

Grip op algoritmische besluitvorming bij de overheid

De rol van de Eerste Kamer



Notitie ter ondersteuning van de werkgroep AI

In de samenleving en de politiek is er groeiende aandacht voor zowel de mogelijkheden van als de zorgen over de inzet van artificiële intelligentie (AI) binnen de overheid. Met behulp van AI is het mogelijk om verbanden en inzichten naar boven te halen die voor de mens moeilijk vindbaar zijn. In het kader van Justitie en Veiligheid (JenV) biedt dit kansen op het gebied van bijvoorbeeld opsporing en intelligence. Het vergt inspanning om dergelijke AI-toepassingen op een verantwoorde manier te ontwikkelen en te implementeren. De commissie JenV heeft daarom de werkgroep AI ingesteld, om te verkennen welke handelingsopties de Eerste Kamer heeft om bij te dragen aan verantwoorde ontwikkeling en implementatie van AI-besluitvorming door de overheid.

Het Rathenau Instituut heeft als taak bij te dragen aan het maatschappelijke debat en de politieke oordeelsvorming over de impact van wetenschappelijke en technologische ontwikkelingen op de samenleving. Op verzoek en met inbreng van de Eerste Kamer

heeft het Rathenau Instituut deze notitie gemaakt, bedoeld om de verkenning van de werkgroep AI te ondersteunen. Bij het borgen van grond- en mensenrechten en democratische controle bij de inzet van AI door de overheid komen dossiers van verschillende commissies bij elkaar. De notitie dient ter voorbereiding van een aantal verdiepingssessies en een Kamerbrede deskundigenbijeenkomst over deze complexe en commissie-overstijgende opgave.

De notitie bestaat uit:

1. Een beknopte samenvatting van een aantal zorgen in de samenleving, de wetenschap en de overheid over algoritmische overheidsbesluitvorming.
2. Een overzicht van de manier waarop het kabinet deze zorgen geadresseerd heeft in de periode 2018-2020.
3. Een opsomming van uitdagingen voor de Eerste Kamer met daarbij vier discussiepunten voor de deskundigenbijeenkomst.

1. Zorgen over algoritmische besluitvorming door de overheid

Het gebruik van ‘slimme’ systemen voor het onderbouwen en nemen van overheidsbesluiten past binnen een bredere digitaliseringstrend die de afgelopen decennia heeft plaatsgevonden. Zo gebruikt de overheid al lange tijd digitale systemen om toeslagen toe te kennen, vergunningen te verlenen of automatisch verkeersboetes te innen. Veelal zijn dat systemen die gebaseerd zijn op zogenaamde ‘als, dan’-instructies, ook wel *op regels gebaseerde AI* genoemd.¹ Zo’n systeem kan uit eenvoudige regels bestaan (bijvoorbeeld een overzichtelijke beslisboom), maar kan door vele variabelen en regels ook zeer complex zijn.

Sinds circa vijf jaar zet de overheid ook zogenaamde ‘slimme’ systemen in, die gebaseerd zijn op *machine learning*. Dit zijn systemen die verbanden ontdekken in grote hoeveelheden data, en op basis daarvan bijvoorbeeld de kans berekenen dat iemand heeft gefraudeerd (Systeem Risico Indicatie, SyRI), of dat in een bepaalde wijk bepaalde vormen van criminaliteit voorkomen (*predictive policing*). Het gaat hier om lerende systemen, waarbij ICT-professionals het systeem leren om patronen te herkennen met behulp van trainingsdata.²

Beide typen systemen kunnen worden ingezet bij overheidsbesluitvorming, zowel ter ondersteuning als voor automatische besluitvorming. Het is niet altijd eenvoudig te

1 Een op regels gebaseerd systeem leert niet van zichzelf, maar vertoont wel intelligent gedrag door de omgeving te analyseren en acties te ondernemen – met enige mate van autonomie – om specifieke doelen te bereiken (zie ook de definitie van AI van de Europese Commissie, EC 2019). Veel aandacht rondom AI gaat nu uit naar zelflerende systemen (zie voetnoot 2), waarbij sommigen op regels gebaseerde systemen niet meer als échte AI beschouwen (Rathenau Instituut, 2019).

2 Lerende algoritmen kunnen op verschillende manieren worden getraind. Bij supervised learning wordt het algoritme getraind aan de hand van gelabelde voorbeelden, bijvoorbeeld een plaatje van een kat heeft het label ‘kat’, een plaatje zonder kat het label ‘geen kat’. Programmeurs leren hoe het algoritme met de aangereikte voorbeelden moet omgaan. Bij unsupervised learning worden de trainingsdata niet gelabeld. Het algoritme wordt dan zo geprogrammeerd dat het na verloop van tijd zelfstandig structuren of patronen in de ingevoerde gegevens kan herkennen (Kulk & van Deursen, 2020).

bepalen of er sprake is van (zelf)lerende systemen bij de overheid, of ‘slechts van intensief datagebruik en -koppeling’. Wanneer sprake is van ‘data-intensieve’ toepassingen is de verwachting dat de volgende stap het gebruik van (zelf)lerende algoritmen zal zijn (Wolswinkel, 2020a; TNO, 2019). In de praktijk zien we ook al mengvormen ontstaan; beslisbomen kunnen worden gecombineerd met algoritmen die worden *getraind*.³

Kader 1 Hoe slim is AI?

De meeste lerende AI-systemen zijn momenteel goed in één specifieke taak. Op sommige terreinen, zoals beeldherkenning, boekt de software snel vooruitgang. Op andere terreinen, zoals het doen van voorspellingen over complexe sociale uitkomsten met betrekking tot fraude of baangeschiktheid, is de vooruitgang minder eenduidig.

Dit hangt samen met de probleemruimte (Schermer, e.a., 2020). Bij een schaakcomputer, bijvoorbeeld, is de omgeving waarbinnen een AI-systeem opereert helder. Er is een speelbord en er zijn eenduidige spelregels waarbinnen specifieke doelstellingen moeten worden bereikt. Ook bij gezichtsherkenningssoftware is de probleemruimte nog redelijk goed afgebakend. Doordat systemen gevoed kunnen worden met heel veel foto's en video's van gezichten, slagen ze er steeds beter in om zeer accuraat gezichten te identificeren.

