstande dat het gebruik van *deep learning* nog in de kinderschoenen staat.

Transparantie: technische transparantie en uitlegbaarheid

Als we in meer algemene zin kijken naar de mate waarin en de wijze waarop algoritmes transparant gemaakt kunnen worden, is het van belang te beseffen dat transparantie geen doel op zichzelf is, maar een middel: transparantie kan bijdragen aan het vertrouwen dat men in het gebruik van algoritmes door de overheid moet kunnen hebben, en aan de mogelijkheden om zich eventueel tegen een uitkomst van het gebruik daarvan te weer te stellen.

Met dit voor ogen is het nuttig onderscheid te maken tussen «technische transparantie» en «uitlegbaarheid».⁷

Bij «technische transparantie» gaat het om wat we in theorie transparant kunnen maken. In algemene zin gaat het hierbij om publicatie van de broncode van de gebruikte algoritmes en analysemethoden, inclusief de gebruikte databestanden. Bij zelflerende algoritmes gaat het bovendien

⁶ Kamerstuk 26 643, nr. 527.

⁷ Dit onderscheid treffen we ook aan in een recent rapport van the House of Lords, «AI in the UK: ready, willing and able?», waarin wordt gesproken over «technical transparency» en «explainability» (<https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf>).

om transparantie van de trainingsdata. Deze kunnen iets zeggen over de context waarin het algoritme het best kan worden toegepast. Als de data niet representatief zijn voor het toepassingsveld of als er in het verleden sprake is geweest van discriminatie en dit in de trainingsdata wordt weerspiegeld, kan het algoritme minder betrouwbaar zijn of zelfs discriminerend. Daarnaast kan transparantie worden gezocht in het model. Dat kan een eenvoudige beslisboom zijn, maar ook een ingewikkeld neurale netwerk. Het gebruikte model kan iets zeggen over het redeneerproces dat aan de uitkomst ten grondslag ligt. Ook de variabelen die zijn meegegeven in het model, kunnen transparant worden gemaakt. Zij kunnen iets zeggen over de kenmerken of indicatoren die zijn meegenomen in het redeneerproces. Aan transparantie kan verder worden bijgedragen door het zekerheidspercentage van de uitkomst te noemen (bijvoorbeeld: «deze persoon is met 70% zekerheid geschikt voor de functie»), de drempelwaarden en de geaccepteerde foutmarge. Tot slot kan inzichtelijk worden gemaakt hoe het algoritme is gevalideerd en wat hiervan de uitkomst was.

Bij uitlegbaarheid gaat het erom dat de uitkomst van een algoritme in begrijpelijke taal kan worden verklaard. Technische transparantie kan hieraan tot op zekere hoogte bijdragen, maar kent ook haar grenzen. Zo kan sprake zijn van een zeer grote hoeveelheid variabelen die worden meegewogen. Sommige algoritmes kunnen duizenden kenmerken meenemen in hun redeneerproces. De uitkomst van het algoritme is dan gebaseerd op zoveel factoren dat het lastig is om in te zien welke factoren doorslaggevend zijn geweest voor de uitkomst. Daarmee wordt het ook lastig om de conclusie te doorgronden of beïnvloeden. In de tweede plaats zijn sommige modellen die bij zelflerende algoritmes worden gebruikt, inherent ondoorzichtig. Dit is met name het geval bij *deep learning* neurale netwerken. Wel zijn er technieken ontwikkeld om achteraf te achterhalen op welke informatie een algoritme zijn uitkomst baseert. Deze technieken zijn op dit moment vaak gebaseerd op sensitiviteitsanalyse. Dit houdt in dat de invoerdata steeds een klein beetje worden aangepast en dat vervolgens wordt gekeken in hoeverre dit een verandering in de uitvoer teweegbrengt. Dit is in het algemeen een tijdrovend proces. Tegelijkertijd is de techniek volop in ontwikkeling, omdat alle partijen (bedrijven, wetenschappers, overheden) het belang zien van navolgbare uitkomsten van algoritmes. Tegen deze achtergrond wordt bij uitlegbaarheid van algoritmes de focus gelegd op het beschrijven van het doel dat met het algoritme wordt nagestreefd, welke variabelen doorslaggevend zijn geweest voor de uitkomst, het type gegevens dat wordt gebruikt, en de eventuele beslisregels. Dit laat uiteraard onverlet dat, zoals hierna nog aan de orde zal komen, er in het licht van de bestaande wet- en regelgeving gronden kunnen zijn om de mate van uitlegbaarheid te beperken.

