Tweede Kamer der Staten-Generaal

2

Vergaderjaar 2019–2020

26 643 Informatie- en communicatietechnologie (ICT)

32 761 Verwerking en bescherming persoonsgegevens

Nr. 642

BRIEF VAN DE MINISTER VAN BINNENLANDSE ZAKEN EN KONINKRIJKSRELATIES

Aan de Voorzitter van de Tweede Kamer der Staten-Generaal

Den Haag, 8 oktober 2019

Mede namens de Staatssecretaris van Binnenlandse Zaken en Koninkrijksrelaties, de Minister van Justitie en Veiligheid, de Minister voor Rechtsbescherming en de Staatssecretaris van Economische Zaken en Klimaat, stuur ik u deze beleidsbrief over AI, publieke waarden en mensenrechten. De brief geeft een overzicht van de kansen en risico's van AI voor publieke waarden die gestoeld zijn op mensenrechten. Het beschrijft tevens bestaande en toekomstige beleidsmaatregelen om risico's voor deze fundamentele publieke waarden te adresseren.

Tegelijkertijd met deze brief zijn het Strategisch Actieplan voor Artificiële intelligentie (SAPAI) (Kamerstukken 26 643 en 32 761, nr. 640) en de brief Waarborgen tegen risico's van data-analyses door de overheid (Kamerstukken 26 643 en 32 761, nr. 641) aan uw Kamer aangeboden. De drie brieven focussen op verschillende onderdelen van het brede vraagstuk ten aanzien van het benutten van kansen en het adresseren van risico's van AI. SAPAI bevat de overkoepelende AI-aanpak van dit kabinet en bevat beleidsmaatregelen om de maatschappelijke en economische kansen van AI te benutten en daarbij de publieke belangen te borgen. SAPAI gaat in spoor 3 kort in op de effecten van AI op publieke waarden. Omdat de effecten van AI op publieke waarden en mensenrechten complex zijn en in potentie ook significant, heeft het kabinet ervoor gekozen om in onderhavige brief nader aandacht te besteden aan beleid op dit vlak. Ditzelfde geldt voor de brief Waarborgen tegen risico's van data-analyses door de overheid, die in het bijzonder ingaat op mogelijke waarborgen tegen de risico's van het gebruik van algoritmes en data-analyses door de overheid.

kst-26643-642
ISSN 0921 - 7371
's-Gravenhage 2019

Onderhavige brief bouwt voort op de kabinetsreactie op het rapport van de Universiteit Utrecht over algoritmes en grondrechten[1] en is tot stand gekomen op basis van input van departementen, medeoverheden en wetenschappers. De brief voorziet tevens in de toezegging in kabinetsreactie op het rapport «Opwaarderen. Borgen van publieke waarden in de digitale samenleving» van het Rathenau Instituut om beleid op dit onderwerp verdergaand te ontwikkelen.[2] Ook wordt met deze brief in voldoende mate invulling gegeven aan de motie van de leden Van Dam en Van der Molen over digitalisering en publieke waarden, aangenomen op 6 juni 2018.[3] De voorstellen voor de versterking van coherentie van beleid en van toezicht en controlemechanismen in deze brief voorzien voor een deel in de initiatiefnota van lid Middendorp, ingediend op 29 mei[4], de motie van de leden Verhoeven en Van der Molen, aangenomen op 29 mei 2019[5] en de motie van de leden Middendorp en Drost, aangenomen op 20 juni 2019.[6]

AI is een dwarsdoorsnijdend thema dat de verantwoordelijkheid van alle departementen raakt. Voor het onderwerp «kansen en risico's van AI voor de bescherming van grondrechten» is de Minister van Binnenlandse Zaken en Koninkrijksrelaties de coördinerende bewindspersoon, terwijl iedere bewindspersoon verantwoordelijk is voor de ontwikkeling van beleid met betrekking tot deze kansen en risico's voor de bescherming van grondrechten op het eigen beleidsterrein, met inbegrip van de daarvoor benodigde wetgeving.[7] In deze brief staat dan ook beleid dat door verschillende departementen wordt ontwikkeld.

1. Mensenrechten als uitgangspunt voor AI-beleid

Nederland kent een hoog beschermingsniveau van mensenrechten, waar een uitgebreid stelsel van regels, voorzieningen, instellingen en verantwoordingsprocedures aan ten grondslag ligt.[8] Het zorgt ervoor dat mensenrechten niet alleen op papier, maar ook in de praktijk worden beschermd en bevorderd. Dat zorgt er vervolgens weer (mede) voor dat mensen graag in Nederland wonen, de samenleving vitaal is, bedrijven zich hier graag vestigen en onze democratische rechtstaat naar behoren kan functioneren. Nederland draagt door zijn mensenrechtencultuur bij aan het functioneren van de Europese Unie als gemeenschap van waarden. Het kabinet hecht zeer aan het beschermen van mensenrechten, ook in Europees en internationaal verband.
Inbreuken op mensenrechten door AI moeten worden voorkomen. Niet alle AI-inzet vraagt echter om een rol van de overheid. De focus van beleid ligt op terreinen waar AI-toepassingen een duidelijke impact op mensen hebben en/of op de maatschappij als geheel. In dergelijke gevallen zet het kabinet zich ervoor in om te zorgen dat AI-toepassingen mensenrechten respecteren en waar mogelijk zelfs versterken. Het kabinet staat dan ook een mensgerichte benadering van AI voor, internationaal ook wel een

---

[1] Kamerstuk 26 643, nr. 601.
[2] Kamerstuk CVIII, S.
[3] Kamerstuk 32 761, nr. 120; Handelingen II 2017/18, nr. 92, item 10.
[4] Kamerstuk 35 212, nr. 2
[5] Kamerstuk 26 643, nr. 610; Handelingen II 2018/19, nr. 94, item 30.
[6] Kamerstuk 35 200 VII, nr. 14; Handelingen II 2018/19, nr. 97, item 22.
[7] Voor specifiek de aspecten «transparantie, toetsbaarheid en rechtsbescherming» zijn de verantwoordelijkheden nader uitgewerkt in de brief van de Minister voor Rechtsbescherming over waarborgen tegen risico's van data-analyses door de overheid, die tegelijkertijd met deze brief naar de Tweede Kamer is verzonden.
[8] College voor de Rechten van de Mens. (2018). Mensenrechten in Nederland 2017 – Jaarlijkse rapportage, https://mensenrechten.nl/nl/publicatie/38613, Kamerstuk 33 826, nr. 25. Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. (2013). Nationaal actieplan mensenrechten, https://www.rijksoverheid.nl/documenten/publicaties/2013/12/10/nationaal-actieplan-mensenrechten.

human centred benadering genoemd, waarbij respect voor publieke waarden gestoeld op mensenrechten het uitgangspunt vormt achter het doel, ontwerp en gebruik van AI-toepassingen.