Maar bij meer complexe en dynamische omgevingen spelen er veel meer variabelen en afhankelijkheden. Er zijn tal van factoren waar bijvoorbeeld een zelfrijdende auto rekening mee moet houden. Ook is het lastig om eenduidig prioriteiten of wenselijke strategieën vast te leggen; als er problemen ontstaan, beschermt de auto dan in eerste instantie de inzittende of medeweggebruikers?

Bij het profileren van menselijk gedrag en het voorspellen van sociale uitkomsten is de probleemruimte *nóg* groter. Er worden AI-systemen ontwikkeld om te selecteren wie de meest geschikte sollicitant is of wanneer iemand crimineel gedrag gaat vertonen. Maar welke gegevens daarbij relevant zijn én welke dimensies te vangen zijn in data, is niet duidelijk. Wanneer de probleemruimte niet meer “uit te schrijven” is, groeit de kans op onvoorziene effecten en fouten.

In deze notitie richten we ons op rechtstatelijke kwesties die ontstaan doordat systemen (a) soms op een weinig transparante, of zelfs onnavolgbare, wijze opereren, (b) vaak weinig ruimte overlaten voor maatwerk en (c) mensen- en grondrechten, zoals privacy en non-discriminatie, kunnen schenden. We gaan kort op elk van deze punten in.

1.1 Algoritmische systemen zijn niet altijd transparant

Voor op regels gebaseerde AI-systemen geldt dat de algoritmen en de gemaakte keuzen door mensen zijn geprogrammeerd, en daarmee in principe voor mensen

3 Zie bijvoorbeeld https://www.dsp-groep.nl/wp-content/uploads/11onbetaal_Onderzoek_gebruik_betaalprofielen_DSP-Decisio_2019.pdf

bekend en te begrijpen zijn. In het geval van systeemontwikkeling door derde partijen – wanneer de software door intellectueel eigendom is beschermd – kan het in de praktijk echter lastig zijn voor de overheid en het parlement om transparantie over gemaakte keuzes te garanderen. Maar ook in andere gevallen kan het lastig zijn om te achterhalen hoe beslissingen van op regels gebaseerde systemen tot stand komen.

Kader 2 Waardeoordelen van ICT-professionals

Tijdens het systeemontwerp maken ontwikkelaars en data-analisten voortdurend keuzes en waardeoordelen – het is onmogelijk om dat niet te doen (Wieringa 2020; RvS 2018). De vertaling van wet- en regelgeving naar algoritmes vraagt vaak om het prioriteren en kwalitatief uitdrukken van waardeoordelen. Als tijdens het wetgevingsproces ook waardeoordelen en politieke keuzes plaatsvinden, worden deze keuzes getoetst en gewogen door het parlement – er vindt een democratisch proces plaats. Dit is niet het geval als IT-professionals de wet moeten interpreteren, en daarbij waardeoordelen moeten maken. Het parlement heeft daar geen zicht op of controle over.

Soms zijn wetteksten zo precies uitgewerkt dat programmeurs de beschrijvingen en uitgangspunten vrijwel naadloos als code kunnen overnemen. In de Wet beslagvrije voet (artikel 475da) bijvoorbeeld, staan voor elke categorie – een alleenstaande, een alleenstaande ouder, gehuwden zonder of met kinderen – precieze bedragen en rekenregels beschreven (Passchier, 2021).

Soms laat wet- en regelgeving echter nog allerlei beoordelingsruimte op het uitvoerende niveau en kan eenzelfde situatie verschillend worden gewogen (Van Eck e.a., 2018). Om te beslissen over het intrekken van een vergunning moet bijvoorbeeld worden beoordeeld welke gebreken ernstig genoeg zijn. Wanneer criteria multi-interpretabel zijn of iemand in verschillende gradaties aan een criterium kan voldoen, dan moeten bij programmeren van een AI-systeem waardeoordelen worden geveld. ICT-professionals die formeel geen besluitvormende bevoegdheid hebben, worden dan *de facto* ook beleidsmakers.

Er zijn verschillende mogelijkheden om de transparantie en democratische controle te vergroten. Enerzijds kan het betrekken van ICT-professionals bij het opstellen van wetteksten helpen voorkomen dat er in een later stadium ruis ontstaat in de vertaling van wetten naar AI-systemen.⁴ Anderzijds kunnen verplichte impact assessments of audits, bijvoorbeeld door organisaties zelf, helpen afdwingen, dat ICT-professionals reflecteren op te maken keuzes en dat deze dusdanig worden vastgelegd dat ze achteraf inzichtelijk en controleerbaar zijn voor burgers, het parlement, de toezichthouder en/of de rechter.

4 Binnen de wetenschap en de overheid wordt gezocht naar manieren om tot “machinaal uitvoerbare wetgeving” te komen. Een voorbeeld is het project Calculemus-FLINT waarbij gekeken wordt hoe bij het opstellen van wetgeving al rekening kan worden gehouden met ontwerpkeuzes. Het parlement kan deze beter uitgewerkte wetgeving toetsen, zodat de democratische legitimiteit van AI-systemen wordt vergroot.

De keuzes van *lerende AI-systemen* zijn op een fundamenteel niveau moeilijker navolgbaar en te verantwoorden dan de keuzes van op regels gebaseerde AI-systemen. Er zijn verschillende *machine learning*-technieken waarbij in feite niet de *ICT-professionals*, maar de *systemen zelf*, bepalen wat relevante criteria of patronen zijn. Dit proces leunt zwaar op kansberekening (statistiek) en gaat gepaard met een enorme 'datahonger'.⁵ Dat wil overigens niet zeggen dat er geen mensen meer bij de besluitvorming door systemen betrokken zijn. Softwareontwikkelaars leren het systeem patronen te herkennen met behulp van trainingsdata, zodat het systeem uiteindelijk in staat is om bijvoorbeeld voor de mens onverwerkbare datasets te analyseren. Het AI-systeem is afhankelijk van de data die zij als input krijgt. Als die data fouten bevatten, dan zal het AI-systeem ook geen zinnige resultaten genereren, ofwel: *garbage in, garbage out*. Daarnaast kunnen er problemen ontstaan wanneer gevonden verbanden (correlaties) als causale verbanden worden geïnterpreteerd. Bij lerende AI-systemen kunnen ontwerpers wel 'onder de motorkap' kijken, en de programmeercode aanpassen. Maar dat geeft vaak maar beperkt inzicht in de manier waarop het systeem allerlei criteria en data hanteert.⁶

Er is dus een risico dat keuzes, die tijdens de softwareontwikkeling worden gemaakt door ICT-professionals en/of AI-systemen, zich onttrekken aan politiek debat en democratische controle. Als onduidelijk blijft hoe (en door wie) waardeoordelen structureel worden omgezet in getallen, bepaald wordt wanneer een systeem 'goed genoeg' is en wat acceptabele foutmarges zijn (Wieringa, 2020), dan heeft dit fundamentele gevolgen voor de democratische legitimiteit van AI-besluiten.