Wettelijke verplichtingen met betrekking tot transparantie van algoritmes

Voor een beoordeling van de motie is mede van belang na te gaan welke verplichtingen voor de overheid met betrekking tot de transparantie van algoritmes uit de wet- en regelgeving voortvloeien.

Uit recente jurisprudentie over het algemene bestuursrecht blijkt dat, als een door een bestuursorgaan genomen besluit geheel of ten dele het resultaat is van een geautomatiseerd proces en de belanghebbende de juistheid van de bij dat geautomatiseerde proces gemaakte keuzes en van de daarbij gebruikte gegevens en aannames wil controleren en zo nodig gemotiveerd betwisten, het bestuursorgaan moet zorgdragen voor de inzichtelijkheid en controleerbaarheid van die keuzes, aannames en

gegevens. Zonder die inzichtelijkheid en controleerbaarheid dreigt een ongelijkwaardige procespositie van partijen te ontstaan. Een belanghebbende kan in geval van besluitvorming op basis van een computerprogramma dat vanuit zijn perspectief is te beschouwen als een «black box» immers niet controleren op basis waarvan tot een bepaald besluit wordt gekomen.⁸ De genoemde verplichting om gegevens openbaar te maken geldt niet zonder meer voor gebruikte standaardinvoergegevens, maar wel voor gebruikte maatwerkgegevens, waarbij voldoende is dat in of met het besluit duidelijk is gemaakt welke keuzen bij de invoer zijn gemaakt ten aanzien van dergelijke maatwerkgegevens.⁹

Met betrekking tot de vraag in hoeverre het gebruik van algoritmes door de overheid transparant dient te zijn, moeten we, indien we bij het gebruik van algoritmes persoonsgegevens verwerken, ook een blik werpen op voorschriften uit de Algemene verordening gegevensbescherming (AVG), die op 25 mei jl. van toepassing is geworden.¹⁰

Een belangrijk beginsel uit de AVG is dat persoonsgegevens verwerkt moeten worden op een wijze die ten aanzien van betrokkenen transparant is (artikel 5, eerste lid, onder a). Dit beginsel wordt uitgewerkt in verschillende informatieverplichtingen. Een algemeen kader daarvoor is te vinden in artikel 12 AVG. Als een overheidsdienst met behulp van algoritmes persoonsgegevens verwerkt, vloeit uit dat artikel voort dat zij over deze verwerking informatie dient te verschaffen die beknopt, transparant, begrijpelijk en gemakkelijk toegankelijk moet zijn. Deze informatie moet in duidelijke en eenvoudige taal worden verschaft, vooral als de informatie voor kinderen is bestemd.¹¹ Als de verwerking het karakter van geautomatiseerde besluitvorming heeft, schrijft de AVG bovendien voor dat de overheidsdienst betrokkene daarover informeert. Zij dient ten minste in het geval dat deze besluitvorming gebaseerd is op profilering, ook nuttige informatie te verschaffen over de onderliggende logica en het belang en de verwachte gevolgen van die verwerking voor de betrokkene (artikelen 13, tweede lid, onder f, en 14, tweede lid, onder g, AVG).¹² Bij «onderliggende logica» zou het bijvoorbeeld kunnen gaan om de beoordelingscriteria die het algoritme gebruikt, de werkwijze van het algoritme of het proces waarin het algoritme wordt gebruikt, afhankelijk van wat voor soort besluit het is en de wijze van gebruik.