Tegelijkertijd worden we geconfronteerd met dilemma's die AI met zich meebrengt als het gaat om de borging van publieke waarden en mensenrechten. Bijvoorbeeld wanneer afwegingen nodig zijn tussen de bescherming van privacy en de bescherming van veiligheid. Of tussen de bescherming van privacy en het voorkomen van fraude. Deze situaties vergen een zorgvuldige en transparante afweging tussen mensenrechten en andere publieke belangen bij het formuleren van beleid en wetgeving.

2. AI-technologie, kansen en risico's en bestaand beleid[9]

AI refereert aan systemen die intelligent gedrag vertonen door hun omgeving te analyseren, de verzamelde data te interpreteren en op basis daarvan te bepalen welke actie het beste is om specifieke doelen te bereiken.[10] Een snel opkomende vorm van AI is *narrow-AI.* Dit zijn systemen die gericht zijn op de uitvoering van één taak (zoals beeld- of spraakherkenning). Voorbeelden hiervan zijn virtuele assistenten (zoals Siri), persoonlijke aanbevelingssystemen (zoals ingezet door Netflix) en patroonherkenning in beelden (zoals ingezet door medici bij diagnose). De ontwikkeling van het geavanceerdere AI-type *Artificial General Intelligence* systemen staat echter nog in de kinderschoenen. Deze zelfdenkende en zelfredenerende systemen bestaan vooralsnog alleen in theorie. Bovendien verschillen experts sterk in hun verwachting of en op welk moment deze vorm van intelligentie wordt gerealiseerd. Deze brief concentreert zich daarom op kansen, risico's en beleid ten aanzien van *narrow-AI* systemen.

Literatuur laat zien dat AI de maatschappij vele kansen biedt. In SAPAI wordt uitgebreid stilgestaan bij de bijdragen die AI kan leveren aan het adresseren van maatschappelijke uitdagingen zoals op het gebied van een effectievere gezondheidszorg, klimaatbeheersing en betere veiligheid. AI biedt daarnaast allerlei kansen om specifieke mensenrechten te versterken. Zo kan het recht op informatie versterkt worden, doordat AI-systemen informatie kunnen personaliseren en hiermee relevanter kunnen maken voor gebruikers. AI kan ook het verbod van discriminatie bevorderen, bijvoorbeeld wanneer het wordt ingezet om vooroordelen uit een selectieproces te elimineren. Ook kunnen AI-toepassingen individuele autonomie versterken. Zo zijn er AI-gebaseerde apps die mensen helpen zelf een medische diagnose te stellen.

Naast kansen, brengen AI-ontwikkelingen ook potentiële risico's met zich mee. Onderstaande tabel geeft een overzicht van risico's die meest dominant in de literatuur aanwezig zijn.

Publieke waarde	Omschrijving	Risico
Verbod van discriminatie	Mensen moeten in gelijke gevallen gelijk behandeld worden, en mogen niet op basis van bepaalde kenmerken ten onrechte worden uitgesloten.	– Bias in onderliggende data, hetgeen leidt tot discriminerende patronen – Bias in een algoritme, hetgeen leidt tot discriminerende patronen – Foutmarges die leiden tot onjuiste classificatie

[9] Deze paragraaf vormt een samenvatting van de bijlage.
[10] Op basis van Kamerstuk 22 112, nr. 2578, p.2, aangepast naar hedendaags inzicht.

Publieke waarde	Omschrijving	Risico
Privacy	Mensen moeten onbevangen «zichzelf» kunnen zijn en doen en laten wat zij willen, zonder bemoeienis van derden.	– Grote hoeveelheid data benodigd voor goede uitkomsten van AI-systemen – Sensitieve data die gegenereerd worden door AI-systemen
Vrijheid van meningsuiting	Iedereen heeft het recht om overtuigingen, gevoelens en meningen onder woorden te brengen en te delen met anderen. Hieronder valt ook het recht op toegang tot (gebalanceerde) informatie.	– Beperkte toegang tot en pluriformiteit van informatie – Onnauwkeurige algoritmen die content te snel verwijderen
Menselijke waardigheid	Het enkele «zijn» van mens gaat gepaard met een bepaalde waardigheid, die een beschermingsniveau ten opzichte van de overheid en derden garandeert. Een belangrijk onderdeel is menselijk contact.	– Afname van intermenselijk (en daarmee de kwaliteit van) contact wanneer AI interactie overneemt
Persoonlijke autonomie	Een mens moet vrijelijk keuzes kunnen maken en grotendeels zelf kunnen bepalen hoe hij zijn leven inricht.	– Ongemerkte beïnvloeding door sturende AI
Recht op een eerlijk proces	Iedereen moet toegang hebben tot het recht; tot informatie, advies, begeleiding bij onderhandeling, rechtsbijstand en de mogelijkheid van een beslissing door een neutrale (rechterlijke) instantie.	– Ondoorzichtigheid van algoritmen waardoor individuen moeilijker kunnen opkomen voor hun recht

Er bestaat al veel beleid om deze risico's te adresseren. De bijlage bij deze brief geeft een uitgebreid overzicht van bestaand beleid. Er wordt door verschillende instanties gewerkt aan standaarden om de kwaliteit van AI-systemen – mede op het punt van bias en foutmarges – te garanderen.[11] Zo brengt de Minister voor Rechtsbescherming gelijktijdig een brief uit over mogelijke wettelijke waarborgen tegen risico's van data-analyses door de overheid en richtlijnen voor de toepassing van algoritmes door overheden, die mede ingaat op de mogelijkheid om kwaliteitswaarborgen in wetgeving en richtlijnen op te nemen. Voor de ontwikkeling en inzet van AI-systemen zijn allerlei *Privacy-by-design* concepten beschikbaar. Het kabinet steunt daarnaast onderzoek naar AI-systemen die minder data nodig hebben om tot een kwalitatief hoogwaardige uitkomst te komen. Verder laat het Ministerie van JenV door Tilburg University onderzoek doen naar maatregelen om risico's van gezichtsherkenning voor de privacy, waarbij AI eveneens een rol kan spelen, te beperken.