1.2 AI-systemen laten weinig ruimte voor maatwerk

Terwijl ICT-professionals meer invloed krijgen om beleid (naar eigen inzicht) te vertalen en uit te voeren, wordt de discretionaire ruimte van ambtenaren steeds kleiner (zie Bovens & Zouridis 2002; Zouridis 2000). Twintig jaar geleden beschreven bestuurskundigen al de transformatie die zich binnen de overheid voltrok door de toepassing van ICT. In de plaats van de uitvoeringsambtenaar die oordeelt over een individueel geval, komt steeds vaker een algoritmisch systeem te staan, dat tot een besluit komt (Bovens & Zouridis 2002). Er ontstaat meer standaardisatie in de besluiten en steeds meer scenario's en oplossingen worden al op systeemniveau voorzien. Ook laten diverse wetten weinig discretionaire ruimte toe, denk aan de wetgeving die een rol speelde in de toeslagenaffaire en geen ruimte liet voor maatwerk. De wendbaarheid kan verloren gaan om voor individuele casussen passende oplossingen te formuleren.

5 De beschikbaarheid van grote hoeveelheden data is cruciaal om AI-systemen te trainen en beter te maken (zie voetnoot 2).

6 De wetenschap werkt aan methodes om dit te verbeteren, zie het Hybrid Intelligence project in het NWO-zwaartekrachtprogramma.

Kader 3 Schrijnende gevallen

Een foutje maakt nog geen fraudeur. Toch kwamen tienduizenden ouders door de toeslagenaffaire in grote financiële problemen. Het rapport 'Ongekend Onrecht' roept op tot kritische reflectie over het functioneren van de hele overheid, waaronder de inzet van algoritmische overheidsbesluitvorming. Voor een specifiek AI-systeem moet gelden dat er snel en effectief kan worden ingegrepen als het misgaat, omdat bij sommige AI-toepassingen kleine fouten in een database grote consequenties kunnen hebben voor burgers.

Een fout in het automatic numberplate recognition (ANPR) bestand kan er bijvoorbeeld toe leiden dat iemand meermaals door de politie wordt aangehouden. Die fout kan zijn ontstaan in de aanlevering van informatie door het Centraal Bureau Rijvaardigheid (CBR), de Dienst Wegverkeer (RDW) of de gemeente. De Nationale ombudsman oordeelde dat de politie desalniettemin verantwoordelijk is voor een betrouwbaar ANPR-bestand en ervoor moet zorgen dat fouten hersteld worden (Nationale Ombudsman, 2015).

Helaas komt het regelmatig voor dat het niet of bijna niet mogelijk is om een fout snel te traceren en die fout met één druk op de knop in alle onderling verbonden systemen te herstellen (Widlak en Peters 2018; Van Eck e.a., 2018).⁷ De Nationale Ombudsman en Stichting Kafbrigade hebben de afgelopen jaren schrijnende gevallen gedocumenteerd van burgers die klem komen te zitten in de digitale overheidsbureaucratie.

Wanneer systemen besluiten nemen op basis van verkeerde, verouderde of onvolledige gegevens, ligt de verantwoordelijkheid al snel bij de burger om fouten waar te nemen en een bezwaarprocedure te starten (WRR, 2017). Ook ontbreekt het vaak aan een menselijk aanspreekpunt en worden burgers zo van het kastje naar de muur gestuurd (TCU, 2021).

Het is van belang ervoor te zorgen dat het mogelijk blijft om in te grijpen wanneer systemen vastlopen of wanneer er onvoorziene effecten optreden die ongewenste uitwerking hebben op burgers. Bij AI-systemen kan in het ontwerp worden bepaald welke mate van menselijke betrokkenheid en flexibiliteit wenselijk is.⁸ Ook valt te denken aan verplichte transparantie over de herkomst van gebruikte gegevens, zodat die ook kunnen worden gecorrigeerd of aangevochten.

Het gaat echter niet alleen om de technologie. Het is cruciaal een optie in te bouwen om een specifieke situatie nader toe te lichten aan een *mens*, die vervolgens ook een

7 Vaak is het wel een optie om een aantekening van onjuistheid te plaatsen, maar dan kunnen er alsnog problemen ontstaan door ketenautomatisering bij de overheid. Bij IT-systemen is het daardoor soms vrijwel onmogelijk om fouten *met terugwerkende kracht* recht te zetten, terwijl dit in het recht een belangrijke principe is.

8 Er wordt vaak een onderscheid gemaakt tussen human-in-the-loop (mens geeft toestemming; AI-systeem als hulpmiddel), human-on-the-loop (mens kan ingrijpen; AI-systeem is te stoppen) en human-out-of-the-loop (geen menselijk toezicht; automatische besluitvorming). Maar ook op het niveau van digitale formulieren kan er bijvoorbeeld veel gewonnen worden door wensen van burgers mee te nemen en te werken aan gebruiksvriendelijkheid (Nationale Ombudsman, 2019).

oordeel kan vellen en kan afwijken van het AI-besluit. Wanneer er niet voldoende bewegingsruimte is voor mensen om het systeem te corrigeren of specifieke uitkomsten zelf te (her)beoordelen, komt de menselijke maat en betekenisvolle menselijke tussenkomst in het gedrang.

1.3 AI-systemen kunnen grondrechten schenden

De afgelopen jaren is uit diverse studies gebleken dat algoritmische overheidsbesluitvorming risico's oplevert voor de bescherming van grond- en mensenrechten (Rathenau Instituut 2017a; Vetzo et al. 2018; WRR 2016). Problemen kunnen ontstaan bij de implementatie van systemen. Zo staat de menswaardige behandeling en het recht op menselijk contact onder druk, wanneer burgers worden geconfronteerd met '*computer says no*'-scenario's (zie ook kader 3). Maar ook bij de keuze om een systeem te gaan bouwen en tijdens de ontwikkelingsfase van het systeem kan het misgaan, waardoor bijvoorbeeld ongewenste discriminatie optreedt (zie kader 4). De afgelopen jaren zijn in diverse gevallen het recht op respect voor de persoonlijke levenssfeer, het recht op gelijke behandeling en het recht op een eerlijk proces vreselijk in het geding gekomen.