De transparantie- en informatieverplichtingen uit de AVG zijn primair bedoeld om betrokkenen te informeren, dat wil zeggen personen wier persoonsgegevens worden verwerkt, teneinde hen in staat te stellen de rechten uit te oefenen die de AVG hun geeft om hun persoonsgegevens te beschermen (inzagerecht, correctierecht etc.). Uit de AVG vloeit dus niet voort dat deze informatie openbaar wordt gemaakt voor het publiek. Integendeel, als het informatie over niet louter de onderliggende logica maar de verwerkte persoonsgegevens zelf betreft, is verstrekking daarvan zonder een wettelijke grondslag als bedoeld in artikel 6, eerste lid, AVG,

⁸ HR 17 augustus 2018, ECLI:NL:HR:2018:1316, r.o. 2.3.3. In hiermee vergelijkbare zin: ABRvS 17 mei 2017, ECLI:NL:RVS:2017:1259 (PAS-zaak).

⁹ ABRvS 18 juli 2018, ECLI:NL:RVS:2018:2454.

¹⁰ Buiten beschouwing wordt hier gelaten dat sommige verwerkingen van persoonsgegevens niet onder de AVG, maar onder de Richtlijn gegevensbescherming opsporing en vervolging (PbEU 2016, L 119) vallen, of anderszins buiten de werking van de AVG vallen, zoals verwerkingen door Defensie en de inlichtingen- en veiligheidsdiensten.

¹¹ Deze voorwaarden zijn nader uitgewerkt in de Guidelines on transparency under Regulation 2016/679 (AVG) van de Article 29 Data protection working party, par. 6–18.

¹² Overigens is op grond van de artikel 22, eerste lid van de AVG het in beginsel niet toegestaan om uitsluitend op basis van geautomatiseerde verwerkingen te komen tot een besluit; er dient sprake te zijn van menselijke tussenkomst als er een besluit wordt genomen waaraan rechtsgevolgen zijn verbonden dan wel indien het betrokkene anderszins in aanmerkelijke mate treft.

niet toegestaan. Maar zelfs informatie over de onderliggende logica behoeft vanuit het perspectief van de AVG niet in alle gevallen waarbij sprake is van geautomatiseerde besluitvorming, aan betrokkenen te worden verstrekt, laat staan openbaar te worden gemaakt voor het publiek. Artikel 41 van de Uitvoeringswet AVG (UAVG) noemt in navolging van artikel 23 AVG verschillende gronden waarop (onder meer) de informatieverplichtingen buiten toepassing kunnen worden gelaten, zoals de nationale veiligheid, de openbare veiligheid en de voorkoming, het onderzoek, de opsporing en de vervolging van strafbare feiten. Dit artikel biedt overigens geen grondslag voor het structureel en categorisch beperken van bepaalde rechten van betrokkenen maar dient louter als vangnet voor een incidenteel geval waarin een afwijking noodzakelijk is. Als er wel sprake is van een noodzaak om structureel af te wijken van de transparantieplicht, dan zal daarvoor een wettelijke grondslag moeten worden gecreëerd welke voldoet aan de vereisten van artikel 23 AVG. Tegen de achtergrond van mijn toezegging tijdens het VAO aan het Kamerlid Van Dam om in deze brief ook aandacht te besteden aan de vraag of de nationale veiligheid openbaarmaking van algoritmes in de weg kan staan¹³, blijkt uit het voorgaande dat in ieder geval de AVG en de UAVG de nationale veiligheid als grond noemen om de informatieverplichting buiten toepassing te laten.

In die gevallen dat er geen uitzondering is gemaakt ten aanzien van de informatieverplichtingen uit de AVG, kan het uit een oogpunt van doelmatigheid zinvol zijn dat een overheidsdienst informatie over het gebruik van algoritmes in een privacystatement op zijn website opneemt. Op deze wijze kunnen niet alleen alle betrokkenen gezamenlijk worden geïnformeerd, maar ook andere burgers, van wie de persoonsgegevens mogelijk later met gebruik van algoritmes worden verwerkt. Deze informatie kan eventueel «gelaagd» op de website worden geplaatst, waarbij men snel kan navigeren naar relevante passages, zonder door het gehele privacystatement te hoeven scrollen.¹⁴ In dit kader kan ook worden overwogen om een op grond van artikel 35 AVG uitgevoerde gegevensbeschermingseffectbeoordeling met betrekking tot de desbetreffende gegevensverwerking te publiceren.¹⁵ Het is uiteindelijk aan de Autoriteit persoonsgegevens en aan de rechter om te beoordelen of de geboden transparantie op grond van de AVG voldoende is.