Hoewel op AI gebaseerde nieuwspersonalisatie in Nederland nog geen grote negatieve impact heeft, investeert het kabinet in mediawijsheid en publiekscampagnes om het publiek bewust te maken van de rol die personalisering kan spelen bij het nieuwsaanbod. Wat betreft het tegengaan van illegale content met behulp van AI, hanteert Nederland een «Notice-And-Take-Down» procedure en neemt hierbij maatregelen om te vermijden dat legale inhoud onbedoeld wordt verwijderd. Als het gaat om menselijke waardigheid, heeft het kabinet in zijn agenda NL DIGIbeter aangegeven dat zinvol contact – en dus niet uitsluitend digitaal – uitgangspunt van overheidsbeleid is. Voor het contact tussen burgers en bedrijven heeft het kabinet de WRR verzocht om te adviseren over de impact van AI op menselijk contact. Onderzoek laat daarnaast zien dat de overheid AI momenteel niet inzet om gedrag te beïnvloeden. Bedrijven doen dat wel steeds vaker. Het kabinet is daarom voornemens onderzoek

[11] Tevens heeft er onlangs een inventarisatie plaatsgevonden omtrent standaarden voor gegevensuitwisseling tussen overheden. Het kabinet bereidt hier een reactie op voor.

te doen naar de effecten hiervan. Wat betreft het recht op een eerlijk proces en de ondoorzichtigheid van algoritmen zijn er verschillende beleidsinitiatieven, zoals de hiervoor genoemde richtlijnen voor de toepassing van algoritmes door overheden en een Transparantielab waarin geëxperimenteerd wordt met vormen van transparantie van algoritmen. Naar aanleiding van de vraag van het lid Futselaar tijdens het plenaire debat op 10 september 2019 (Handelingen II 2018/19, nr. 106, item 26) zal in relatie tot de norm uitlegbaarheid tevens aandacht worden besteed aan de mate en gevallen waarin de beslisregels van het algoritme voor de burger inzichtelijk worden gemaakt.

Naast bovenstaand beleid dat veelal voor de borging van een specifiek publieke waarde wordt ingezet, zijn er meer algemene beleidslijnen om publieke waarden en mensenrechten bij AI-ontwikkelingen te borgen. Deze liggen onder meer op het gebied van het versterken van begrip en bewustzijn onder burgers middels dialoog, stimuleren van vormen van zelfregulering zoals gedragscodes en op het gebied van internationale agendering en onderhandeling

3. Aanvullende maatregelen

Voorgaande paragraaf laat zien dat het kabinet kansen en risico's in beeld heeft en daar actief beleid op voert. Toch kan het beleid op een aantal punten verder versterkt worden, met name als het gaat om de coördinatie van de borging van publieke waarden en mensenrechten. Mijn ministerie heeft hier een belangrijke verantwoordelijkheid in. Beleid kan aan effectiviteit winnen, wanneer het in meer samenhang wordt ontwikkeld. Ook kan de concretisering, van de nu nog abstracte *human centred AI* concepten, verschillende instanties helpen om rechten beter mee te nemen in de ontwikkeling van AI-toepassingen. Bovendien kunnen toezicht en controlemechanismen mogelijk versterkt worden en kunnen standpunten internationaal sterker uitgedragen worden. Hieronder worden daartoe voorstellen gedaan.

Meer samenhang in beleid

Om de kansen van AI te kunnen benutten en daarnaast goed voorbereid te zijn op de risico's van AI-toepassingen is aandacht nodig voor de organisatie van de borging en versterking van publieke waarden. Zoals gesteld worden reeds veel maatregelen genomen om de risico's van AI voor publieke waarden te adresseren. Tegelijkertijd constateert het kabinet dat de samenhang tussen beleidsinitiatieven beter kan, omdat deze in gevallen overlappen of juist versnipperd zijn. Overheden kunnen effectiever opereren wanneer kennis wordt gedeeld en beleid intensiever wordt afgestemd. In het komende jaar wordt daarom door mijn ministerie een samenwerkingsplatform voor de overheid opgericht ten aanzien van het onderwerp AI en publieke waarden. Via dit platform wordt kennis uitgewisseld, beleidsafstemming gefaciliteerd, verbinding gelegd met de wetenschap en toegewerkt naar een gezamenlijke onderzoeksprogrammering.

Concretisering human centred AI naar systeemconcepten

Zoals gesteld staat het kabinet een mensgerichte benadering van AI voor, hetgeen in lijn is met de *human centred AI* benadering in Europa.[12] Huidige *human centred AI* concepten zijn echter erg abstract en algemeen. Het is voor ontwikkelaars van AI-toepassingen daardoor lastig om concrete publieke waarden en mensenrechten in hun ontwerp mee te

[12] https://ec.europa.eu/digital-single-market/en/news/coordinated-plan-artificial-intelligence

nemen. Hiervoor dient duidelijk te zijn hoe specifieke waarden en rechten in een AI-ontwerp vertaald kunnen worden. Mijn ministerie zal in concrete use cases kijken hoe publieke waarden en mensenrechten geoperationaliseerd kunnen worden tot systeemprincipes en deze beschikbaar stellen aan overheden en bedrijfsleven. Om te beginnen op het gebied van non-discriminatie.

Versterking controlemechanismen

Voorgaande paragraaf laat zien dat verschillende partijen bezig zijn met het verbeteren van controlemechanismen op AI (zoals standaarden en audits).[13] Het is nog te vroeg om te beoordelen of deze controlemechanismen effectief en wenselijk zijn, laat staan om er nu al een verplichtend karakter aan te geven. Toezicht is één van de controlemechanismen op het juiste gebruik van algoritmes. Bij toezicht op algoritmes is een veelheid aan toezichthouders betrokken. De Staatssecretaris van BZK en de Minister voor Rechtsbescherming gaan samen onderzoeken of toezichthouders voldoende toegerust zijn om toezicht op algoritmes te kunnen houden en of er toch nog blinde vlekken zijn in het toezichtlandschap. Daarnaast willen de Staatssecretaris van BZK en de Minister voor Rechtsbescherming de start die al is gemaakt om toezichthouders in een samenwerkingsverband van elkaars expertise op het gebied van algoritmes en AI te laten leren verdergaand stimuleren.