Kader 4 Vormen van discriminatie

Er zijn allerlei manieren waarop vooroordelen in een AI-systeem kunnen sluipen en het AI-systeem als gevolg daarvan systematisch discriminerende besluiten neemt (Barocas en Selbst, 2016; Zuiderveen-Borgesius 2019). Er kunnen problemen optreden door de moeilijkheid om de opdracht van het systeem en relevante kenmerken of variabelen te definiëren (zie ook kader 1). Maar het kan ook misgaan op het niveau van de trainingsdata. Die kunnen bijvoorbeeld niet-representatief zijn of al menselijke vooringenomenheid bevatten.

In de ICT-industrie zijn bijvoorbeeld van oudsher meer mannen dan vrouwen werkzaam. Zo'n historisch gegeven kan ertoe leiden dat een AI-systeem voor het selecteren van sollicitanten onbedoeld een 'oude' vooringenomenheid reproduceert.⁹ Een Ivy League-diploma als selectiecriteria kiezen, kan ook problematisch zijn. Deze universiteiten zijn immers in hun selectieproces ook niet (altijd) inclusief (geweest). Volgens wetenschappers zijn diploma's bovendien geen beste indicator voor toekomstig succes.¹⁰ Maar zelfs een criterium als 'nooit te laat komen op je werk' kan discriminerend uitpakken, aangezien bepaalde groepen meer risico lopen op vertragingen door hun afhankelijkheid van het openbaar vervoer of betaalbare woonruimte buiten de stad.

Ten slotte kunnen bepaalde kenmerken ook gecodeerd zijn in andere gegevens. Als informatie over geslacht of huidskleur niet is opgenomen in de trainingsdata, kunnen dergelijke kenmerken alsnog meewegen in het besluit van een AI-systeem, bijvoorbeeld doordat correlaties bekend zijn, bijvoorbeeld tussen postcodes en

9 Dit overkwam Amazon bij het ontwikkelen van een systeem voor het automatisch selecteren van sollicitanten. Zelfs op het moment dat de ontwikkelaars zich bewust werden van dit specifieke vooroordeel in het systeem, kregen ze het er niet uit en zag het bedrijf zich genoodzaakt te stoppen met de ontwikkeling van het systeem.

10 Zie ook: Rathenau Instituut (2020a).

etniciteit. Er kan bovendien ook een bevooroordeelde dataset ontstaan door de manier waarop trainingsdata worden verzameld. Wanneer de politie meer aandacht richt op bepaalde etnische groepen of bepaalde buurten, levert dat geen eerlijke steekproef op. Het is dan niet verwonderlijk dat die groepen of buurten systematisch oververtegenwoordigd zijn in politieregisters.

Diverse AI-systemen zijn vanwege hun mogelijke inbreuk op grond- en mensenrechten onderwerp van maatschappelijke en politieke discussie. Sommige systemen zijn door de rechter verboden, zoals het SyRI-systeem dat algoritmes gebruikt om burgers te profileren om fraude op te sporen. Volgens de rechter waren het maatschappelijk belang en de inbreuk op de burgerrechten niet met elkaar in balans. Het in gebruik nemen van AI-systemen vraagt daarom om zorgvuldige toetsing op grondrechten.

2. De gereedschapskist van het kabinet

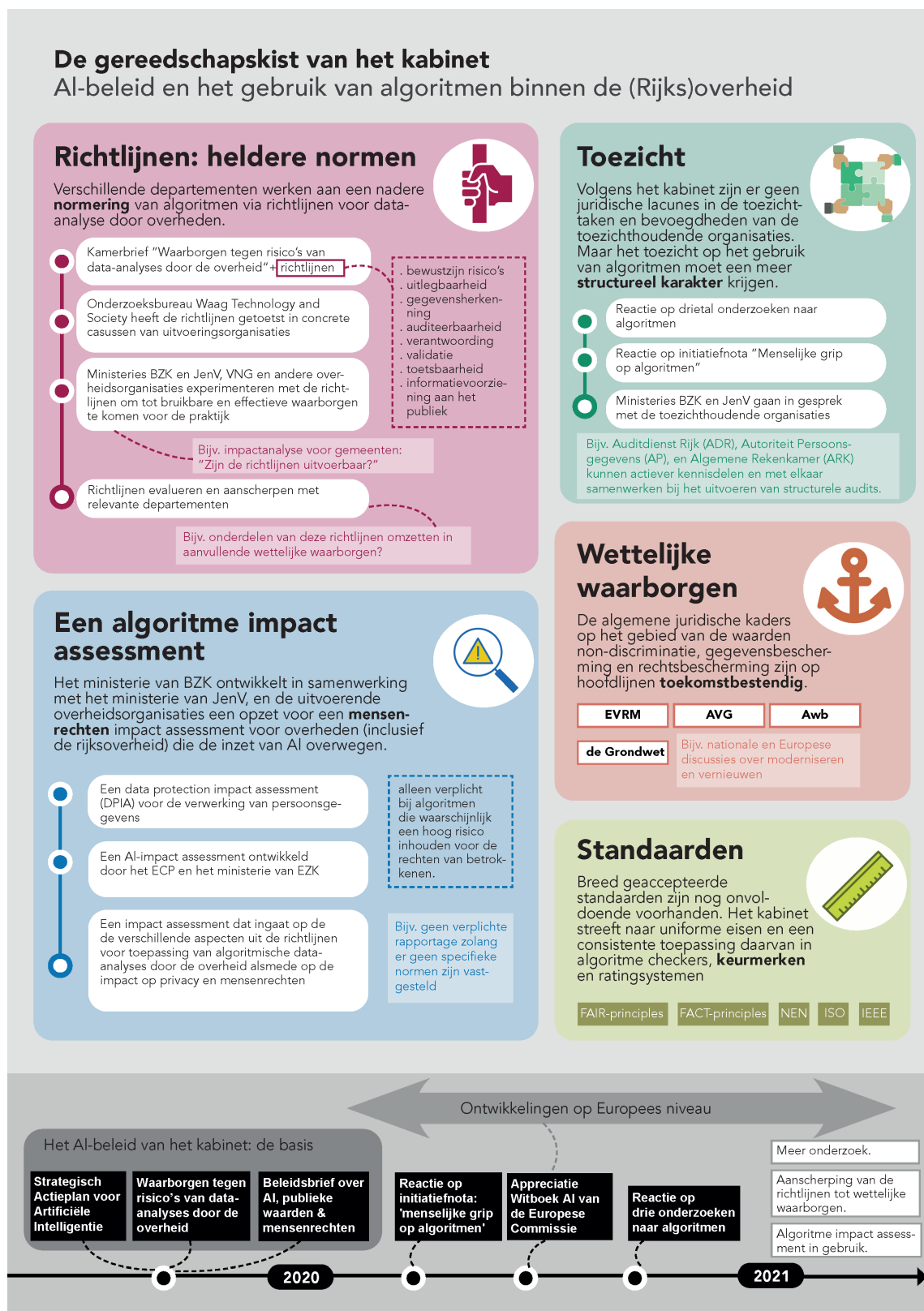
Gezien de hierboven genoemde problemen, beschouwen velen een betere democratische verantwoording van de ontwikkeling en het gebruik van algoritmische besluitvorming als noodzakelijk (CAHAI 2020; EC 2019). Ook het kabinet heeft de hierboven beschreven zorgen grotendeels erkent. Het kabinet heeft de afgelopen drie jaar in diverse onderzoeken en Kamerbrieven bekeken of de bestaande wettelijke kaders en het bijbehorende toezichtstelsel – de ‘gereedschapskist’ van het kabinet – voldoende toekomstbestendig is en welke middelen nodig zijn om de zorgen over verantwoording, menselijke maat en mensenrechten te verkleinen. We geven de belangrijkste acties van het kabinet weer in figuur 1.