Transparantie van het gebruik van algoritmes door de overheid kan ook worden ingegeven door de Wet openbaarheid van bestuur (Wob). Een ieder kan op grond van artikel 3 van die wet om openbaarmaking van onder de overheid berustende informatie over bestuurlijke aangelegenheden voor een ieder vragen. Daaronder kan ook informatie worden begrepen over het gebruik van algoritmes die (persoons)gegevens verwerken. Ingevolge artikel 8 dient een bestuursorgaan deze ook uit eigen beweging openbaar te maken, zodra dat in het belang is van een goede en democratische bestuursvoering. Wel bevatten de artikelen 10 en 11 Wob gronden om het verstrekken van informatie (gedeeltelijk) te weigeren. In het licht van mijn eerdergenoemde toezegging aan het Kamerlid Van Dam merk ik op dat openbaarmaking van informatie bijvoorbeeld achterwege blijft, voor zover dit de veiligheid van de Staat zou kunnen schaden. Andere gronden voor weigering zijn onder meer

¹³ Handelingen II 2017/18, nr. 90, item 3, p. 5.

¹⁴ Guidelines on transparency under Regulation 2016/679 (AVG) van de Article 29 Data protection working party, par. 30.

¹⁵ Idem, par. 35.

gelegen in het belang van de opsporing en vervolging van strafbare feiten en het belang van inspectie, controle en toezicht door bestuursorganen.¹⁶

Een en ander samenvattend kunnen de volgende conclusies worden getrokken:

1. Als een door een bestuursorgaan genomen besluit geheel of ten dele het resultaat is van een geautomatiseerd proces en de belanghebbende de juistheid van de bij dat geautomatiseerde proces gemaakte keuzes en van de daarbij gebruikte gegevens en aannames wil controleren en zo nodig gemotiveerd betwisten, moet het bestuursorgaan zorgdragen voor de inzichtelijkheid en controleerbaarheid van die keuzes, aannames en gegevens.
2. Als een overheidsdienst met behulp van algoritmes persoonsgegevens verwerkt en deze verwerking het karakter van geautomatiseerde besluitvorming heeft, dient zij betrokkene daarover te informeren. Zij dient, tenminste in het geval dat deze besluitvorming mede gebaseerd is op profilering, ook nuttige informatie te verschaffen over de onderliggende logica en het belang en de verwachte gevolgen van die verwerking voor de betrokkene. Uit een oogpunt van doelmatigheid mag de dienst ervoor kiezen de informatie over het gebruik van algoritmes in een privacystatement op zijn website op te nemen. De verstrekking van informatie kan achterwege blijven als bepaalde, in de AVG of UAVG opgenomen gronden (waaronder nationale veiligheid) daartoe noodzaken.
3. Op grond van de Wob kan het geboden zijn om algoritmes die de overheid gebruikt, openbaar te maken, maar openbaarmaking kan achterwege blijven als sprake is van een of meer in die wet genoemde weigeringsgronden.

Het is van belang bij deze conclusies te benadrukken dat de genoemde verplichtingen in de gevallen, genoemd onder 1 en 2, betrekking hebben op deels geautomatiseerde besluitvorming.¹⁷ Dat impliceert dat de overheid in andere gevallen waarin zij algoritmes gebruikt, niet op voorhand gehouden is deze inzichtelijk en controleerbaar te maken c.q. nuttige informatie over de onderliggende logica te verschaffen. Te denken valt aan gevallen waarin de overheid toezicht houdt en in dat kader met behulp van algoritmes risico's taxeert dat personen zich niet aan de wet houden. Het gebruik van dergelijke risicotaxatie-instrumenten kan bijdragen aan een efficiëntere en effectievere inzet van capaciteit. In dergelijke gevallen zou wel met een beroep op de Wob om informatie kunnen worden gevraagd, maar zou de overheid het verzoek mogelijk op een van de uitzonderingsgronden uit die wet mogen afwijzen.