Internationaal uitdragen belang publieke waarden in human centred AI

Een laatste punt betreft het internationaal uitdragen van het belang van publieke waarden en mensenrechten en *best practices* op het gebied van *human centred AI*. Door vele landen wordt stevig geïnvesteerd in AI. Zo heeft China de ambitie om in 2030 wereldleider in AI-innovatie te zijn en zegt de VS zijn leiderschap in AI-innovatie te willen behouden. Nederland zal – samen met andere Europese landen – nadrukkelijker uitdragen dat AI publieke waarden en mensenrechten moet versterken in plaats van verzwakken en hiervan best practices laten zien. *Human centred AI* kan net als andere waarden-gedreven technologieën (zoals op het gebied van wind- en zonne-energie, maar ook in de landbouw) een essentieel exportproduct worden. Nederland zou daarmee tot de koplopers op het gebied van *human centred AI* kunnen behoren.

De Minister van Binnenlandse Zaken en Koninkrijksrelaties,
K.H. Ollongren

[13] Voor meer voorbeelden, zie ook pagina 8.

Artificiële intelligentie

AI refereert aan systemen die intelligent gedrag vertonen door hun omgeving te analyseren, de verzamelde data te interpreteren en op basis daarvan te bepalen welke actie het beste is om specifieke doelen te bereiken.[14] Er worden grofweg twee typen AI-systemen onderscheiden, namelijk: *»rule-based»* en *»machine learning»*. *Rule-based*-systemen komen tot beslissingen op basis van vóóraf gedefinieerde regels en leren niet van de data die ze verwerken. *Machine learning*-systemen «leren» daarentegen wel regels door patronen af te leiden uit data.[15] Vooral op het gebied van *machine learning* zijn de afgelopen jaren grote stappen gezet door nieuwe technologische inzichten, de grotere verwerkingscapaciteit van computers en de grotere beschikbaarheid van data.

Machine learning-systemen kunnen weer onderverdeeld worden in *narrow AI*- en *Artificial General Intelligence* systemen. Waar *narrow AI*-systemen gericht zijn op de uitvoering van één taak (zoals beeldherkenning) zijn *Artificial General Intelligence* systemen in staat om algemene taken uit te voeren. De ontwikkeling van *Artificial General Intelligence* systemen staat echter nog in de kinderschoenen. Deze zelfdenkende, zelfredenerende, zelflerende en wellicht zelfs zelfbewuste systemen bestaan vooralsnog alleen in theorie. Bovendien verschillen experts sterk in hun verwachting of en op welk moment deze vorm van intelligentie wordt gerealiseerd.[16]

De mate waarin de technologie klaar is voor gebruik, is van steeds meer *narrow AI*-systemen hoog. De prestaties van deze systemen benaderen die van de mens of overtreffen die zelfs. Het aantal *narrow AI*-toepassingen is de afgelopen jaren dan ook explosief gegroeid. Voorbeelden zijn virtuele assistenten (mondelinge beantwoording van vragen, zoals Siri en Cortana), persoonlijke aanbevelingssystemen (zoals ingezet door Netflix en Amazon) en patroonherkenning in beelden (zoals ingezet door medici bij diagnose).

Kansen, risico's en bestaand beleid per publieke waarde

Deze paragraaf benoemt kansen en risico's[17] van AI voor op mensenrechten gestoelde publieke waarden en beschrijft bestaand beleid per publieke waarde.

---

[14] Op basis van Kamerstuk 22 112, nr. 2578, p.2, aangepast naar hedendaags inzicht.

[15] Leren betreft hier het berekenen van gewichten in een wiskundige formule en wordt ook wel trainen genoemd.

[16] Müller, V. & Bostrom, N. (2016). Future progress in artificial intelligence: A survey of expert opinion. In Fundamental Issues of Artificial Intelligence (pp. 553–571). Berlin: Springer.

[17] Op basis van literatuur en rapporten zoals: Vetzo, M.J., Gerards, J.H. en Nehmelman, R. (2018), Algoritmes en grondrechten, Universiteit Utrecht. Kool, L., J. Timmer, L. Royakkers en R. van Est, Opwaarderen – Borgen van publieke waarden in de digitale samenleving. Den Haag, Rathenau Instituut 2017, Van Est, R. & J.B.A. Gerritsen, with the assistance of L. Kool, Human rights in the robot age: Challenges arising from the use of robotics, artificial intelligence, and virtual and augmented reality – Expert report written for the Committee on Culture, Science, Education and Media of the Parliamentary Assembly of the Council of Europe (PACE), The Hague: Rathenau Instituut 2017, Raso, Filippo and Hilligoss, Hannah and Krishnamurthy, Vivek and Bavitz, Christopher and Kim, Levin Yerin, Artificial Intelligence & Human Rights: Opportunities & Risks (September 25, 2018). Berkman Klein Center Research Publication No. 2018–6. Algemene, maatschappelijke kansen van AI worden onder meer vermeld in: Jha, K. et al. (2019). A comprehensive review on automation in agriculture using artificial intelligence. Artificial Intelligence in Agriculture. 2 (juni). 1–12. Topol. E. (2019). High-performance medicine: the convergence of human and artificial intelligence. Nature Medicine. 25. 44–56. Zhang et al. (2016). A study on key technologies of unmanned driving. CAAI Transactions on Intelligence Technology. 1 (1). 4–13.

Verbod van discriminatie

> Het verbod van discriminatie verbiedt ongelijke behandeling van mensen in gelijke gevallen. Ongelijke behandeling is vooral problematisch wanneer sprake is van een benadeling van het ene (individu/groep of geval) ten opzichte van het andere zonder goede rechtvaardiging.

Kansen en risico's

AI biedt kansen voor het tegengaan van discriminatie. Zo bestaan er AI-toepassingen om vooroordelen in selectieprocessen te verminderen (zoals instrumenten die werkgevers waarschuwen voor potentieel bevooroordeelde taal in functiebeschrijvingen). AI-toepassingen kennen echter ook risico's voor het verbod van discriminatie, met name wanneer sprake is van bias in onderliggende data, bias in het algoritme of bij onjuiste classificatie.[18] Er is sprake van bias in data wanneer er (onbedoeld) discriminerende patronen in voorkomen. Wanneer een algoritme voor werving van medewerkers wordt getraind met een dataset waarin (aanzienlijk) meer mannen promotie hebben gekregen, kan dit er bijvoorbeeld toe leiden dat het algoritme de kansen op geschiktheid bij vrouwen lager inschat. Er is sprake van bias in een algoritme wanneer (onbewuste) vooroordelen van experts worden vertaald in het ontwerp van het algoritme of in de keuze van variabelen. Ook hier kan de bias leiden tot onrechtmatige benadeling van groepen. Daarnaast kennen AI-systemen die classificeren per definitie een foutmarge. Dit komt omdat deze systemen constateringen voor groepen toepassen op het individu. Er zijn daardoor altijd gevallen die volgens het model onder een bepaalde categorie vallen, maar in werkelijkheid daar niet onder passen. Hierdoor kunnen mensen verkeerd worden ingedeeld; bijvoorbeeld onterecht als kredietwaardig (*false positive*), of onterecht niét als kredietwaardig (*false negative*).

