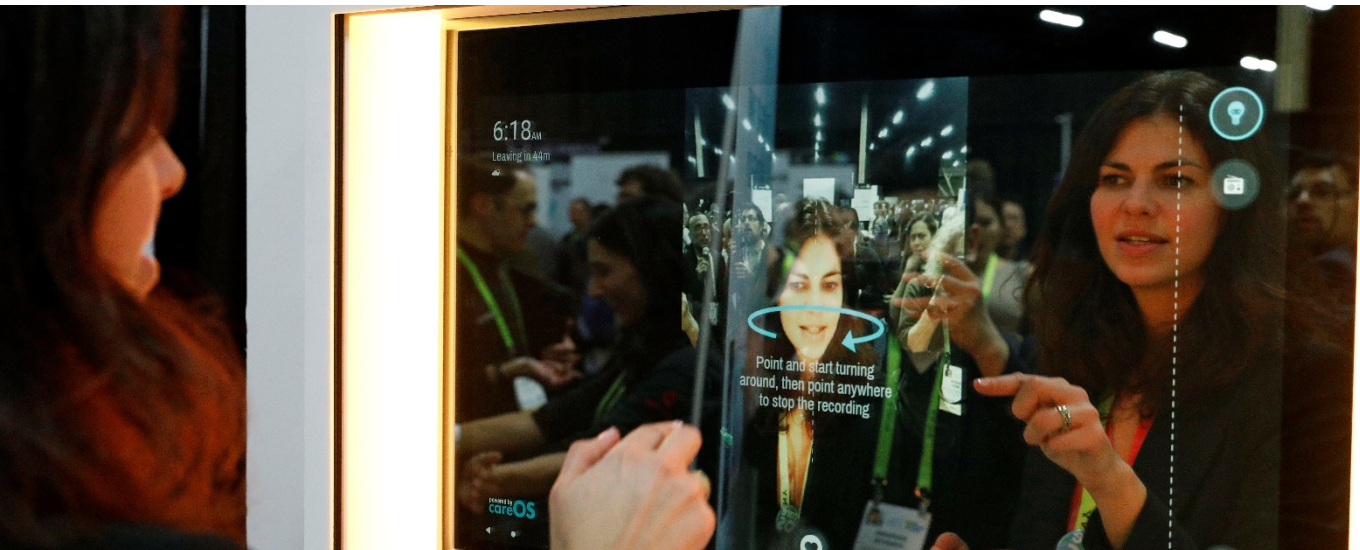


Zeven acties voor verantwoord innoveren met AI



Bericht aan het Parlement

Artificiële Intelligentie (AI) is van toenemend belang voor onze maatschappij en economie. Volgens het kabinet dient de inzet van AI plaats te vinden met inachtneming van publieke waarden en mensenrechten. Dat vraagt om een effectieve toepassing van juridische en ethische kaders in de praktijk van AI. Het Rathenau Instituut stelt 7 acties voor om maatschappelijk verantwoorde AI-innovatie mogelijk te maken.

Inleiding

Artificiële intelligentie (AI) kent al vele toepassingen. Je kunt muziek afspelen die is geselecteerd door slimme algoritmes. De politie gebruikt AI-systemen voor gezichtsherkenning en advies over de inzet van eenheden. Scholen maken steeds meer gebruik van *learning analytics*. En AI-software wordt gecombineerd met virtual reality-toepassingen.

Het kabinet beschouwt AI als een sleuteltechnologie voor de toekomst. Het is een technologie waarmee maatschappelijke opgaven, van zorg en duurzame energie tot veiligheid, beantwoord kunnen worden, en waarmee Nederland geld kan verdienen.

Het kabinet ziet ook dat het gebruik van AI-systemen maatschappelijke uitdagingen oplevert. AI-systemen kunnen namelijk onbedoelde effecten hebben en misbruikt worden, bijvoorbeeld door criminelen. AI-systemen kunnen discrimineren en AI kan de privacy van burgers onnodig aantasten.

Het kabinet wil daarom, net als de Europese Commissie, investeren in ethische AI: AI-systemen die publieke waarden en mensenrechten, zoals privacy en gelijke behandeling, in acht nemen *en* bijdragen aan de economische en maatschappelijke ontwikkeling van Nederland.