Overige aspecten met betrekking tot transparantie van algoritmes

Zoals ik aan het begin van deze brief al heb verwoord, onderschrijf ik het uitgangspunt dat algoritmes die de overheid gebruikt, zo transparant mogelijk dienen te zijn. De verschillende departementen moeten daarvoor zelf passende maatregelen treffen. Met het oog daarop kunnen aanbestedende overheidsdiensten eisen stellen aan bedrijven met betrekking tot

¹⁶ Met het oog op de transparantie van het gebruik van algoritmes door de overheid kan tot slot het initiatiefvoorstel voor een Wet open overheid nog relevant worden. Omdat de initiatiefnemers en het kabinet over dit wetsvoorstel nog overleg voeren, zou het echter prematuur zijn in deze brief aan te geven wat de gevolgen van dat voorstel zouden zijn voor de mate van transparantie van algoritmes die de overheid gebruikt (Kamerstuk 33 328, K).

¹⁷ In dit verband is van belang dat de Staatssecretaris van Binnenlandse Zaken en Koninkrijksrelaties nog een brief aan de Kamer zal zenden waarin hij zal ingaan op de vraag of het mogelijk en noodzakelijk is de openbaarmaking van de broncode van die geautomatiseerde besluitvorming in meer gevallen mogelijk te maken (Aanhangsel Handelingen II 2017/18, 1612, p. 2).

het bieden van transparantie in de formules en algoritmes die ze gebruiken.¹⁸ Bij dit al dient in de eerste plaats rekening te worden gehouden met de hiervoor vermelde wettelijke verplichtingen die met betrekking tot transparantie van algoritmes gelden. Er zijn echter ook nog andere, niet juridische aspecten die in dit verband zullen moeten worden meegewogen. Die aspecten komen in deze paragraaf aan bod.

Nu algoritmes door de opkomst van Big Data en de toename in rekenkracht steeds complexer worden, neemt de behoefte toe om te begrijpen hoe algoritmes tot een bepaalde uitkomst komen. Er zijn echter ook verscheidene situaties waarin deze behoefte niet lijkt te bestaan. Te denken valt aan algoritmes die worden gebruikt bij de programmering van onderhoud van infrastructuur, het voorkomen van files of het voorspellen van het weer.

Voor zover er wel behoefte aan transparantie van algoritmes bestaat, is het wenselijk nader te specificeren waarop deze berust. Zo kan er behoefte zijn aan meer inzicht in de betrouwbaarheid van het algoritme, bijvoorbeeld in hoe goed het algoritme toepasbaar is in de context waarin het gebruikt wordt. Transparantie kan in dat geval de controleerbaarheid vergroten en daarmee het vertrouwen in het algoritme. De behoefte aan transparantie kan ook berusten op de vraag hoe uit het algoritme een bepaald resultaat is voortgekomen. Zijn de uitkomsten van het algoritme eerlijk of is er mogelijk sprake van discriminatie? Daarbij kan het uitmaken of het algoritme onderdeel van besluitvorming is of sturingsinformatie voor het optreden van de overheid geeft. Als het algoritme onderdeel van besluitvorming is, is bovendien relevant wat de impact op de burger is. Heeft de inzet van het algoritme rechtsgevolgen voor de burger of anderszins een aanmerkelijk impact? Of heeft het dergelijke gevolgen niet?

Het hangt van de specifieke behoefte af, welke aspecten transparant gemaakt moeten worden. Zo kan transparantie van de trainingsdata en de variabelen inzicht bieden in de eerlijkheid van het algoritme. Als in het verleden sprake is geweest van discriminatie en dit in de trainingsdata zit, kan het gevolg zijn dat het algoritme ook discrimineert. Indien er vooral behoefte is aan een uitlegbaar algoritme, kan worden gedacht aan transparantie van het model (hoe redeneert het algoritme?), de variabelen (wat wordt meegewogen?) en de sensitiviteitsanalyse (wat was doorslaggevend?). Naarmate de impact van de uitkomst van een algoritme op een burger groter is, zal de behoefte aan transparantie van een of meer van genoemde aspecten navenant groter zijn.