Bestaand beleid

Er wordt door verschillende instanties gewerkt aan standaarden om de kwaliteit van AI-systemen – mede op het punt van bias en foutmarges – te garanderen. Zo werkt het internationale Institute of Electrical and Electronics Engineers (IEEE)[19] aan standaarden om ongewenste bias in data en algoritmes te voorkomen.[20] Daarnaast heeft het Nederlandse normeninstituut NEN eind 2018 een normcommissie Artificial Intelligence en Big Data opgericht. Via deze commissie wil de NEN invloed uitoefenen op de ISO standaardisatie, onder andere ten aanzien van het beperken van bias in AI-systemen, risk management in AI, betrouwbaarheid en robuustheid van AI.[21] Het Ministerie van JenV brengt gelijktijdig een brief uit over mogelijke wettelijke waarborgen tegen risico's van data-analyses door de overheid en richtlijnen voor de toepassing van algoritmes door overheden. Deze brief gaat mede in op de mogelijkheid om kwaliteits-waarborgen in wetgeving en richtlijnen op te nemen. Onder leiding van mijn ministerie wordt gewerkt aan ethische dataprincipes voor de overheid, waar het voorkomen van discriminatie onderdeel van is.[22] Deze dataprincipes worden, naast andere principes en richtlijnen, getest in het

---

[18] Bij deze risico's dient de kanttekening geplaatst te worden dat ook mensen niet vrij zijn van vooroordelen en daardoor soms (onbewust) discrimineren. Een belangrijk risico is echter dat AI-systemen dit kunnen repliceren en automatiseren.

[19] Internationale vereniging van professionals op het gebied van elektronica, elektriciteit en informatica.

[20] https://standards.ieee.org/project/7003.html

[21] https://www.iso.org/committee/6794475.html

[22] Kamerstuk 26 643, 597.

Transparantielab dat mijn ministerie initieert. Ook decentrale overheden trachten hun controle op AI-systemen te vergroten. Zo kijkt de gemeente Amsterdam of de kwaliteit van AI-systemen kan worden gescreend.[23] De inspectie SZW laat onderzoek verrichten naar geautomatiseerde systemen bij werving- en selectieprocedures, de risico's van discriminatie daarbij en mogelijkheden voor toezicht. Naast overheden werken bedrijven aan het voorkomen van bias en foutmarges. IBM ontwikkelde bijvoorbeeld de AI Fairness 360 toolkit om bias in datasets en algoritmes te detecteren.[24]

Privacy en gegevensbescherming

> In Nederland heeft iedereen recht op een privéleven en de eerbiediging van de persoonlijke levenssfeer, veelal privacy genoemd. Individuen kunnen daarbij onbevangen zichzelf zijn en doen en laten wat zij willen, zonder bemoeienis van derden. Voor AI-systemen is met name de bescherming van persoonsgegevens van belang, omdat AI-systemen voor hun functioneren vaak van data over personen afhankelijk zijn.

Kansen en risico's

AI-toepassingen kunnen bijdragen aan het versterken van privacy. Bijvoorbeeld wanneer zij worden ingezet om privacyverklaringen van bedrijven te evalueren op de mate van naleving van de Algemene Verordening Gegevensbescherming (AVG). Zo'n programma kan bedrijven helpen hun privacybeleid te verbeteren, kan mensen ondersteunen bij het opkomen voor hun privacybelangen en kan instellingen aanzetten tot actie.[25] Er zijn echter ook risico's op het gebied van privacy, doordat AI-systemen vaak gebruik maken van grote hoeveelheden data om antwoorden, voorspellingen en patronen te genereren. Het gebruik van meer gegevens leidt in potentie tot betere resultaten en tot nieuwe inzichten. Dit kan een prikkel zijn voor organisaties die AI-systemen inzetten om zo veel mogelijk gegevens te verzamelen, wat leidt tot een grotere inbreuk op de privacy. Daarnaast kunnen AI-systemen op basis van combinaties van (niet sensitieve) gegevens zeer persoonlijke kenmerken van individuen in kaart brengen. Zo konden wetenschappers op basis van Facebook-likes voorspellingen doen over o.a. seksuele geaardheid en politieke voorkeur.[26] Sommige AI-systemen zijn in staat om emoties te herkennen.[27]Dergelijke zeer sensitieve data – die voorheen voor derden niet kenbaar waren – kunnen kenbaar worden. Ook dit zorgt voor een inbreuk op de privacy.

Bestaand beleid

Voor de ontwikkeling en inzet van AI-systemen zijn allerlei *Privacy by Design* principes, raamwerken en handleidingen beschikbaar. Zo heeft het Centrum voor Informatiebeveiliging en Privacybescherming een

---

[23] https://fd.nl/ondernemen/1291305/amsterdam-wil-eerlijke-computers-in-de-stad#

[24] https://www.ibm.com/blogs/research/2018/09/ai-fairness-360/

[25] Zie: CLAUDETTE meets GDPR Automating the Evaluation of Privacy Policies using Artificial Intelligence 2018, http://www.claudette.eu/gdpr/#, en Hamza Harkous c.s., Polisis: Automated Analysis and Presentation of Privacy Policies Using Deep Learning,https://www.wired.com/story/polisis-ai-reads-privacy-policies-so-you-dont-have-to/

[26] Kosinski, M., Stillwell, D. en Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior, Proceedings of the National Academy of Sciences of the United States of America, 110 (15) 5802–5805.