Belangrijke kaders zijn de Grondwet, de Algemene wet bestuursrecht (Awb) en de Algemene Verordening Gegevensbescherming (AVG). Het globale beeld dat uit de brieven naar voren komt, is dat de juridische kaders en het toezicht grotendeels op orde zijn, maar dat er nog wel lacunes te dichten zijn (Kamerstukken II 2019-2020, 26643, nr. 726). Daarnaast acht het kabinet *structureel toezicht* nodig op algoritmische overheidsbesluitvorming door de Algemene Rekenkamer, de Autoriteit Persoonsgegevens en de Auditdienst Rijk (Kamerstukken II 2019-2020, 35212, nr. 3).¹¹

Het kabinet ziet met name aandachtspunten bij de vraag of de huidige informatie- en transparantieverplichtingen ten aanzien van geautomatiseerde besluitvorming en profilering voldoende zijn (Kamerstukken 2019-2020, 26643, nr. 641). Zo vereist artikel 22 van de AVG dat *geheel* geautomatiseerde besluitvorming aan verschillende waarborgen moet voldoen. Maar de AVG maakt niet duidelijk welke waarborgen precies gelden voor *gedeeltelijk* geautomatiseerde besluitvorming, terwijl veel algoritmische besluitvorming door de overheid binnen deze categorie vallen.

¹¹ Het kabinet heeft een ambtelijke werkgroep aangesteld die een onderzoeksagenda opstelt op het gebied van normering van en toezicht op algoritmen.

Figuur 1: Overzicht kabinetsbeleid algoritmische besluitvorming bij de overheid



Een ander probleem doet zich voor bij zogenaamde ‘gebiedsgebonden’ data-analyse. Dergelijke risicoanalyses zijn niet gericht op individuen, maar op wijken of straten. Ze vallen daarmee niet onder de AVG-definitie van profilering.¹² Het kabinet geeft aan dat sommige risico’s van profilering, bijvoorbeeld de kans op vooroordelen, ook bij deze data-analyses kunnen voorkomen.

Het kabinet kijkt inmiddels met name naar de Europese Commissie voor nadere regulering van AI-systemen (Kamerstukken II 2019-2020, 26643, nr. 726).¹³ Naar verwachting zullen aan ‘hoog-risico’ AI-toepassingen nadere wettelijke eisen worden gesteld (EC, 2020; Kamerstukken II 2019-2020, 26643 / 32761, nr. 697).

Ondertussen werkt het kabinet aan op de ontwikkeling van richtlijnen, ontwerpprincipes en *impact assessments* om systeemontwerpers en uitvoeringsinstanties te helpen om vooraf mogelijke mensenrechtenschendingen in te schatten en te voorkomen, en de gemaakte keuzes tijdens het systeemontwerp beter te kunnen verantwoorden (Kamerstukken 2019-2020, 26643, nr. 641; Kamerstukken II 2019-2020, 26643, nr. 726). Het kabinet kiest voor een ‘lerende aanpak’ en financiert bijvoorbeeld het project ‘Calculus-FLINT’, waarin een systematiek wordt uitgewerkt om op transparante wijze wet naar algoritme-regels te vertalen.

Ten aanzien van de navolging van de Algemene wet bestuursrecht (Awb) zijn er ook aandachtspunten. De Awb regelt onder meer dat burgers met hun klachten terecht kunnen bij de Ombudsman en dat de Algemene Rekenkamer en de Auditdienst Rijk de overheid controleert. De beginselen van behoorlijk bestuur, vastgelegd in de Awb, zijn ook van toepassing op algoritmische besluitvorming. Het motiverings- en zorgvuldigheidsbeginsel bepalen dat de overheid zorg moet dragen voor het inzicht in en de controleerbaarheid van door algoritmes gemaakte keuzes.¹⁴

3. Discussiepunten voor de Eerste Kamer

Naast de inspanningen van het kabinet zijn er ook manieren voor de Eerste Kamer om háár rol ten aanzien van algoritmische overheidsbesluitvorming te versterken. Ter ondersteuning van een deskundigenbijeenkomst hierover en op basis van verschenen onderzoeken, adviezen en ministeriele brieven, stelt het Rathenau Instituut hieronder vier discussiepunten voor.

12 De politie gebruikt gebiedsgebonden-analyse bijvoorbeeld om informatie te verzamelen over waar bepaalde vormen van criminaliteit in een stad voorkomen, en op basis hiervan de benodigde politie-inzet te bepalen.

13 Meer specifiek naar de risico-benadering van AI door de Europese Commissie, waar verwacht wordt dat aan ‘hoog-risico’ AI-toepassingen nadere wettelijke eisen zullen worden gesteld (EC, 2020), en de kabinetsappreciatie Witboek kunstmatige intelligentie (Kamerstukken II 2019-2020, 26643 / 32761, nr. 697).