Een eventuele technische transparantie- of uitlegbaarheidseis voor (bepaalde) algoritmes kan ook keerzijdes hebben, die in vijf categorieën kunnen worden ingedeeld:

1. *Uitlegbaarheid kan de toegevoegde waarde van algoritmes beperken.* Juist complexe algoritmes bieden tal van maatschappelijke kansen. Zo kunnen algoritmes worden ingezet bij het verminderen van verkeersdrukke en bij het verbeteren en betaalbaar houden van de gezondheidszorg. Deze kansen kunnen beperkter worden verzilverd wanneer algoritmes uitlegbaar moeten worden gemaakt: uitlegbaarheid kan grenzen stellen aan de complexiteit die een algoritme heeft en daarmee aan de mate waarin algoritmes toegevoegde waarde kunnen hebben. Hierbij geldt wel dat de techniek volop in ontwikkeling is en veel bedrijven en wetenschappers sterk inzetten op algoritmes die uitlegbaar zijn.

¹⁸ Kamerstukken 26 643 en 32 761, nr. 543

2. *Technische transparantie kan ten koste gaan van innovatie.* Technische transparantie van het algoritme kan bedrijfsgeheimen schaden en kan daardoor ten koste gaan van innovatie. Het vooruitzicht van exclusief bezit van een algoritme geeft namelijk een prikkel om te investeren in de ontwikkeling van algoritmes. Dat geldt ook voor algoritmes die bedrijven in opdracht van de overheid ontwikkelen.
3. *Technische transparantie kan de privacy schaden.* Het transparant maken van de trainingsdata van algoritmes kan op gespannen voet staan met privacywetgeving. In sommige gevallen bestaan de trainingsdata namelijk uit persoonsgegevens. Dergelijke data kunnen geanonimiseerd en geaggregeerd transparant gemaakt worden. Dit beperkt echter de mate waarin transparantie het algoritme werkelijk uitlegbaar maakt.
4. *Technische transparantie en uitlegbaarheid kunnen de mogelijkheden voor derden vergroten om de uitkomsten van het algoritme ongewenst te beïnvloeden.* Naarmate een algoritme meer transparant is, kan dit mogelijkheden bieden om uitkomsten te beïnvloeden. Denk aan beroepsfraudeurs die zich met kennis van bepaalde drempelwaarden in een risicotaxatie-instrument dat door een overheidsdienst wordt gebruikt, aan toezicht en opsporing kunnen onttrekken («gaming the system»).
5. *Uitvoeringslasten.* Een transparantie- of uitlegbaarheidseis in aanvulling op de al bestaande informatieverplichtingen brengt ontegenzeggelijk extra uitvoeringslasten mee. Dat geldt zowel voor eenvoudige algoritmes als voor complexe algoritmes. De uitvoeringslasten worden in het eerste geval vooral ingegeven door de relatief grote hoeveelheid van eenvoudige algoritmes, in het tweede geval door het relatief grote volume aan elementen waaruit complexe algoritmes bestaan.

Hoe de weging tussen een gespecificeerde behoefte aan transparantie en de bij deze «keerzijdes» genoemde tegenargumenten moet uitpakken, hangt sterk af van de context waarbinnen deze moet plaatsvinden. In sommige situaties zal de behoefte aan transparantie groot zijn, omdat de werking van het algoritme een directe impact op mensen kan hebben, zoals bij het toekennen of intrekken van een uitkering of bij toepassing in het toezicht, de opsporing en de rechtspraak. Daar staat tegenover dat juist in het toezicht en de opsporing openbaarmaking van variabelen en drempelwaarden bezwaarlijk kan zijn, omdat daarmee bijvoorbeeld de mogelijkheden voor derden zouden worden vergroot om de uitkomsten van het algoritme ongewenst te beïnvloeden.

Conclusies met betrekking tot de motie-Verhoeven c.s.

In deze brief heb ik opnieuw het uitgangspunt onderschreven dat algoritmes die de overheid gebruikt, zo transparant mogelijk moeten zijn. Dit komt overeen met het uitgangspunt dat aan de motie ten grondslag ligt. De vraag die dan rijst is hoe dit uitgangspunt het best kan worden ingevuld.