[27] Cambria, Erik (March 2016). «Affective Computing and Sentiment Analysis». IEEE Intelligent Systems. 31 (2): 102–107, Sharef, Nurfadhlina Mohd; Zin, Harnani Mat; Nadali, Samaneh (1 March 2016). «Overview and Future Opportunities of Sentiment Analysis Approaches for Big Data». Journal of Computer Science. 12 (3): 153–168.

handleiding ontwikkeld om organisaties te helpen *Privacy by Design* principes toe te passen.[28]Tegelijkertijd neemt het aantal technische mogelijkheden toe om gegevens te verwerken zonder inbreuk te maken op privacy, bijvoorbeeld door de toepassing van *secure multi-party computation*, en *self sovereign identity*. Onderzoek laat bovendien zien dat toekomstige AI-systemen mogelijk minder data nodig hebben om tot een kwalitatief hoogwaardige uitkomst te komen.[29] Het kabinet ondersteunt dit type onderzoeken. Gezichtsherkenning is een AI-toepassing die specifieke aandacht vergt, omdat burgers hiervan steeds meer gebruik maken, onder meer via apps en sociale media. Het eerder genoemde onderzoek hiernaar dat wordt uitgevoerd door Tilburg University, zal mede ingaan op maatregelen om risico's te beperken en zal naar verwachting eind dit jaar gereed zijn.[30]
Het gebruik van persoonsgegevens voor AI-systemen bevestigt ook de noodzakelijkheid om de digitale veiligheid op orde te hebben. Organisaties die werken met persoonsgegevens kunnen immers slachtoffer worden van hackaanvallen door cybercriminelen of statelijke actoren. Om dit risico tegen te gaan moet naast het tegengaan van de dreiging worden ingezet op de verhoging van de digitale weerbaarheid. Het kabinet heeft daartoe maatregelen aangekondigd in de Nederlandse Cyber Security Agenda en zet zoals aangekondigd in de beleidsreactie op het Cybersecurity Beeld Nederland 2019 in op verdere versterking van de digitale weerbaarheid van vitale infrastructuur en de rijksoverheid.[31]

Vrijheid van meningsuiting

> Iedereen heeft het recht om overtuigingen, gevoelens en meningen onder woorden te brengen en te delen met anderen. Hierbij hoort ook het recht op toegang tot informatie.

Kansen en risico's

Door de inzet van AI kunnen nieuwssites, zoekmachines en aanbevelingssystemen informatie personaliseren. Dit helpt de gebruiker om relevante informatie te vinden, te delen en beoordelen, wat het recht op informatie en de vrijheid van meningsuiting kan bevorderen. Hieraan zijn echter ook risico's verbonden. De inzet van AI-systemen bij het aanbieden van informatie kan ertoe leiden dat individuen een beperkt informatieaanbod krijgen. Zo geven zoekmachines populaire bronnen, Engelstalige sites en commerciële informatiebronnen veelal een dominantere (hogere) positie in zoekresultaten.[32] Daarnaast kunnen gebruikers door personalisatie vooral informatie te zien krijgen die past bij hun denkbeelden en hiermee deze denkbeelden versterken (*confirmation bias, filter bubble*).
AI-systemen kunnen beperkt worden ingezet voor het verwijderen van strafbare en onrechtmatige content. Bij het detecteren van kinderporno kan AI een belangrijke rol spelen; bij het opsporen van terrorisme, discriminatie en onrechtmatige content – waarbij vooral de context waarin de content wordt gepresenteerd een belangrijke rol speelt – is nog altijd menselijke interventie noodzakelijk. Algoritmes kunnen in die gevallen wel leiden tot een alertering, die vervolgens door een menselijke assessor beoordeeld kan worden. Het kabinet ondersteunt initiatieven om algoritmes te ontwikkelen die meer nuance kunnen meenemen.

---

[28] https://www.rijksoverheid.nl/documenten/rapporten/2018/02/19/handleiding-privacy-by-design
[29] https://hbr.org/2019/01/the-future-of-ai-will-be-about-less-data-not-more
[30] Kamerstuk 34 926, nr. 8.
[31] Nederlandse Cybersecurity Agenda (Kamerstuk 26 643, nr. 536), Cybersecuritybeeld Nederland 2019 en Voortgangsrapportage NCSA (Kamerstuk 26 643, nr. 614).
[32] T. Gillespie, «The Relevance of Algorithms», in: T. Gillespie, P. J. Boczkowski & K. A. Foot (red.), Media technologies: Essays on communication, materiality, and society, Cambridge: MIT Press 2014, p. 167–194.

Bestaand beleid

De effecten van AI op de vrijheid van meningsuiting werden door het vorige kabinet reeds onderkend.[33] Het kabinet beschreef eerder de toenemende personalisering in de media (o.a. door de inzet van AI) en het toenemend aandeel van de bevolking dat nieuws consumeert via sociale media. Uit recent onderzoek[34] blijkt echter dat er in Nederland nog geen grote negatieve impact van nieuwspersonalisatie is, o.a. vanwege het sterke mediabestel en de pluriforme nieuwsconsumptie. Het risico is relatief klein dat online «filter bubbels» ontstaan waarin mensen zich eenzijdig informeren. Om risico's verder te minimaliseren investeert het kabinet o.a. in mediawijsheid en publiekscampagnes (zoals www.blijfkritisch.nl) om het publiek bewust te maken van de rol die personalisering kan spelen bij het nieuwsaanbod dat het te zien krijgt. In mei 2019 heeft de Raad voor het Openbaar Bestuur een adviesrapport uitgebracht met de titel «Zoeken naar waarheid». Daarin wordt de invloed van digitalisering op de democratie beschreven, vanuit het oogpunt van waarheidsvinding. Mijn ministerie bereidt een reactie voor op dit rapport die naar verwachting voor het einde van het jaar verschijnt. Wat betreft het tegengaan van illegale online content hanteert Nederland een «Notice-And-Take-Down» procedure. Via deze procedure worden burgers en bedrijven gestimuleerd en/of verplicht om illegale online content zo snel mogelijk van hun platform te verwijderen. Hoewel het verwijderen van illegale content door private partijen acceptabel kan zijn als dit noodzakelijk is om zwaarwegende algemene belangen te beschermen, is het van belang dat dit met waarborgen omkleed is. De Europese Commissie heeft maatregelen aanbevolen om te vermijden dat legale inhoud onbedoeld wordt verwijderd.[35] Zo zou een aanbieder van content bijvoorbeeld op de hoogte moeten worden gesteld van verwijdering en bezwaar daartegen kunnen aantekenen. Daarnaast moeten bedrijven doeltreffende en passende waarborgen inbouwen om ongeoorloofde inperkingen op de vrijheid van meningsuiting zo veel mogelijk te voorkomen.

Menselijke waardigheid

> Het enkele «zijn» van een mens gaat gepaard met een bepaalde waardigheid, waarvoor een bepaald beschermingsniveau ten opzichte van de overheid en derden moet worden gegarandeerd. Menselijke waardigheid wordt voornamelijk gehanteerd in de context van fysieke en psychologische integriteit, individuele autonomie, toegang tot de rechter, materiële levensomstandigheden en gelijkheid.