Er zijn al goede stappen in de richting van ethische AI gezet. Nieuwe wetgeving, zoals de Algemene Verordening Gegevensbescherming (AVG), is in 2018 in werking getreden. In april 2019 publiceerde de Europese **High Level Expert Group on AI** ethische richtlijnen voor betrouwbare AI. In februari 2020 heeft ook de Europese Commissie haar strategie gepresenteerd, onder meer in het **White Paper On Artificial Intelligence**. Ook diverse bedrijven stelden ethische codes op. In Nederland geeft het kabinet invulling aan ethische AI via onder andere het **Strategisch Actieplan AI** en de Beleidsbrief **AI, publieke waarden en mensenrechten**. Dit bericht aan het parlement omarmt deze initiatieven en geeft handvatten om juridische en ethische kaders in de praktijk te brengen. Hiervoor stellen we **7 acties** voor:

1. Pas proportionaliteit en subsidiariteit toe.
2. Verhelder juridische kaders.
3. Verbied ethisch zeer ongewenste AI-toepassingen.
4. Neem internationale verantwoordelijkheid.
5. Maak beleid voor de gehele datawaardeketen.
6. Betrek de sector bij AI-innovatie.
7. Verminder technologische afhankelijkheid.

1. Pas proportionaliteit en subsidiariteit toe

De principes proportionaliteit en subsidiariteit spelen een belangrijke rol binnen de mensenrechtenrechtspraak en zijn ook verwerkt in de privacywet (AVG). Bij de beoordeling of er sprake is van een inbreuk op mensenrechten vormen ze een belangrijke toets voor de rechter. Een AI-toepassing is proportioneel als de inbreuk die de toepassing maakt op een mensenrecht of publieke waarde in balans is met de maatschappelijke opbrengst van die toepassing. Een AI-toepassing is subsidiair als er geen andere toepassing voorhanden is die hetzelfde resultaat oplevert zonder inbreuk te maken op het mensenrecht of de publieke waarde.

Proportioneel en subsidiair handelen vraagt om zorgvuldige afwegingen. In de praktijk gebeurt dat regelmatig niet. Zo is het SyRI-systeem, dat algoritmes gebruikt om burgers te profileren om fraude op te sporen, onlangs verboden. Volgens de rechter waren het maatschappelijk belang en de inbreuk op de burgerrechten niet met elkaar in balans.

Ook in andere domeinen worden deze afwegingen te weinig gemaakt, bijvoorbeeld in het onderwijs. Zo kan AI in de vorm van *learning analytics* het leertraject van een individuele leerling verbeteren. Tegelijkertijd kan het ook ongelijkheid vergroten, als een systeem geschikter blijkt voor de ene leerling dan voor de andere. Bovendien gebruikt een leerling een AI-systeem vaak individueel achter de computer, wat mogelijk ten koste kan gaan van leren samenwerken. Ten slotte moeten leerlingen in sommige systemen met hun eigen naam inloggen. Dan loopt de leerling het risico dat bepaalde gevoelige data, bijvoorbeeld over een leerprobleem, tot hem of haar herleid kunnen worden, en het leerprobleem hem of haar nog jaren achtervolgt. Het is dus niet vanzelfsprekend dat *learning analytics* voldoet aan principes van proportionaliteit en subsidiariteit. Hier wordt echter nog nauwelijks discussie over gevoerd.

De principes van proportionaliteit en subsidiariteit zijn voor iedere AI-toepassing van belang. Bedrijven, publieke instellingen en overheden dienen daarom deze principes zich zodanig eigen maken, dat ze deze bij iedere beslissing over de ontwikkeling en inzet van AI-systemen meewegen. Ook de Tweede Kamer kan het kabinet vragen naar de onderbouwing van de proportionaliteit en subsidiariteit van AI-systemen, en vragen naar alternatieve opties.

2. Verhelder juridische kaders

Diverse juridische kaders reguleren de inzet van AI-technologie. Het gaat daarbij om mensenrechtenverdragen en de grondwet, en om specifieke wetgeving zoals de AVG, consumentenrecht of mededingingsrecht. Om deze kaders up-to-date te houden, werken het kabinet en de Europese Commissie aan vernieuwing ervan. De AVG is een

van de kaders die recentelijk vernieuwd zijn. Het is belangrijk dat de Tweede Kamer de concrete invulling van de AVG monitort.

Na aanpassing dienen de nieuwe juridische kaders effectief te worden toegepast. Dit zal mede afhangen van de inspanningen van toezichthouders en de rechtspraak. Het vraagt om intensievere samenwerking en onderlinge afstemming tussen toezichthouders. Ook het kabinet kan via beleid sturen op de toepassing van juridische principes. We geven twee voorbeelden, die onder meer spelen in de context van de AVG: de principes van **dataminimalisatie** en **transparantie**.