14 Deze transparantie is niet absoluut. Dat wil zeggen, inzicht en controleerbaarheid in de door het systeem gemaakte keuzes, betekenen niet dat openbaarmaking van de volledige broncode nodig is (zie Kamerstukken 2018 – 2019, 26 643, nr. 570). Voor een nadere bespreking van de toekomstbestendigheid van het bestuursrecht in het kader van zelflerende systemen, zie bijvoorbeeld Wolswinkel (2020a; 2020b).

De kerntaak van de Eerste Kamer bestaat uit het toetsen van de kwaliteit van wetgeving in formele zin. Zij zal dus alert moeten zijn op verschillende wetsvoorstellen en -wijzigingen die langskomen en deze kritisch moeten toetsen op essentiële criteria, zoals rechtmatigheid, uitvoerbaarheid en handhaafbaarheid. Maar lang niet ieder AI-systeem wordt geïntroduceerd door middel van een wet of wetswijziging, bijvoorbeeld omdat het AI-systeem gereguleerd wordt door bestaande juridische kaders. *Het eerste discussiepunt gaat daarom over de verschillende manieren om algoritmische overheidsbesluitvorming steviger te verankeren in de wet.* Sommige wetsvoorstellen en -wijzigingen kunnen bijdragen aan deugdelijke ontwikkeling en gebruik van alle AI-systemen. Alertheid op wijzigingen van dergelijke overkoepelende wet- en regelgeving is dus cruciaal.

Het gebeurt ook dat een nieuwe wet of wetswijziging direct ruimte creëert om bepaalde digitale systemen te ontwikkelen en toe te passen. Dit was bijvoorbeeld het geval bij SyRI, dat verbonden is aan de wet Structuur Uitvoeringsorganisatie Werk en Inkomen (SUWI). Op het moment dat de Eerste Kamer aan zet is om een dergelijk wetsvoorstel te beoordelen, spelen drie aspecten een belangrijke rol voor een zorgvuldige toetsing. Deze aspecten vormen *de volgende drie discussiepunten*:

- de proportionaliteit van de voorgestelde wetgeving,
- de manier waarop menselijke tussenkomst en toezicht geregeld is, en
- de mate waarin het parlement na goedkeuring van de wet op de ontwikkeling van het AI-systeem kan toezien.

3.1 Hoe kunnen AI-systemen steviger in de wet verankerd worden?

Hierboven noemden we al dat niet ieder AI-systeem wordt geïntroduceerd door middel van een wet of wetswijziging. Zo valt het Criminaliteits Anticipatie Systeem (CAS) onder de bestaande wet Politiegegevens. CAS is dus nooit in de Eerste Kamer besproken, terwijl er inmiddels controverse bestaat over de inbreuk die het systeem maakt op de privacy van burgers, over vermeende discriminerende effecten en over de effectiviteit van de voorspellingen van CAS. De vraag rijst of systemen met een dergelijke impact op grond- en mensenrechten meer parlementaire controle vereisen.

Een manier om die parlementaire controle te vergroten, is via nadere algemene juridische inkadering. Bestaande wetten zien namelijk al vaak toe op het gebruik van AI-systemen door de overheid, of stellen eisen aan het handelen van de overheid. De Uitvoeringswet AVG en de Wet Politiegegevens zien bijvoorbeeld toe op dataverwerking door verschillende overheidsdiensten, terwijl de Algemene Bestuurswet eisen stelt aan de besluiten die de overheid neemt – ook als deze ondersteund worden door algoritmes. Deze wetten regelen niet de komst van een specifiek systeem, maar geven wel aan wat de overheid of uitvoeringsinstanties mogen of kunnen doen. Zijn in deze wetten aanscherpingen of wijzigingen nodig?

- ❖ ***Tijdens de deskundigenbijeenkomst** kunnen de Eerste Kamerleden verschillende manieren bespreken om AI-systemen steviger in de wet te verankeren. Hierbij kan onder meer gedacht worden aan eventuele updates of aanscherpingen van algemene wetten die de Eerste Kamer heeft te toetsen, zoals de Grondwet, Awb, WoB, AVG en de wet Politiegegevens. Wat zijn bijvoorbeeld de voor- en nadelen van het advies van wetenschappers en de Raad van State om softwarecode een hogere wettelijke status¹⁵ te geven?*

3.2 Hoe kan de Eerste Kamer beter ondersteund worden bij de proportionaliteitstoets van AI-gerelateerde wetgeving?

Het is de taak van de Eerste Kamer om in het kader van de bescherming van grond- en mensenrechten de beginselen van proportionaliteit en subsidiariteit toe te passen. Daarbij is het zaak af te wegen of een eventuele inbreuk op deze rechten door aan AI gerelateerde wetgeving gerechtvaardigd is, gezien de verwachte maatschappelijke opbrengst van het voorgestelde systeem of type systemen. Het is ook van belang om daarbij na te gaan of er alternatieven voorhanden zijn die hetzelfde resultaat op zouden kunnen leveren, maar minder of geen inbreuk zouden maken op grond- en mensenrechten.

Om dit zorgvuldig te kunnen toetsen, is informatie nodig over de effectiviteit en de ontwikkeling van een voorgesteld AI-systeem. Dit kan vooraf lastig zijn, zie discussiepunt 4 (paragraaf 3.4). Ook is het van belang om na te gaan hoe het systeem precies in de wet verankerd wordt. Dit zal deels bestaan uit het kritisch bestuderen van de adviezen van de Raad van State en de betrokken toezichthouders, en van de reacties van het kabinet hierop. De Eerste Kamer organiseert verder regelmatig expertmeetings om zich voor te bereiden op bepaalde wetgevingstrajecten. Ook inzichten uit mensenrechten *impact assessments*, bijvoorbeeld gemaakt door een uitvoeringsorganisatie, kunnen Eerste Kamerleden helpen bij de beoordeling.