Zoals beschreven, hanteren overheidsdiensten soms wel honderden tot zelfs duizenden algoritmes. Het lijkt hier veelal om betrekkelijk simpele algoritmes in de vorm van een beslisboom te gaan die een één-op-één vertaling vormen van wetgeving en beleidsregels, die beide gepubliceerd zijn. Openbaarmaking van de beslisregels van het algoritme zou in dergelijke gevallen weinig tot niets toevoegen, maar wel extra uitvoeringslasten meebrengen. Dit effect mag naar mijn oordeel niet worden onderschat.

De overheid gebruikt echter ook complexere algoritmes, die niet louter een één-op-één-vertaling van wetgeving en beleidsregels geven. Vooral met het oog op dergelijke algoritmes is het van belang dat een bestuursorgaan in geval van geheel of gedeeltelijk geautomatiseerde besluitvorming verplicht is zorg te dragen voor de inzichtelijkheid en controleerbaarheid van in deze algoritmes gemaakte keuzes, gedane aannames en gebruikte gegevens. Juist dan kan het algoritmes betreffen waarop de motie betrekking heeft: algoritmes die een aanmerkelijke impact op burgers hebben. Overigens geldt dat verstrekking van informatie over gebruikte algoritmes achterwege kan blijven als bepaalde in wetgeving opgenomen gronden (waaronder nationale veiligheid) daartoe noodzaken.

Buiten gevallen waarin sprake is van een proces waarin geheel of gedeeltelijk geautomatiseerde besluitvorming plaatsvindt, is de overheid niet op voorhand gehouden om algoritmes transparant te maken. Te denken valt aan het gebruik van risicotaxatie-instrumenten, die sturingsinformatie voor het handelen van de desbetreffende overheidsdienst geven.

Met het oog op transparantie van algoritmes is verder van belang dat uitlegbaarheid van het algoritme, dat wil zeggen de verklaring daarvan in begrijpelijke taal, de voorkeur kan hebben boven «technische» transparantie waarbij de gebruikte broncode en analysemethoden en de gebruikte databestanden volledig openbaar worden gemaakt. Andere aspecten die bij het transparant maken van algoritmes dienen te worden meegewogen, zijn onder meer de vraag waarop de behoefte aan transparantie precies berust. Als het vooral om inzicht in de eerlijkheid (het niet-discriminerende effect) van het algoritme gaat, kan transparantie van de trainingsdata en de variabelen daar aan bijdragen. Er dienen echter ook aspecten te worden meegewogen die mogelijk tot de conclusie leiden dat de transparantie van een algoritme beperkt moet blijven, bijvoorbeeld om de mogelijkheden voor burgers te beperken om de uitkomsten van het algoritme ongewenst te beïnvloeden. Hoe de weging tussen de in de brief genoemde aspecten uitpakt, hangt sterk af van de context waarbinnen deze moet plaatsvinden en zal daarom per geval verschillen.

Bij deze conclusies wil ik het niet laten. Zo is op dit moment een tweetal werkgroepen bezig met het opstellen van richtlijnen voor het inzichtelijk maken van algoritmes die de overheid gebruikt voor toezichthouders en de rechtspraak, respectievelijk voor publieksvoorlichting over Big Data analyses door de overheid. In eerstbedoelde richtlijnen zal, zoals in de motie wordt gevraagd, ook aandacht worden geschonken aan het uitvoeren van technische audits met betrekking tot de kwaliteit van algoritmes. Verder is een werkgroep bezig met een onderzoek naar mogelijke wettelijke kwaliteitswaarborgen rond het gebruik van Big Data door de overheid.¹⁹

In het licht van deze trajecten acht ik het nu nog te vroeg om een oordeel over de motie te geven. Om dat te kunnen doen wil ik eerst de resultaten van eerder genoemde werkgroepen afwachten. Ik verwacht u in het begin van 2019 daarover te kunnen informeren en u dan van een oordeel over de motie in kennis te kunnen stellen. Met het oog daarop verzoek ik u om de motie tot dat moment nog aan te houden.

De Minister voor Rechtsbescherming,
S. Dekker

¹⁹ Het gaat hier om trajecten ter uitvoering van het kabinetsstandpunt over het rapport «Big Data in een vrije en veilige samenleving» van de WRR (Kamerstukken 26 643 en 32 761, nr. 426, p. 9).