Dataminimalisatie

De AVG vereist dat overheden en bedrijven inzetten op dataminimalisatie. Dit betekent dat het verzamelen en verwerken van persoonsgegevens gebonden moet zijn aan een bepaald doel (doelbinding). Ook mogen er niet meer gegevens verzameld en verwerkt worden dan noodzakelijk zijn om dat doel te behalen (minimale gegevensverwerking). Dit principe verkleint het risico dat persoonlijke data in de verkeerde handen vallen of gebruikt worden voor oneigenlijke doelstellingen. Ontwikkelaars dringen er echter op aan dat veel data nodig zijn om AI-systemen te trainen. Maar AI-systemen kunnen ook zuiniger met data omgaan. Denk aan toepassingen voor *e-health* op basis van de *Personal Health Train* die geen gegevens van de apparaten van burgers afhalen. Of aan nieuwe AI-ontwikkelingen zoals *federated learning*, waarbij algoritmes leren van decentraal opgeslagen datasets. Het is van belang dat het kabinet deze vormen van AI aanmoedigt. De Tweede Kamer kan het kabinet vragen dataminimalisatie op te nemen als voorwaarde in het eigen inkoopbeleid, bij financieringsopties en het opzetten van *Public-Private Partnerships*.

Transparantie

De AVG geeft aan dat burgers bij geautomatiseerde beslissingen het recht hebben om van organisaties inzage te ontvangen in de categorieën persoonsgegevens die worden verzameld en verwerkt en wat daarmee gebeurt. Sommige organisaties beantwoorden dit soort verzoeken nu door het openbaar maken van ingewikkelde gegevenslijsten waaruit leken weinig kunnen opmaken. Het is van belang dat het kabinet vereist dat partijen die persoonsgegevens gebruiken, helder kunnen aangeven om welke gegevens het gaat en waarvoor ze worden gebruikt.

Transparantie is niet alleen belangrijk bij het geven van inzicht, maar ook bij het uitleggen van hoe (deels) geautomatiseerde besluiten tot stand komen. De Europese Commissie ziet in haar *White Paper On Artificial Intelligence* deze uitlegbaarheid als een van de grootste opgaven. Het kabinet heeft verschillende voorstellen gedaan om de uitlegbaarheid van (deels) geautomatiseerde besluiten te waarborgen. Het is van belang dat de Tweede Kamer de inspanningen van het kabinet op dit punt nauwlettend volgt. Wanneer het uitleggen van besluiten, vanwege de complexiteit van het AI-systeem, momenteel nog niet mogelijk is, dienen alternatieve systemen te worden ingezet.

3. Verbied ethisch zeer ongewenste AI-toepassingen

Sommige AI-systemen brengen zulke grote risico's met zich mee dat het zeer de vraag is of ze ethisch verantwoord zijn. Neem bijvoorbeeld de killer robots. Dit zijn gewapende robots die autonoom menselijke doelwitten kunnen uitkiezen en aanvallen. Er is een sterke internationale beweging die *killer robots* wil verbieden. Het kabinet ondersteunt zo'n moratorium niet, maar vindt dat er sprake moet zijn van 'betekenisvolle menselijke controle'. Dit uitgangspunt wordt uitgewerkt binnen internationale overleggen zoals de *Group of Governmental Experts on Lethal Autonomous Weapons Systems* maar is vooralsnog te onduidelijk om in de praktijk aan te geven wat wel en niet mag. De Tweede Kamer kan aandringen op meer duidelijkheid.

Eenzelfde voorzichtigheid aan de kant van het kabinet is zichtbaar in de discussie over de massale inzet van gezichtsherkenning in de publieke ruimte door de politie of burgers. Deze software kan het persoonlijke leven van burgers ernstig onder druk zetten. Verschillende maatschappelijke organisaties hebben daarom opgeroepen om gezichtsherkenning niet in de publieke ruimte toe te passen. De Tweede Kamer heeft het kabinet verzocht om de verschillende toepassingen van gezichtsherkenning te inventariseren en de wettelijke regels hierover te actualiseren.

Het is belangrijk dat de Tweede Kamer de inspanningen van het kabinet ten aanzien van ethisch zeer riskante AI-toepassingen nauwlettend blijft volgen. Daarbij is het goed in te zetten op Europese en internationale afspraken en afstemming, maar die inzet mag een besluitvaardige houding niet in de weg zitten.