De besluitvorming van de Eerste Kamer zelf is verdeeld in vaste commissies, maar één AI-systeem kan het werkveld van meerdere commissies beslaan. Een goede toetsing op proportionaliteit vraagt dus om meer integrale behandeling van een wetsvoorstel, waarbij meerdere commissies zijn betrokken. De Eerste Kamer kan verkennen of, en op welke manier, zij haar grip op dit besluitvormingsproces kan versterken. Zo heeft de Tweede Kamer naar aanleiding van de tijdelijke commissie Digitale Toekomst besloten om een eigen vaste commissie voor Digitale Zaken in te stellen, die de vaste commissies kan ondersteunen bij besluiten over AI (Tweede Kamer, 2020).¹⁶

¹⁵ De beslisregels in software zouden bijvoorbeeld de status van beleidsregels (in de zin van de Awb) of van een Algemene Maatregel van Bestuur (AMvB) kunnen krijgen (Van Eck et al. 2018; RvS, 2018).

¹⁶ Deze commissie kan een bewindspersoon verantwoordelijk voor digitalisering controleren, toekomstige commissieoverstijgende ontwikkelingen verkennen en agenderen, andere vaste commissies informeren, en het voortouw nemen bij de behandeling van commissieoverstijgende digitaliseringsvraagstukken, en fungeren als een aanspreekpunt voor digitaliseringskwesties voor zowel het kabinet als maatschappelijke groepen, bedrijfsleven, wetenschap en anderen (Tweede Kamer 2020).

- ❖ *Tijdens de deskundigenbijeenkomst kunnen de leden bespreken hoe de Eerste Kamer haar kennis en expertise verder uit kan bouwen. Welke informatie is nodig: zouden rapportages als de mensenrechten impact assessment verplicht naar het parlement moeten worden gestuurd? Is een integrale behandeling van wetsvoorstellen nodig? Is een eigen vaste commissie voor Digitale Zaken bij de Eerste Kamer wenselijk? Is samenwerking met de Tweede Kamer een optie?*¹⁷

3.3 Hoe kan de Eerste Kamer effectief toezicht en ruimte voor maatwerk beter borgen?

Behalve op rechtmatigheid, toetst de Eerste Kamer ook op uitvoerbaarheid en handhaafbaarheid. De mate van beoordelingsruimte of zelfs beleidsvrijheid op het uitvoerende niveau is hierbij een belangrijk aspect. Het streven is om potentiële ruis in de vertaling van de wet naar de praktijk zoveel mogelijk te voorkomen. Het is immers onwenselijk om besluitvormende bevoegdheid, zonder democratische verantwoording, aan ICT-professionals of zelfs AI-systemen te geven. Wel is het cruciaal om ruimte te laten voor maatwerk. Ambtenaren moeten een passende oplossing kunnen formuleren voor een uitzonderlijke casus. Om te voorkomen dat burgers worden geconfronteerd met ‘*computer says no*’-scenario’s, moet niet alles op systeemniveau worden dichtgetimmerd. Het is dus zaak om te controleren of nieuwe systemen binnen ambtelijke organisaties voldoende ruimte bieden voor menselijke tussenkomst en maatwerk.

In dit kader is het belangrijk dat de Eerste Kamer kijkt naar het toezicht. De verschillende toezichthouders, zoals de Autoriteit Persoonsgegevens en de Algemene Rekenkamer, hebben vanuit hun functie en bevoegdheden mogelijkheden om na te gaan in hoeverre de algoritmische besluitvorming bij de overheid verantwoord plaatsvindt. Toch is de praktijk weerbarstig; de vraag rijst in hoeverre er soms sprake is van een papieren werkelijkheid. Te vaak blijkt het toezicht niet toereikend om slecht functionerende geautomatiseerde besluitvorming tijdig te detecteren en waar nodig aan de bel te trekken. Het is daarom zaak dat de Eerste Kamer in staat is om zorgvuldig te beoordelen of een wetsvoorstel genoeg ruimte biedt voor menselijke tussenkomst en controleert of effectief toezicht mogelijk is. Ook hiervoor geldt dat de Kamer interne en externe expertise moet opbouwen en benutten.

- ❖ *Tijdens de deskundigenbijeenkomst kunnen de leden bespreken hoe de Eerste Kamer meer grip kan krijgen op de controle van AI-systemen die worden gebruikt bij overheidsbesluitvorming: is vooraf duidelijk welke verantwoordelijkheden diverse partijen krijgen? Op welke manier kan de bewegingsruimte van ambtenaren ten aanzien van geautomatiseerde*

¹⁷ In het buitenland zijn er diverse voorbeelden van samenwerking van parlementaire kamers te vinden. In het Verenigd Koninkrijk bestaan er gezamenlijke commissies tussen beide Kamers, en hebben onderzoekscommissies van het Hogerhuis (inquiry committees) invloed op de politieke discussies in het Lagerhuis. Om meer grip te krijgen op digitaliseringsvraagstukken heeft de Tweede Kamer een inventarisatie laten maken van de werkwijze van buitenlandse parlementen (Rathenau Instituut 2020b).

besluitvorming wettelijk worden vastgelegd? En zijn de toezichthouders geëquipeerd om hun taak effectief uit te voeren?

3.4 Welke mate van parlementaire betrokkenheid is wenselijk na invoering van wetsvoorstellen?

De toetsing van een AI-gerelateerd wetsvoorstel gebeurt doorgaans vóór de ontwikkeling en daadwerkelijke introductie van dat AI-systeem. We gaven hierboven al aan dat tijdens het systeemontwerp wetsbepalingen naar algoritmen moeten worden vertaald. Het parlement heeft geen zicht op en controle over interpretaties en waardeoordelen die systeemontwerpers genoodzaakt zijn te maken. De discretionaire ruimte van ICT-professionals neemt toe, maar wordt nog weinig genormeerd. Met andere woorden: rondom de uitwerking van de wet in (toekomstige) AI-systemen is momenteel weinig democratische controle georganiseerd.

Ook is er vooraf doorgaans weinig bekend over de effectiviteit van slimme systemen. Tot nu blijken er bijvoorbeeld weinig aanwijzingen te zijn dat SyRI leidde tot een betere opsporing van fraude, of dat *predictive policing* beter werkt dan traditionele opsporingsmethodes (Meijer en Wessels, 2019). Vaak zal meer informatie over de effectiviteit gedurende de looptijd van het systeem moeten worden verzameld, bijvoorbeeld via een pilot of experiment. Maar als de Eerste Kamer daar niet vooraf om vraagt, komt deze informatie, en de mogelijkheid om opnieuw proportionaliteit te wegen en te toetsen, dus niet meer bij de Eerste Kamer.