Kansen en risico's

AI-toepassingen kunnen bijdragen aan menselijke waardigheid, bijvoorbeeld wanneer zij gevaarlijk of mensonterend werk overnemen. AI kan echter ook tot risico's leiden wanneer zij contact tussen mensen vervangt. Persoonlijke digitale assistenten en *chat bots* die contact hebben met mensen roepen bijvoorbeeld de vraag op of voor mensen steeds duidelijk moet zijn dat zij met een mens of een machine communiceren.
AI-systemen die worden ingezet voor de werving van personeel roepen de vraag op in welke mate sollicitanten recht hebben op menselijk contact. In

[33] Kamerstuk 32 827, nr. 116.
[34] Zie o.a. Rathenau, 2018, «Digitalisering van het nieuws; online nieuwsgedrag, desinformatie en personalisatie» van het Rathenau Instituut en Commissariaat voor de Media, bijlage 3 bij Kamerstuk 32 827, nr. 127.
[35] Kamerstukken 22 112, nr. 2420.

dergelijke gevallen is de vraag hoe de menselijke waardigheid gewaarborgd blijft.

Bestaand beleid

In zijn agenda NL DIGIbeter[36] stelt de Staatssecretaris van BZK dat zinvol contact met de overheid het vertrekpunt is bij digitaal overheidsbeleid. Het streven is dat informatie en dienstverlening van de overheid toegankelijker, begrijpelijker, inclusiever en persoonlijker wordt – en niet uitsluitend digitaal.[37] Voor wat betreft AI-ontwikkelingen en zinvol contact tussen burgers en bedrijven heeft het kabinet de Wetenschappelijke Raad voor het Regeringsbeleid (WRR) verzocht om onder meer te adviseren over de impact van AI op menselijk contact. Het onderzoek van de WRR is in het najaar van 2018 van start gegaan en zal naar verwachting in 2020 gereed zijn.

Persoonlijke autonomie

> Persoonlijke autonomie houdt in dat een mens vrijelijk keuzes kan maken en grotendeels zelf kan bepalen hoe hij of zij het leven inricht. Het individu dient in een rechtsstaat ruimte te hebben om zijn eigen afwegingen te mogen maken; waar hij woont, hoe hij leeft, welk geloof hij wel of niet aanhangt, welke opleiding hij kiest, welke baan hij accepteert, etc.

Kansen en risico's

AI-toepassingen kunnen de autonomie versterken. Zo helpen AI-gebaseerde digitale coaches mensen om de kwaliteit van hun leven te verbeteren en zijn er AI-gebaseerde apps die mensen helpen zelf een medische diagnose te stellen. Er zijn echter ook AI-toepassingen die de autonomie van mensen beperken door hen (ongemerkt) ongewenst te beïnvloeden. Dit speelt vooral in gevallen waarin systemen personen beperkte of sturende informatie en/of keuzes voorleggen of (al dan niet ongemerkt) in de richting van bepaalde voorkeuren, keuzes en gedragingen duwen (nudging).

Bestaand beleid

De overheid zet AI momenteel niet in om gedrag te beïnvloeden.[38]Er zijn wel duidelijke uitgangspunten voor gedragsbeïnvloeding als overheidsinstrument. Het kabinet heeft in eerdere brieven aangegeven dat de overheid bij gedragsbeïnvloeding (zoals het bevorderen van een gezondere levensstijl) altijd rekening moet houden met de normatieve overwegingen en rechtsstatelijke aspecten (legaliteit, proportionaliteit en behoorlijke bestuur).[39] Ook moet de overheid transparant zijn over de beleidsdoelen van de gedragsbeïnvloeding. Bedrijven zetten AI wel steeds vaker voor gedragsbeïnvloeding in.[40] Het kabinet ziet deze ontwikkeling en de risico's die daarmee samenhangen en is voornemens onderzoek te doen naar de effecten van (met name persuasieve technologieën) op autonomie.

---

[36] Kamerstukken 26 643, nr. 549.
[37] Kamerstukken 26 643, nr. 583.
[38] Onderzoek van TNO «Quick-scan AI in de publieke dienstverlening» laat zien dat bijna alle onderzochte overheidstoepassingen van AI in de sfeer van «predictive analytics» liggen.
[39] Kamerstuk 34 000 XIII, nr. 140.
[40] Kool, L., J. Timmer, L. Royakkers en R. van Est, Opwaarderen – Borgen van publieke waarden in de digitale samenleving. Den Haag, Rathenau Instituut 2017.

Procedurele rechten

> Voor een goed functionerende rechtsstaat moet iedereen toegang hebben tot het recht. Mensen moeten toegang hebben tot informatie, advies, begeleiding bij onderhandeling, rechtsbijstand en de mogelijkheid van een beslissing door een neutrale (rechterlijke) instantie.

Kansen en risico's

In de rechtspleging biedt de inzet van AI kansen om de efficiëntie te vergroten.[41] Zo voert het OM een experiment uit om met AI jurisprudentie te analyseren met als doel een officier voor te bereiden op een zaak of zitting (project Jurisprudentierobot). AI-toepassingen kunnen echter ook leiden tot risico's voor procedurele rechten. Algoritmes die worden toegepast in AI-systemen kunnen complex zijn. Hierdoor is de uitleg van, of de controle op de werking van AI-systemen lastig, of soms helemaal niet mogelijk. Daarnaast worden algoritmes vanwege commerciële redenen in sommige gevallen geheimgehouden. Dit maakt het voor burgers lastig om bezwaar te maken tegen de werking of uitkomst van deze AI-systemen.

Bestaand beleid

Beleidsmatig gezien zijn er verschillende initiatieven.[42] De eerdergenoemde brief over mogelijke wettelijke waarborgen tegen risico's van data-analyses door de overheid en richtlijnen voor de toepassing van algoritmes door overheden van het Ministerie van JenV is ook hier relevant. Deze brief biedt richtlijnen voor transparantie van algoritmen voor overheden. Het door mijn ministerie ingerichte Transparantielab toetst de richtlijnen en zal deze vervolgens verder operationaliseren in een online applicatie die de gebruiker ondersteunt bij het toepassen ervan. Daarnaast werkt mijn ministerie aan een Code Goed Digitaal Openbaar Bestuur, waarin onder meer wordt ingegaan op het principe van transparantie. De Code zal een plek krijgen in de bestaande Code Goed Openbaar Bestuur.