4. Neem internationale verantwoordelijkheid

AI-systemen worden wereldwijd verkocht, en Nederland en Europa maken deel uit van die markt. Het is daarom belangrijk dat het kabinet voorkomt dat Nederlandse en Europese AI-systemen in verkeerde handen vallen. Denk bijvoorbeeld aan buitenlandse actoren die cyberwapens stelen om er aanslagen mee te plegen, of aan autoritaire regimes die dissidenten surveilleren met AI-technologie. Zo heeft het Nederlandse Noldus de *FaceReader*, die wordt ingezet voor emotieherkenning, verkocht aan organisaties in China. Omdat de Chinese overheid een omvangrijk surveillancesysteem heeft opgebouwd om dissidenten in de gaten te houden wordt gevreesd dat dit met dergelijke technologie gebeurt. Het bedrijf Noldus geeft aan dat dit andere technologie betreft, maar heeft niettemin maatregelen genomen om ander dan bedoeld gebruik tegen te gaan. (*geüpdatet 2 februari 2021*)

Het is daarom van belang dat Nederland zich blijft inzetten om binnen Europa de regels aan te scherpen voor de ontwikkeling van en handel in riskante AI-systemen. Als deze aanscherping, bijvoorbeeld ten aanzien van de *dual use*-lijst, onmogelijk blijkt, is het zaak de nationale reguleringsopties grondig te verkennen. Belangrijk is ook om in het

wetenschappelijk onderzoek nieuwe afwegingskaders op te nemen over het delen van riskante technologie met het buitenland of met buitenlandse studenten.

De Tweede Kamer kan op deze punten de inzet van het kabinet monitoren. Ten aanzien van het groeiende internationale cyberconflict is het van bijzonder belang dat de Tweede Kamer zo veel mogelijk transparantie vraagt over de Nederlandse inzet, en hierover parlementair debat voert.

5. Maak beleid voor de gehele datawaardenketen

Veel actoren zijn bezig met het verzamelen, analyseren en toepassen van data. In die zogeheten datawaardenketen zitten de burgers van wie de data worden verzameld, ontwikkelaars van AI-technologie zoals universiteiten en bedrijven, overheden en partijen die producten en diensten aanbieden op bijvoorbeeld Facebook of Android-systemen en gebruik maken van persoonlijke gegevens.

Het is van belang om de juridische en ethische kaders op al deze partijen toe te passen. Het is goed dat verschillende bedrijven zich aan ethische AI hebben gecommitteerd maar een wirwar van verschillende ethische codes is niet wenselijk. Er is behoefte aan gezamenlijke actie, ook op internationaal niveau. Organisaties als ISO (International Organisation for Standardisation) en UNESCO werken daarom aan collectieve standaarden.

Ethische AI moet onderdeel worden van het DNA van bedrijven en deel uitmaken van maatschappelijk verantwoord ondernemen (mvo). Dit betekent dat bedrijven niet alleen afspraken maken voor het eigen bedrijf of de eigen branche, maar ook verantwoordelijkheid nemen voor alle partijen in de productieketen, zoals buitenlandse toeleveranciers. De OESO speelt een belangrijke rol in de concrete uitwerking van juridische en ethische kaders op dit vlak en gaat zich nu ook richten op maatschappelijk verantwoorde datastromen.

Ten slotte is het van belang dat de gehele datawaardenketen zo min mogelijk in het bezit komt van één of enkele partijen. De afgelopen jaren is gebleken hoe moeilijk dit is te voorkomen. Bedrijven als Google en Facebook hebben op verschillende terreinen een enorm marktaandeel opgebouwd. Zij beschikken nu over zoveel gebruikers, data en geld dat nieuwe bedrijven lastig met hen kunnen concurreren. Aanscherping van het mededingingsrecht is daarom nodig. Nederland heeft op dit punt veranderingen voorgesteld, bijvoorbeeld ten aanzien van het van tevoren (*ex ante*) nemen van mededingingsmaatregelen en de aanpassing van fusiedrempels. Het is nu zaak deze beleidsvoornemens op Europees niveau in wetgeving om te zetten.

De Tweede Kamer kan de voortgang op het gebied van mededingingsrecht en de standaardisering van codes monitoren.