Bovendien kunnen de omstandigheden waaronder een systeem van start ging – en de politieke keuzes en afwegingen die aanvankelijk gemaakt werden – in de loop der tijd veranderen. AI-software is dermate complex dat het vrijwel altijd onvoorziene effecten heeft die mogelijk een ongewenste uitwerking hebben op burgers en bedrijven. Ook kunnen uitvoeringsinstanties de grens opzoeken van hun mandaat en kan het gegeven mandaat achteraf te ruim blijken.

Om al deze redenen kan het wenselijk zijn om de weging die de Eerste Kamer aanvankelijk deed opnieuw te bezien.

- ❖ ***Tijdens de deskundigenbijeenkomst*** kunnen de leden verkennen op welke manieren de Eerste Kamer een vinger aan de pols kan houden en relevante wetgeving op enig moment na invoering kan (her)evalueren. Hierbij kan onder meer gedacht worden aan het bespreken van de voor- en nadelen van tijdelijk geldige wetgeving, het opnemen van horizonbepalingen en het invoeren van voorhangconstructies.

Literatuurlijst

Barocas, S. en A. Selbst (2016). "Big Data's Disparate Impact." *Californian Law Review* 104, 671.

Bovens, M. en S. Zouridis (2002). "Van street-level bureaucratie naar systeem-level bureaucratie Over ICT, ambtelijke discretie en de democratische rechtsstaat." *Nederlands Juristenblad* 77(2): 65-74.

Europese Commissie (2020). *White paper on Artificial Intelligence: a European approach to excellence and trust*. https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en

Europese Commissie (2019). *A definition of AI: main capabilities and disciplines*. High-Level Expert Group on Artificial Intelligence. <https://ec.europa.eu/digital-single-market/en/news/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines>

Kamerstukken II 2019-2020. 26643, nr. 726.

Kamerstukken II 2019-2020. 35212, nr. 3.

Kamerstukken II 2019-2020. 26643, nr. 641.

Kamerstukken II 2019-2020. 26643 / 32761, nr. 697.

Kulk, S. en M. van Deursen (2020). *Juridische aspecten van algoritmen die besluiten nemen – een verkennend onderzoek*. Den Haag: WODC 2020.

Meijer, A. en M. Wessels (2019). "Predictive Policing: Review of Benefits and Drawbacks." *International Journal of Public Administration* 42(12): 1031-1039.

Nationale Ombudsman, Jaarverslag 2019: 'Regel regie!'.

Nationale Ombudsman, Jaarverslag 2015: 'Burgerperspectief: een manier van kijken'.

Passchier, R. (2021). *Artificiële intelligentie en de rechtsstaat. Over verschuivende overheidsmacht, Big Tech en de noodzaak van constitutioneel onderzoek*. Den Haag: Boom Juridisch.

Raad van State (2018). *Ongevraagd advies over de effecten van de digitalisering voor de rechtsstatelijke verhoudingen*. <https://www.raadvanstate.nl/@112661/w04-18-0230/>

Rathenau Instituut (2020a). *Werken op waarde geschat. Grenzen aan digitale monitoring op de werkvloer door middel van data, algoritmen en AI*. Den Haag: Rathenau Instituut

Rathenau Instituut (2020b). *Meer grip op digitalisering. Een internationale vergelijking van parlementaire werkvormen*. Den Haag: Rathenau Instituut

Rathenau Instituut (2019). "Zo brengen we AI in de praktijk vanuit publieke waarden." <https://www.rathenau.nl/nl/digitale-samenleving/zo-brengen-we-ai-de-praktijk-vanuit-europese-waarden>

Rathenau Instituut (2017a) *Opwaarderen. Borgen van publieke waarden in de digitale samenleving*. Den Haag: Rathenau Instituut.

Rathenau Instituut (2017b) *Human rights in the robot age. Challenges arising from the use of robotics, artificial intelligence, and virtual and augmented reality*. Den Haag: Rathenau Instituut.

Schermer, B. J. van Ham en K. W. Falkena (2020). *Onvoorziene effecten van zelflerende algoritmen*. Amsterdam: Considerati.

TNO (2019) *Quick scan AI in de publieke dienstverlening* (eindpresentatie 8 april 2019). <https://www.rijksoverheid.nl/documenten/rapporten/2019/04/08/quick-scan-in-de-publieke-dienstverlening>.

Tweede Kamer der Staten-Generaal (2021a). *Klem tussen balie en beleid*. Rapport van de Tijdelijke Commissie Uitvoeringsorganisaties (TCU).

Tweede Kamer der Staten-Generaal (2021b). *Ongekend Onrecht*. Rapport van de Parlementaire ondervragingscommissie Kinderopvangtoeslag (POK).

Tweede Kamer der Staten-Generaal (2020). *Update vereist. Naar meer parlementaire grip op digitalisering*. Rapport van de Tijdelijke Commissie Digitale Toekomst (TCDT).

Van Eck, M., M. Bovens en S. Zouridis (2018). "Algoritmische rechtstoepassing in de democratische rechtsstaat." *Nederlands juristenblad* 93(40), 3008-3017.

Vetzo, M.J., J. Gerards en R. Nehmelman (2018). *Algoritmes en grondrechten*. Den Haag: Boom Juridisch.

Widlak, A. & R. Peeters (2018). *De digitale kooi. Over (on)behoorlijk bestuur door informatiearchitectuur. Of: hoe we de burger weer centraal zetten in een digitaliserende overheid*. Den Haag: Boom Bestuurskunde.

Wieringa, M. (2020). "What to account for when accounting for algorithms. A systematic literature review on algorithmic accountability." In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*: 1-18.

Wolswinkel, J. (2020a). "AR meets AI. AI: Een bestuursrechtelijk perspectief op een nieuwe generatie besluitvorming." *Computerrecht*, 2020(1): 22-29.

Wolswinkel, J. (2020b). *Willekeur of algoritme? Laveren tussen analoog en digitaal bestuursrecht*. Tilburg University.

WRR (2017). *Weten is nog geen doen – Een realistisch perspectief op redzaamheid*. Den Haag: WRR.

WRR (2016). *Big data in een vrije en veilige samenleving*. Den Haag: WRR.

Zouridis, S. (2000). *Digitale disciplineren. Over ICT, organisatie, wetgeving en het automatiseren van beschikkingen* (diss. Tilburg). Delft: Eburon 2000.

Zuiderveen Borgesius, F. (2018). *Discrimination, artificial intelligence, and algorithmic decision-making*. Strasbourg: Council of Europe, Directorate General of Democracy.