Algemeen beleid om publieke waarden en mensenrechten te borgen

Naast het hierboven beschreven beleid per publieke waarde, zijn er meer algemene beleidslijnen ingezet om publieke waarden en mensenrechten bij AI-ontwikkelingen te borgen. In deze paragraaf wordt een opsomming van beleidslijnen gegeven.

Dialoog en mediawijsheid

Bij technologische ontwikkelingen is het van belang dat burgers weten wat de ontwikkelingen inhouden en wat ze kunnen doen om kansen te benutten en risico's voor rechten te adresseren. Het kabinet lanceert daarom bewustwordingscampagnes. Een voorbeeld is «http://www.blijfkritisch.nl» waarbij burgers alert worden gemaakt op (door AI gegenereerde) online desinformatie.[43] Daarnaast organiseert het kabinet

---

[41] Kamerstuk 34 775 VI, nr. AH, Uitgangspunt van het Ministerie van JenV bij AI in de rechtspraak is dat er op incrementele wijze experimenten plaatsvinden, voorlopig buiten de context van reële, aanhangige rechtszaken of geschillen.

[42] In de brief «AI en algoritmes in de rechtspleging» wordt uitgebreider ingegaan op regelgeving en beleid t.a.v. AI in de rechtspleging.

[43] https://www.rijksoverheid.nl/onderwerpen/desinformatie-nepnieuws

dialogen met burgers over AI-ontwikkelingen. Zo was er tijdens de Conferentie Nederland Digitaal[44] een burgerdialoog over de kansen en risico's van AI voor grondrechten. Het kabinet ontwikkelt ook voorlichtingsmateriaal voor specifieke doelgroepen. Mijn ministerie bracht een speciale editie van de Donald Duck uit over AI en grondrechten. Het Ministerie van OCW voert algemeen beleid om mediawijsheid onder jongeren te vergroten. In maart lanceerde het Ministerie van OCW de digitaliseringsagenda primair en voortgezet onderwijs, die ingaat op diverse aspecten van digitale geletterdheid (een combinatie van mediawijsheid, ICT-basisvaardigheden, *computational thinking* en informatievaardigheden).[45] Dit is onderdeel van een curriculumherziening die onderwijsprofessionals in http://www.curriculum.nu aan het uitwerken zijn.

Zelfregulering en gedragscodes

Het kabinet verwacht van het bedrijfsleven dat het zich houdt aan de *UN Guiding Principles on Business and Human Rights*.[46] Dit betekent onder meer dat bedrijven bij het ontwerpen of aanschaffen van AI-toepassingen zorgvuldig moeten kijken naar mogelijke negatieve effecten op mensenrechten en maatregelen moeten treffen om die effecten te voorkomen of te repareren.
Een instrument dat daarbij kan helpen is zelfregulering. Het kabinet is er groot voorstander van dat bedrijven hier hun eigen verantwoordelijkheid nemen. Zelfregulering is mogelijk met instrumenten als gedragscodes, certificering en bindende bedrijfsvoorschriften. De afgelopen jaren zijn er op nationaal en internationaal niveau verschillende initiatieven tot zelfregulering ontstaan. Tijdens de Conferentie Nederland Digitaal presenteerde Nederland ICT bijvoorbeeld de Ethische Code AI.[47] De code is gebaseerd op de richtsnoeren voor ethische omgang met AI die zijn opgesteld door de Europese High Level Expert Group On Artificial Intelligence (HLEG).[48] Het bevat onder meer principes om publieke waarden te borgen en transparant te zijn over de inzet en het functioneren van AI. Momenteel loopt een pilot fase voor de toepassing van de Europese ethische richtsnoeren waarvan de resultaten begin 2020 zullen worden gepresenteerd.[49] Het kabinet blijft de ontwikkelingen nauwgezet volgen om te bezien of verdere stappen nodig zijn om respect voor mensenrechten door de private sector te garanderen.

Internationale agendering en onderhandeling

Het grensoverschrijdende karakter van AI-ontwikkelingen noopt tot een internationale aanpak. Het kabinet zet zich in het internationale speelveld in voor een duurzame en effectieve borging van mensenrechten. Zo zet Nederland zich in diverse gremia binnen de Raad van Europa actief in voor het ontwikkelen van beleidsinstrumenten die beogen mensenrechten bij de voortschrijdende digitalisering te borgen. Daarnaast werkt Nederland in Europees verband samen aan activiteiten voortvloeiend uit de AI-strategie van de Europese Commissie, vastgelegd in het «Coordinated Action Plan».[50] Dit plan gaat uit van een *human-centred* AI-ontwikkeling waarbij voor mensen belangrijke waarden centraal staan.

---

[44] https://www.nederlanddigitaal.nl/conferentie-nederland-digitaal
[45] Kamerstuk 33 009, nr. 13.
[46] https://www.ohchr.org/documents/publications/GuidingprinciplesBusinesshr_eN.pdf
[47] https://www.nederlandict.nl/ethischecodeai/
[48] De HLEG heeft als algemene doelstelling de uitvoering van de Europese strategie inzake AI te ondersteunen. Dit omvat mede de uitwerking van ethische, juridische en maatschappelijke kwesties met betrekking tot AI.
[49] https://ec.europa.eu/futurium/en/register-piloting-process
[50] https://ec.europa.eu/digital-single-market/en/news/coordinated-plan-artificial-intelligence

Bij de presentatie van de eerdergenoemde ethische richtsnoeren heeft de Europese Commissie een aanvullende mededeling gepresenteerd.[51] Voor de periode 2019–2024 hebben regeringsleiders aangegeven dat de EU beleid moet ontwikkelen dat onze maatschappelijke waarden belichaamt, inclusiviteit bevordert en verenigbaar blijft met onze manier van leven. Gaat het om AI en veiligheid dan benadrukt het kabinet dat multilateraal overleg leidend is. Met name de toepassing van AI in wapensystemen en de ethische vraagstukken die dit tot gevolg heeft, is een belangrijk punt van internationaal overleg. Sinds 2013 staat het op de agenda van de Convention on Certain Conventional Weapons van de Verenigde Naties. Nederland bepleit hierin dat de ontwikkeling en toepassing van wapensystemen te allen tijde onder betekenisvolle menselijke controle moet staan en dat de toepassing ervan enkel kan plaatsvinden in overeenstemming met het internationaal oorlogsrecht.

[51] Kamerstukken 22 112, nr. 2799.