6. Betrek de sector bij AI-innovatie

We schreven het hierboven al: AI speelt een rol in uiteenlopende maatschappelijke praktijken. Er is daarom een sectorspecifieke aanpak nodig om juridische en ethische kaders in de praktijk te brengen. De medische wereld, het onderwijs en de veiligheidssector hebben ieder hun eigen regelgeving en werkprocessen die up-to-date gebracht moeten worden. Bovendien blijkt uit Rathenau-onderzoek naar missiegedreven innovatiebeleid dat het cruciaal is om als overheid te profiteren van de kennis die dokters, politieagenten en leraren al hebben over het inpassen van technologie in hun werk. Als ze goed worden gecoördineerd en ingebed, kunnen sectorspecifieke, lokale *living labs* oplossingen bieden van hoge kwaliteit, waarbij zorgvuldig onderzocht wordt wat bepaalde ethische, juridische, technische en sociale aspecten precies vereisen in de praktijk.

Uit onderzoek van het Rathenau Instituut naar missiegedreven innovatie blijken focus en kleinschaligheid, waarbij verantwoordelijke partijen samenwerken in testomgevingen, de succesfactoren. Dat geldt zeker voor (semi)publieke sectoren als de zorg en het onderwijs. Deze inzet zal ook vragen om nieuwe beleidsinstrumenten, zoals kleine fondsen die aantrekkelijk zijn voor sociale ondernemingen. De Tweede Kamer kan het kabinet vragen om deze beleidsinstrumenten te ontwikkelen.

Ten slotte kan het maatschappelijk middenveld een belangrijke rol spelen. Juist met hulp van beroepsverenigingen en consumentenorganisaties kunnen lokale successen een grotere maatschappelijke impact hebben.

7. Verminder technologische afhankelijkheid

De energiecrises uit de jaren 70 lieten zien dat Europa te afhankelijk was van olie uit het Midden-Oosten. Nu groeien de zorgen over de geringe onafhankelijkheid van Europa op het gebied van data en AI. De digitale reuzenbedrijven zoals Google en Huawei zijn vrijwel allemaal Amerikaans of Chinees en hebben grote macht op het gebied van data en AI. Deze digitaal-technologische afhankelijkheid vermindert de controle van Europa over haar eigen digitale toekomst.

Het is daarom hoog tijd om de technologische afhankelijkheid van Nederland en Europa te verminderen. Dat vraagt om Europese bedrijvigheid. Overheden kunnen die bedrijvigheid stimuleren door een goed gecoördineerd en intensief innovatiebeleid te voeren. We schreven hierboven al dat het belangrijk is om daarbij te profiteren van sectorale kennis. Vervolgens is het nodig om lokale experimenten in te bedden in bredere, gefocuste programma's, en de successen op nationaal en Europees niveau te delen. Maar ook binnen overheden dient de kennis en kunde over omgang met data en AI aanzienlijk versterkt te worden.

Meer technologische onafhankelijkheid kan ook de ontwikkeling van ethische AI versterken, iets dat als het fundament van de Nederlandse en Europese economische ontwikkeling wordt gezien. Bedrijven die binnen Europa actief zijn, hebben zich aan onze juridische en ethische standaarden te houden. In een wereld waarin Chinese en Amerikaanse AI-systemen te vaak ten kwade worden gebruikt, vormt hoogwaardige ethische AI een unieke Europese propositie. De introductie van de AVG toonde al dat als Europa het juiste voorbeeld laat zien, de rest van de wereld kan volgen.

Kader 1 **Relevante publicaties van het Rathenau Instituut**

Voorbij lokaal enthousiasme; Lessen voor de opschaling van living labs (rapport, 2020)

Gezondheid centraal; Zorgvuldig data delen in de digitale samenleving (rapport, 2019)

Missiegericht innovatiebeleid in uitvoering (bericht aan het parlement, 2019)

Zo brengen we AI in de praktijk vanuit Europese waarden (artikel, 2019)

Cyberspace zonder Conflict; De zoektocht naar de-escalatie van het internationale informatieconflict (rapport 2019)

Kennis in het vizier; De gevolgen van de digitale wapenwedloop voor de publieke kennisinfrastructuur (rapport, 2019)

Bijdrage rondetafelgesprek Digitalisering in het hoger onderwijs (bericht aan het parlement, 2019)

Uitdagingen voor de regulering van drones en killer robots (nieuws, 2019)

Doelgericht digitaliseren; Hoe Nederland werkt aan een digitale transitie waarin mensen en waarden centraal staan (rapport, 2018)

Waardevol digitaliseren; Hoe lokale bestuurders vanuit publiek perspectief mee kunnen doen aan het 'technologiespel' (rapport, 2018)

Coördinatie van publiek-privaat onderzoek (rapport, 2013)
